Kostas Pentikousis
Ramón Agüero
Marta García-Arranz
Symeon Papavassiliou (Eds.)

68

Mobile Networks and Management

LNICST

Second International ICST Conference, MONAMI 2010
Santander, Spain, September 2010
Revised Selected Papers

Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 68

Kostas Pentikousis
Ramón Agüero
Marta García-Arranz
Symeon Papavassiliou (Eds.)

Mobile Networks and Management

Second International ICST Conference, MONAMI 2010
Santander, Spain, September 22-24, 2010
Revised Selected Papers

Springer

Volume Editors

Kostas Pentikousis
Huawei Technologies Düsseldorf GmbH
European Research Centre, 10587 Berlin, Germany
E-mail: k.pentikousis@huawei.com

Ramón Agüero
Network Planning and Mobile Communications Laboratory
Department of Communications Engineering
University of Cantabria, 39005 Santander, Spain
E-mail: ramon@tlmat.unican.es

Marta García-Arranz
Network Planning and Mobile Communications Laboratory
Department of Communications Engineering
University of Cantabria, 39005 Santander, Spain
E-mail: marta@tlmat.unican.es

Symeon Papavassiliou
Network Management and Optimal Design Laboratory
School of Electrical and Computer Engineering
National Technical University of Athens, 15780 Athens, Greece
E-mail: papavass@mail.ntua.gr

ISSN 1867-8211 e-ISSN 1867-822X
ISBN 978-3-642-21443-1 e-ISBN 978-3-642-21444-8
DOI 10.1007/978-3-642-21444-8

Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011928827

CR Subject Classification (1998): C.2, H.4, D.2, H.3, H.5, I.2

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume is the result of the Second International ICST Conference on Mobile Networks and Management (MONAMI), which was held in Santander, Spain, during 22–24 September 2010, hosted by the University of Cantabria. Santander, a modern cosmopolitan city with a rich historical past, appealing social and cultural background, and high-quality service offerings, was the perfect backdrop for the second edition of the conference.

MONAMI aims at closing the gap between hitherto considered separate and isolated research areas, namely, multiaccess and resource management, mobility management, and network management. Although these have emerged as core aspects in the design, deployment, and operation of current and future networks, there is still little to no interaction between the experts in these fields. MONAMI enables cross-pollination between these areas by bringing together top researchers, academics, and practitioners specializing in the area of mobile network and service management. MONAMI 2010 more than doubled in terms of size and attendance when compared to the first edition in Athens, but remained a focused event. We are happy to announce that the third edition of MONAMI will be hosted by the University of Aveiro in September 2011. Our goal is to establish the conference over the years as a reference point for the research community.

The technical program featured 29 papers, which were selected after a thorough peer-review process based on their relevance to the scope of the conference and their technical merit. Thirty-six Technical Program Committee members made sure that each submitted paper was reviewed by at least three competent researchers. This volume is organized by subject in seven parts as follows. Papers 1 to 4 address "Routing and Virtualization" in Part I. "Autonomic Networking" aspects are discussed in Part II (papers 5-8). Papers 9 through 16 present new approaches for "Mobility Management" (Part III) and "Multiaccess Selection" (Part IV). Papers 17 to 24 consider "Wireless Network Management" and "Wireless Networks," in Parts V and VI, respectively. Finally Part VII includes five papers presenting "Future Research Directions." All papers were orally presented in a single-track format, with sufficient time allocated for discussion following each presentation, fostering active participation by all attendees.

The conference opened with a full-day tutorial on "Machine-to-Machine (M2M) Communication" by domain experts Mischa Dohler and Jesús Alonso-Zárate of CTTC. Joan Serrat of Universitat Politècnica de Catalunya opened the second day of the conference with a keynote on "Service Management in Future Networks: The C3SEM Vision." Last but not least, José Manuel Hernández of Telefónica R+D gave the keynote on the third day of the conference entitled "SmartCities, the Silent IoT Revolution."

We acknowledge the vital role that the Technical Program Committee members and additional referees played during the review process. Their efforts ensured that all submitted papers received a proper evaluation. We thank CreateNet for technically co-sponsoring the event and the University of Cantabria for hosting MONAMI 2010 as well as for providing organizational, logistics, and media support. Finally, we thank all delegates for attending MONAMI 2010 and making it such a vibrant conference!

December 2010

Kostas Pentikousis
Ramón Agüero
Marta García-Arranz

Executive Committee

Organizing Committee

General Chairs

Kostas Pentikousis	Huawei Technologies European Research Center, Germany
Ramón Agüero	University of Cantabria, Spain

Technical Program Committee Co-chairs

Marta García-Arranz	University of Cantabria, Spain
Symeon Papavasileiou	National Technical University of Athens, Greece

Tutorial Chair

Oliver Blume	Alcatel-Lucent Bell Labs, Germany

Publication Chair

Carlo Giannelli	University of Bologna, Italy

Publicity Chairs

Jarno Pinola	VTT Technical Research Centre of Finland, Finland
Carlo Giannelli	University of Bologna, Italy

Steering Committee

Co-chairs

Imrich Chlamtac	CREATE-NET, Italy
Kostas Pentikousis	Huawei Technologies European Research Center, Germany

Technical Program Committee

Ramón Agüero	University of Cantabria, Spain
Rui Aguiar	University of Aveiro, Portugal
Toufik Ahmed	University of Bordeaux 1, France
Nadeem Akhtar	Centre of Excellence in Wireless Technology (CEWiT), India
Faouzi Bader	CTTC, Spain
Hussein Badr	Stony Brook University, USA
Roland Bless	Karlsruhe Institute of Technology, Germany
Oliver Blume	Alcatel-Lucent Bell Labs, Germany
Mischa Dohler	CTTC, Spain
Stephen Farrell	Trinity College Dublin, Ireland
Marta García-Arranz	University of Cantabria, Spain
Raffaele Giaffreda	CREATE-NET, Italy
Carlo Giannelli	University of Bologna, Italy
Tobias Hossfeld	University of Wüerzburg, Germany
Jussi Kangasharju	University of Helsinki, Finland
Theo G. Kanter	Mid-Sweden University, Sweden
Vasileios Karyotis	National Technical University of Athens, Greece
Timotheos Kastrinogiannis	National Technical University of Athens, Greece
Polychronis Koutsakis	Technical University of Crete, Greece
Zhaojun Li	Fujitsu Labs, UK
Emmanuel Lochin	ISAE - LAAS-CNRS, France
Raquel Morera	Telcordia Technologies, USA
Antonio de la Oliva	Carlos III University, Spain
Symeon Papavassiliou	National Technical University of Athens, Greece
Kostas Pentikousis	Huawei Technologies European Research Center, Germany
Miguel Ponce de Leon	Waterford Institute of Technology, Ireland
Antonio Puliafito	University of Messina, Italy
Anand R. Prasad	NEC Corporation, Japan
Javier Rubio-Loyola	National Polytechnic Institute, Mexico
Susana Sargento	University of Aveiro, Portugal
Peter Schoo	Fraunhofer-SIT, Germany
Joan Serrat	Polytechnic University of Catalonia, Spain
Haitao Tang	Nokia Siemens Networks, Finland
Andreas Timm-Giel	Hamburg University of Technology, Germany
Kurt Tutschku	University of Vienna, Austria
Christos Verikoukis	CTTC, Spain

Additional Reviewers

Georgios Aristomenopoulos	National Technical University of Athens, Greece
Nicola Baldo	CTTC, Spain
Javier Baliosian	Universidad de la Republica, Uruguay
Roksana Boreli	NICTA, Australia
Dario Bruneo	Universita' di Messina, Italy
Paul Bucknell	Fujitsu, United Kingdom
Antonio Celesti	University of Messina, Italy
Johnny Choque	University of Cantabria, Spain
Marta Domingo	University of Cantabria, Spain
Eirini Eleni Tsiropoulou	National Technical University of Athens, Germany
Alberto Eloy Garcia Gutierrez	University of Cantabria, Spain
Shane Fox	TSSG Waterford Institute of Technology, Ireland
Jochen Furthmüller	Karlsruhe Institute of Technology, Germany
Jens Gebert	Alcatel-Lucent Bell Labs, Germany
Michael Georgiades	Cyprus University of Technology, Cyprus
Victor Gil	University Carlos III, Spain
Diogo Gomes	Instituto de Telecomunicacoes - University of Aveiro, Portugal
Mary Grammatikou	National Technical University of Athens, Greece
Christian Haas	Karlsruhe Institute of Technology, Germany
Matthias Hartmann	University of Würzburg, Germany
Mythri Hunukumbure	Fujitsu, United Kingdom
Vitor Jesus	Instituto de Telecomunicacoes - University of Aveiro, Portugal
Fernando Jose Velez	IT-DEM, Universidade da Beira Interior, Portugal
Ahmed Kazi	Pennsylvania State University, USA
Ivan Lopez-Arevalo	Cinvestav Tamaulipas, Mexico
Aarne Mammela	VTT Technical Research Centre of Finland, Finland
Jaume Nin	CTTC, Spain
Mikko Pervilä	University of Helsinki, Finland
Henrik Petander	NICTA, Australia
Manuel Ricardo	INESC Porto, Portugal
Luis Sanchez	University of Cantabria, Spain
Roberto Sanz	University of Cantabria, Spain
Christian Schwartz	University of Vienna, Austria
Eleni Stai	National Technical University of Athens, Greece
Alessandra Toninelli	University of Bologna, Italy
Fatima Zarinni	Stony Brook University, USA

Table of Contents

Part I: Routing and Virtualization

Part II: Autonomic Networking

Part III: Mobility Management

Part IV: Multiaccess Selection

Part V: Wireless Network Management

Part VI: Wireless Networks

Part VII: Future Research Directions

Generic Connectivity Architecture for Mobility and Multipath Flow Management in the Future Internet

Amanpreet Singh[1], Christoph Nass[1], Andreas Timm-Giel[2], Peter Schefczik[3], Horst Roessler[3], and Michael Scharf[3]

[1] Center for Computing and Communication Technologies (TZI), University of Bremen, Germany
[2] Institute of Communication Networks, Hamburg University of Technology, Germany
[3] Bell Labs Germany, Alcatel-Lucent, Germany
{aps,chn}@comnets.uni-bremen.de, timm-giel@tuhh.de, {peter.schefczik,horst.roessler, michael.scharf}@alcatel-lucent.com

Abstract. With the evolution of the Internet, the vast majority of the traffic is generated by information-centric applications, which would benefit from enhanced data transport paradigms. This paper presents the development and implementation of the Generic Connectivity architecture, a new communication flow abstraction that is based on the Generic Path architecture developed within the European Research Project 4WARD. The Generic Connectivity mechanisms allow for a high degree of flexibility by covering both existing and new protocol paradigms, which are particularly beneficial for wireless access networks. This paper shows that the Generic Connectivity architecture can realize new network mechanisms beyond the features of the current Internet protocol architecture. It is thus a promising clean-slate approach for the Future Internet. The relevant aspects of the architecture are implemented with the OMNET++ 4.0 network simulation tool. Using simulations, the advantages of the Generic Connectivity architecture are shown for several new use cases, including an adaptive protocol selection, mobility, multicast and multipath connectivity over heterogeneous wireless networks. Furthermore, it is also demonstrated that the architecture inherently supports guaranteed Quality-of-Service (QoS) agreements and traffic distribution over dynamic channels.

Keywords: Future Internet, Mobility, Multipath, Quality of Service.

1 Introduction

The current Internet architecture is challenged by the rapid growth of the number of nodes and the increasing traffic volume, as well as the fact that more and more content is accessed via mobile devices. These are major issues for the evolution towards the Future Internet [1].

The 4WARD project [2] addresses these challenges by a "clean-slate" architectural approach. 4WARD provides a flexible framework that allows a number of different networks to co-exist and inter-operate by realizing multiple virtual networks, e.g.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 1–13, 2011.

information-centric networks. Moreover “default-on” management capabilities within the network are incorporated and the network path is made an active element called the "Generic Path".

A Generic Connection is a new networking communication abstraction. It is set up between two or more communicating end-points and organizes the cooperation between nodes for a wide range of communication services. Unlike the existing Internet architecture, the Generic Connection inherently integrates mobility, multipath transport, multicast support, as well as QoS mechanisms.

This paper shows the benefits of the Generic Connectivity architecture and investigates how it could be implemented in a Future Internet, using the 4WARD concepts as a basis. As a proof-of-concept, a discrete event simulator implementation of the newly developed architecture is presented. The key advantages of the new architecture are demonstrated by experiments with adaptive error and flow control and multipath traffic distribution over heterogeneous wireless access networks.

The rest of this paper is structured as follows: Related ongoing work in Future Internet research is briefly summarized in section 2. The Generic Connectivity architecture is presented in section 3. The flexibility of the Generic Connectivity architecture can be seen by applying concepts of mobility, multipath and multicast in section 4. The simulation tool implementation and evaluations of different scenarios are depicted in section 5 and 6, respectively. Finally, section 7 concludes the paper and provides an outlook.

2 State of Art

There are numerous ongoing research activities for designing the Future Internet, e.g., the NSF Future Internet Design (FIND) program and the Global Environment for Network Innovations (GENI) platform in the US. The latter’s purpose will be to implement and test a wide range of research proposals in distributed global testbeds.

Sensor and mobile wireless networks are a key challenge for Future Internet design. This has also triggered research activities on fundamentally new protocol architectures, for instance in the European 4WARD project. The 4WARD approach to mobility is described in [3].

A clean slate approach combining routing with content data was triggered by Van Jacobsen and others [4]. The content centric network (CCN) proposal makes mobility management for certain services easier by putting content Ids into the forwarding table of the routers and making the content itself move. A related approach is targeted by network information objects within the 4WARD project.

Both trends result in a need for more flexible data transport mechanisms. This requirement has also been identified in [5], and is addressed there by introducing a separate flow layer and factoring endpoint addressing into a separate endpoint layer.

3 Generic Connectivity Architecture

A generic architecture for communication paths in the Future Internet must support the growing diversity of applications and network technologies, allow multiple points

of attachment and maintain seamless connectivity for mobile hosts and networks. Due to the varied requirements, it is not possible to envision a single transport solution but a family of communication paradigms differing in their characteristics and types.

The Generic Path (GP) framework developed in the 4WARD project abstracts and generalizes a number of transport connections from physical wires, wireless mediums, optical fibers and virtual connections. In addition to the data transfer capabilities, the GP architecture inherently allows for data transformation such as aggregation, encapsulation, encryption, translation, coding and transcoding [6].

3.1 Overview and Terminology

The Generic Connectivity (GC) architecture extends the object-oriented programming technique of inheritance (consisting of base classes, methods and procedures) proposed in the 4WARD project. Similar to the GP architecture, the GC architecture also enables modular GC services, allowing for a recursive architecture where complex or advanced (higher level) GC services can be obtained from simple and basic GC services. The different objects of the Generic Connectivity architecture are shown in Fig. 1 and are explained in the following:

Information Object (InfObj) - InfObj is a client of a Generic Connection, capable of consuming/producing information.

Entity (ENT) - An Entity may be a process, a thread in a process or a set of processes. It can communicate with other Entities of the same compartment by means of a Generic Connection.

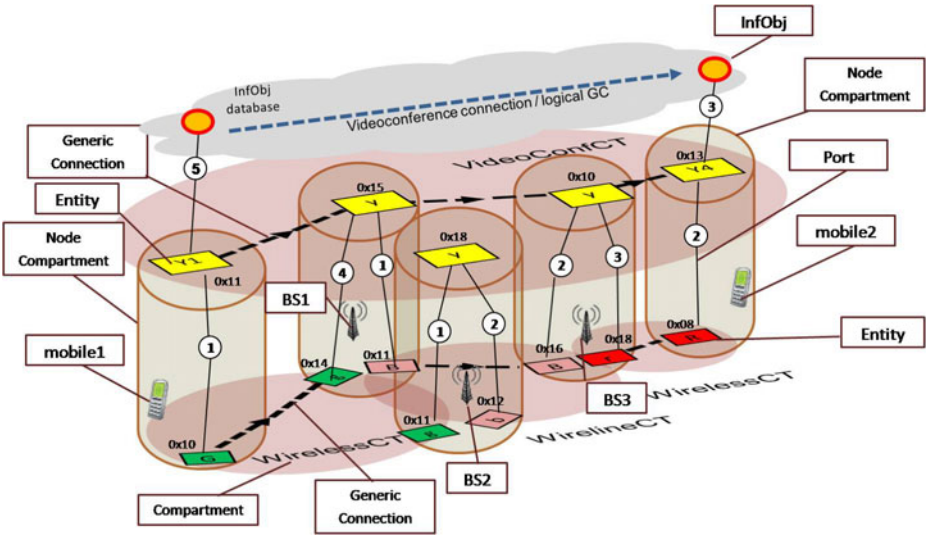


Fig. 1. Simplistic Generic Connection Scenario

Compartment (CT) - A compartment has its own set of protocols for operations such as routing, forwarding, authentication, security, authorization, etc. and may even follow strict policy enforcements. In addition, addressing is unique within the scope

of a compartment and therefore a compartment can be seen as a 'name space'. Like the Generic Connection, different types of Compartments are described by inheritance in the object-oriented model.

Node Compartment (NodeCT) - A node compartment is composed of the software capable of supporting tasks that can all atomically reference the same memory space [6]. A node compartment manages a name space to address Entities. A Node Compartment can relate to a processing system as defined in the Network Inter Process Communication Architecture (NIPCA) of [7].

Ports - A Port is a handle to an Entity, an identifier local to the Node Compartment.

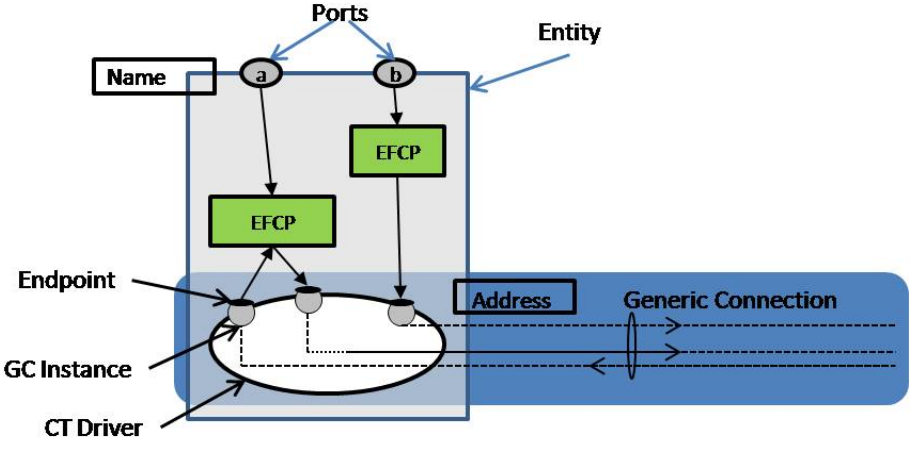


Fig. 2. The Entity Internal Structure

While the GP architecture [6] does not describe node internal structures, the following additional structures were defined within the GC architecture for a real implementation (refer to Fig. 2):

Generic Connection Instance (GC Instance) - Data transfer including forwarding and manipulation of data is executed by GC Instances. A GC Instance keeps the local state information of a GC. It is a thread or process executing a data transfer protocol machine. GC Instances are created by an Entity and may access shared information of that Entity.

Error and Flow Control Protocol (EFCP) - EFCP takes care of the reliable transmission of the messages over the Generic Connection.

Endpoint (EP) - A Generic Connection terminates at an Endpoint within an Entity. The EFCP injects or retrieves packets from the GC via the Endpoint.

Compartment Driver (CT Driver) - CT Driver identifies the Entity within a compartment. An Entity can only join a single Compartment at a time.

3.2 Naming and Addressing

In the Generic Connectivity architecture framework, the entities are assigned names and compartment specific addresses. To resolve an entity, the name resolution is

done. The entity always keeps its name even if it moves and joins a new compartment, in which case the same entity will have a new address. Within a Node Compartment also the different entities have to be identified by addresses or some other forms of identifiers. In addition, there is also a need to uniquely identify a port with port numbers.

3.3 Generic Connection Setup

To account for the design goals along with the optimal combinations of the communication protocols, GC services are compartmentalized [6]. Therefore, when an entity wants to set up a Generic Connection in order to initiate communication, it needs to be a member of the appropriate compartment. In order to do so, the entity gets the compartment information from the Node Compartment, searches for the specific compartment or creates the compartment itself and advertises it. In addition to being a namespace, the compartment can be seen as a signaling control plane which provides a topological view and specifies the rules to be followed by its members. Once the entity is a member of the compartment, a Generic Connection can be established as the compartment can obtain the resource information from the underlying and/or neighboring compartments.

4 Mobility, Multipath and Multicast in the Generic Connectivity Architecture

Mobility can be classified into different types like device mobility, network mobility, session mobility, etc. The Generic Connectivity framework can be extended by specialized mobility management mechanisms with respect to the specific characteristics of a compartment. In the following, the mobility, multipath and multicast concepts of the Generic Connectivity architecture are summarized.

The *mobility* solution is shown in Fig. 3. Therein, a Provider compartment is distributed over wireless, access and IP compartments. The mobile device, assumed to be capable on supporting parallel wireless connections, is first a member of the wirelessCT1 and as it moves it appears in the vicinity of another compartment wirelessCT2 it will have to perform a handover from wirelessCT1 to wirelessCT2. Since wirelessCT2 is a new compartment, the entity in the Provider compartment instantiates a new entity for the wirelessCT2 to set up a Generic Connection in the wirelessCT2. The Generic Connection in the Provider compartment is unaware as it will not identify anything whereas the Provider compartment is aware that it now has more resources and more options. In order to use the 2nd available path, a new Generic Connection needs to be established within the access and IP compartments. This *mobility* solution of utilizing multiple paths is of the form of “make-before-break” mobility management.

For uplink traffic, the handover is realized in the GC architecture by switching ports to the wireless compartments. Different from today’s networks, this switching uses the same mechanisms for a handover between two wireless compartments with the same technology or for an inter-technology handover. For downlink traffic, the optimal point from where the downlink traffic is diverted to the new entity needs to be identified.

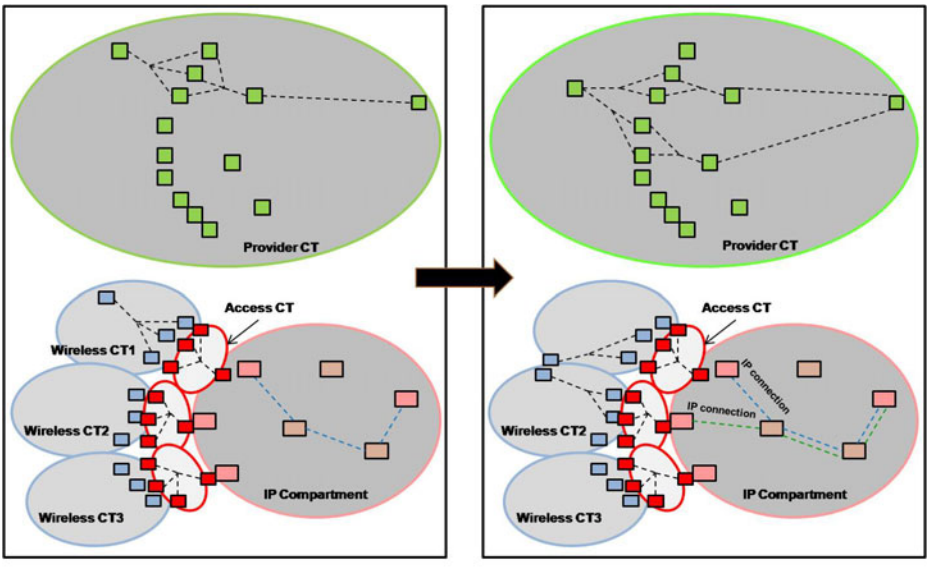


Fig. 3. Mobility – Compartment View

One of the novel features provided by the Generic Connectivity framework is the *multipath* transport. The Generic Connection is spread over multiple links to form the end-to-end transport connection within a compartment. While establishing the Generic Connection, it may request multipath transport from the compartment, if available. In doing so, the Generic Connection can choose amongst the various combinations that exist for multiple paths.

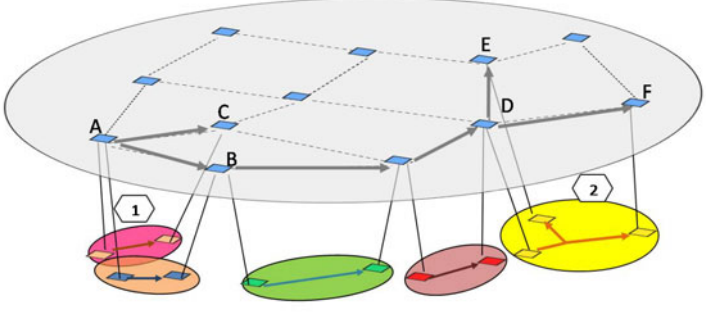


Fig. 4. Multicast – Compartment View

In Fig. 4, an example of a *multicast* Generic Connection is illustrated. The interesting aspect is that the Generic Connection in the higher layer compartment can span over multiple lower layer compartment types. On the left hand side, the

realization of multicast requires that the entity *A* has to duplicate the packet and send them out on both ports. In contrast, on the right hand side, the underlying compartment is a broadcast compartment or already contains a multicast Generic Connection. Hence, the entity *D* only has to send the packet to a single port and the lower compartment will forward it to both entity *E* and *F*.

5 Simulation Tool Implementation

In order to show the feasibility of the developed Generic Connectivity concept, a demonstrator based on the network simulation tool OMNeT++ 4.0 [7] was developed. It implements the GC concept and methods for testing and validating the characteristics of the Generic Connectivity architecture.

5.1 Simulated Scenario

Fig. 5 depicts the network topology that was simulated, including a video conferencing scenario. It is assumed that the mobile1 is within the range of both base stations bs1 and bs3. The connectivity between the mobile1 and the two base stations bs1 and bs3 is represented by link-A and link-B, respectively.

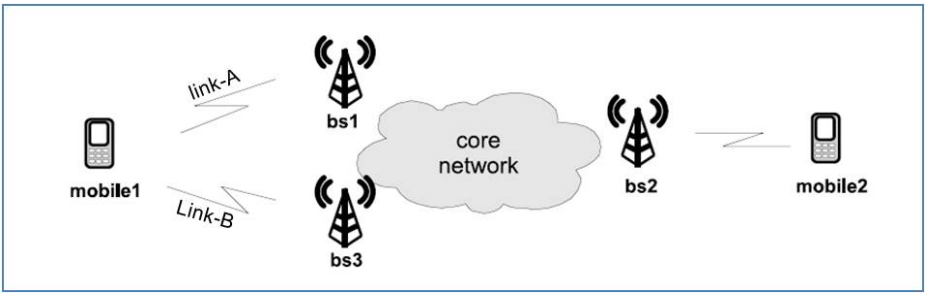


Fig. 5. Simulated Network Topology

5.2 QoS Constrained Multipath Approach

To demonstrate the flexibility of the Generic Connectivity architecture, a heuristic multipath approach for two paths is implemented, where a minimum bandwidth usage policy is used to distribute the traffic with respect to the link quality (packet loss probability) to achieve a pre-defined QoS constraint (packet loss ratio, PLR). To achieve the target PLR, a simple mechanism of duplicating traffic is used. The minimum bandwidth will be used if all traffic can be sent without any duplication. If the link quality is not good enough data duplication is performed. This heuristic algorithm always exploits the best quality path (with least packet losses) to the maximum. If the traffic exceeds the available bandwidth on the best path, the remaining traffic is sent over the second best path.

5.3 Adaptive Error and Flow Control Selection

In the current Internet transport mechanisms, the Error and Flow Control Protocol (EFCP) is integrated into the Transmission Control Protocol (TCP) and cannot be changed with the changing network dynamics. Due to the modular design, the Generic Connectivity framework is much more flexible. This flexibility allows having an adaptive EFCP mechanism within a Generic Connection entity. For the simulations, a *Simple* EFCP module and a *Stop-and-Wait (SnW)* EFCP module were implemented. In Fig. 6, the wireless entities at the mobile1 and base station bs1 (internal-view) are presented. In order to be able to adaptively change the EFCP model, the management module in the entity has to monitor the performance on the link and take appropriate actions depending on the decision algorithm. Once the mobile1 decides to switch between the EFCPs, eventually it has to inform the receiving base station bs1 using a management Generic Connection.

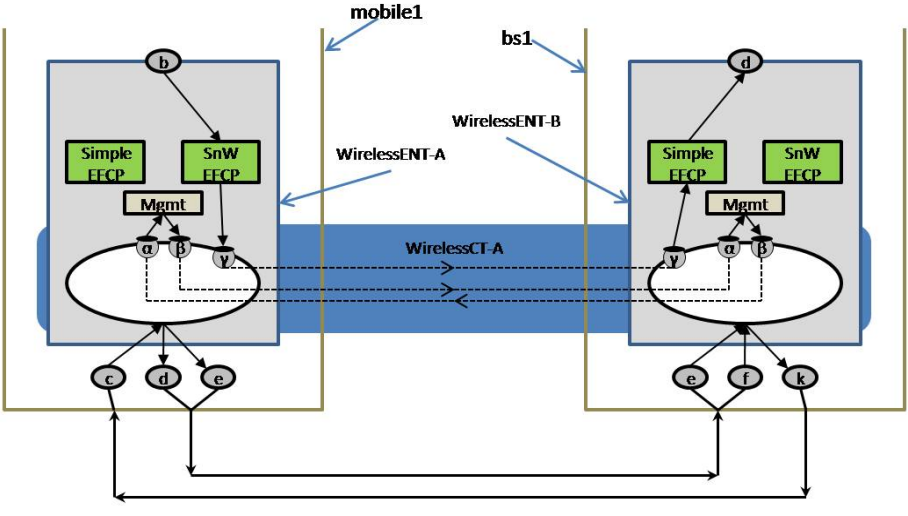


Fig. 6. Wireless Entity view for Adaptive EFCP

5.4 Dynamic Flow Management

For the video conference application considered in the simulated scenario there are two streams – audio and video, as shown in Fig. 7. Within the Generic Connectivity framework, these two data streams can be handled separately by having multiple Generic Connections established for them, even though they are part of the same application data stream. This is another feature that cannot be easily realized in the current Internet.

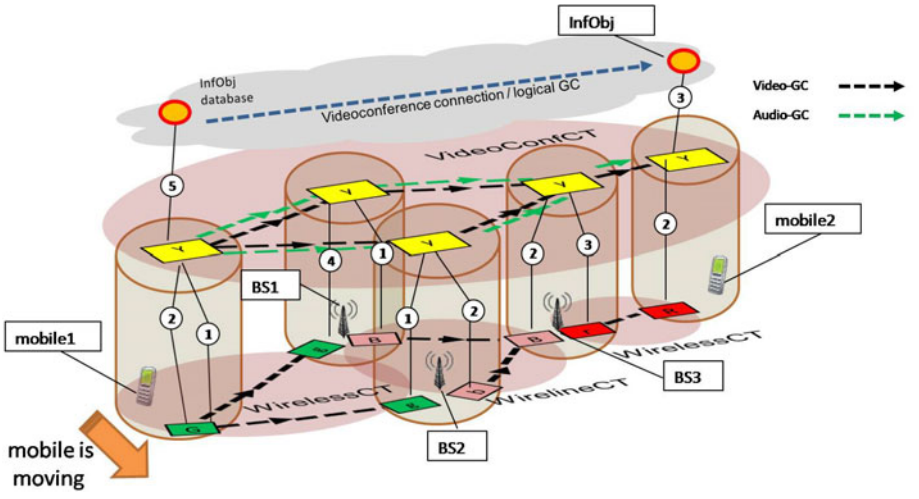


Fig. 7. Audio & Video Multipath Generic Connectivity Management

6 Simulation Results

The following sections present example results for the novel features of the GC architectures and its benefits compared to the existing Internet protocol stack.

6.1 Adaptive Error Correction Mechanisms

For the adaptive EFCP simulations, a packet loss ratio threshold of 0.05 was selected. When the calculated packet loss ratio (depicted in Fig. 8) was less than the set threshold, *Simple* EFCP was used, otherwise *Stop-and-Wait* EFCP was used. The overall packet loss ratio obtained for the adaptive EFCP is presented in Fig. 9and it can be seen that the performance is consistent over the entire simulation period even though the channel quality on link-A was dynamically changing (Fig. 8). In contrast, TCP would always retransmit data, even if this is not required.

6.2 Flexible Multipath Flow Management

In the following, the GC multipath features are illustrated. Fig. 10 depicts the assumed variable packet loss probability of the two links link-A and link-B as mobile1 moves in the left hand side wireless compartment (Fig. 7).

The audio traffic has a data rate of 100 kbit/s and the packet size is 1 kbyte. The target packet loss ratio to be achieved for the audio traffic is set to be 0.025. On the other hand, the video traffic data rate is 1 Mbit/s with the packet size being 2 kbyte and a target packet loss ratio of 0.04. The resources allocated to the audio Generic Connection are 150kbit/s on link-A and also 150kbit/s on link-B. The rest of the resources are allocated to the video Generic Connection i.e., 1.05Mbit/s on link-A and 1.25Mbit/s on link-B.

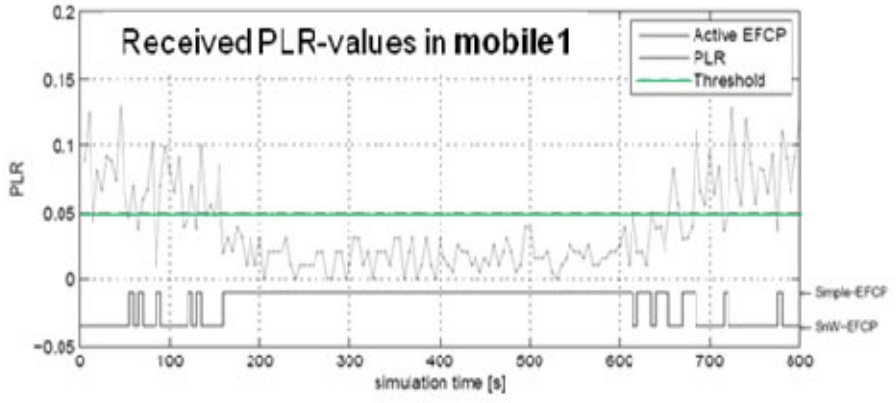


Fig. 8. Adaptive EFCP Switching

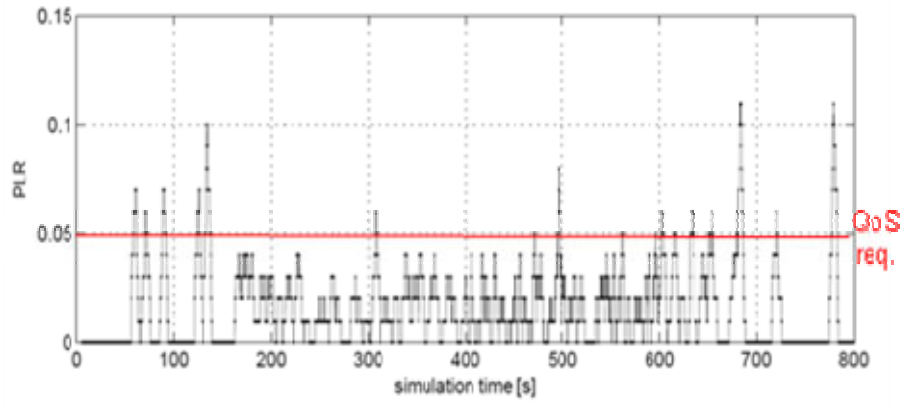


Fig. 9. Packet Loss Ratio seen by the Generic Connection

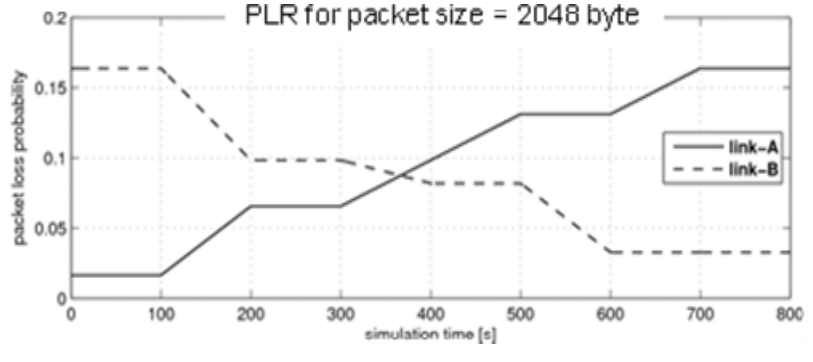
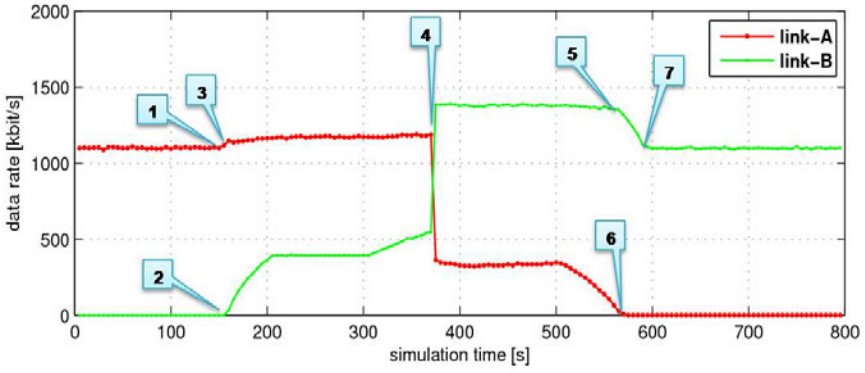


Fig. 10. Packet loss probability in Left-Hand Side Wireless Compartment



1. Video-GP begins to send duplicated packets on link-A
2. Video-GP begins to send duplicated packets also on link-B
3. Audio-GP begins to send duplicated packets on link-A.
4. Link-B is now the "better link"
5. Audio-GP sends single traffic only
6. Video-GP only uses better link
7. Video-GP sends single traffic only

Fig. 11. Combined Data Traffic over the Left-Hand Side Wireless Compartment

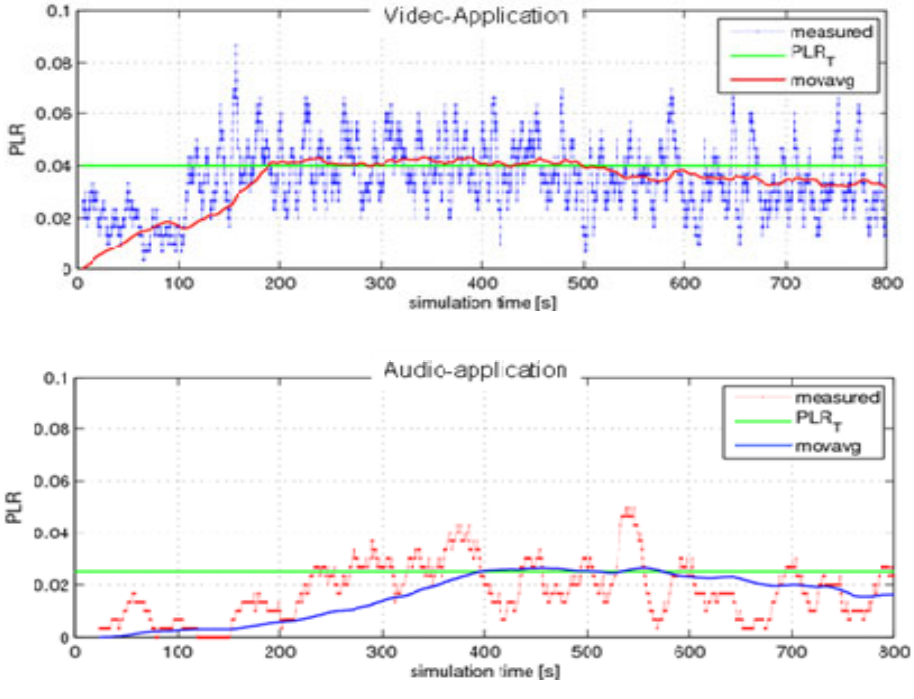


Fig. 12. Loss Ratio for Generic Connections

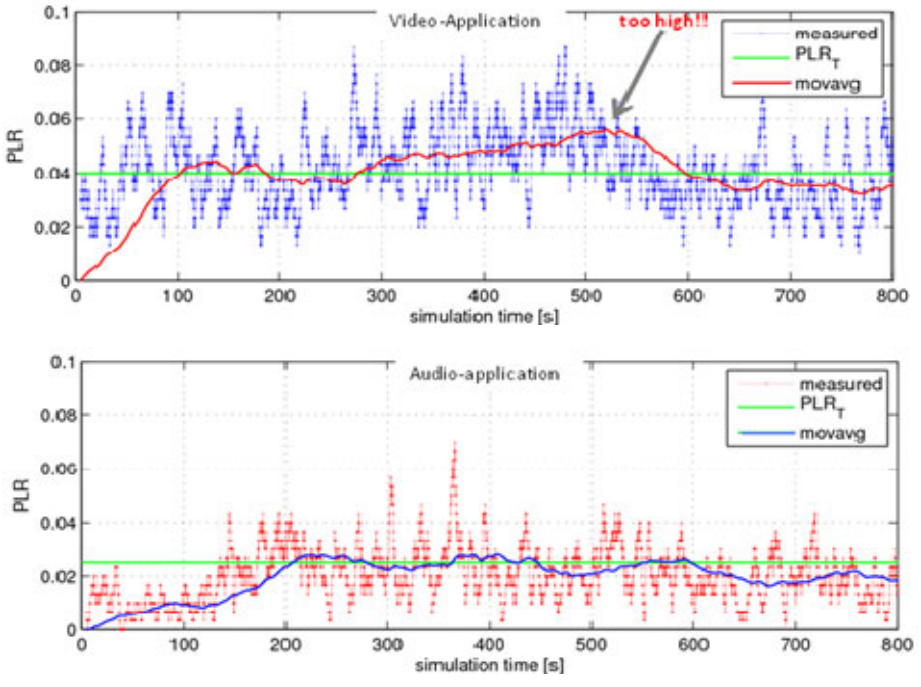


Fig. 13. Packet Loss Ratio for Generic Connections in Bandwidth Limited Scenario

As depicted in Fig. 11, initially both the video and audio Generic Connections use the same link-A for all their traffic. The video traffic requires duplication due to its target PLR and decreasing link-A quality (earlier than the audio traffic) therefore, it switches to duplication mode first and then very quickly starts using additional the resources on link-B (after link-A is fully loaded). After approximately 370s, link-B is the better link and hence it is utilized to the maximum by the two Generic Connections and small portion of data traffic is put on link-A. As the link-B's quality improves the duplicated traffic on link-A is reduced and finally only link-B is used alone by both the Generic Connections.

Fig. 12 depicts the obtained packet loss ratio for both the audio and video Generic Connection and it can be seen that the computed moving average is always in line with the target packet loss ratio (PLR_T).

Finally, a bandwidth constrained example is presented. Now the audio Generic Connection has a higher priority than the video Generic Connection during the resource allocation of the GC architecture. The audio traffic data rate is now 500kbit/s and the allocated bandwidth on both link-A and link-B is 600kbit/s. The video traffic data rate is still 1Mbit/s while the allocated bandwidth on link-A and link-B is only 600Kbit/s and 800Kbit/s, respectively. The packet size and the target packet loss ratio are same for the two Generic Connections as in case of the previous example.

Fig. 13 depicts the packet loss ratio seen by the two Generic Connections. As the video Generic Connection lacks resources, the packet loss ratio is higher than the

target value during the simulation run. In contrast, for the audio Generic Connection, there was enough available bandwidth and hence it is always able to conform to its target packet loss ratio. Such a traffic differentiation is still hardly possible in the Internet.

7 Conclusion and Outlook

In this paper, the Generic Connectivity concept was briefly introduced and it was shown that this clean-slate protocol framework is very flexible and powerful. The feasibility and advantages of this framework are demonstrated in several scenarios. The results show that the Generic Connectivity mechanisms support innovative networking paradigms that cannot easily be realized by the current Internet protocol architecture, such as the automatic cross-layer adaptation of error correction mechanisms or flexible per-flow multipath routing over heterogeneous access networks.

Further work is needed to address some remaining research issues such as the design of the signaling mechanisms to realize the Generic Connectivity, as well as further extensions of the supported protocol mechanisms, e. g., for resource management or incremental deployment.

References

1. Banniza, T.-R., Boettle, D., Klotsche, R., Schefczik, P., Soellner, M., Wuenstel, K.: A European Approach to a Clean Slate Design for the Future Internet. Bell Labs Technical Journal (2009)
2. 4WARD project page, http://www.4ward-project.eu
3. Bertin, P., Aguiar, R., Folke, M., Schefczik, P., Zhang, X.: Paths to Mobility Support in the Future Internet., ICT-MobileSummit, Santander, Spain (June 2009)
4. Van Jacobson, K.: Introduction to Content Centric Networking. In: FISS 2009, Bremen, Germany (June 2009)
5. Ford, B., Iyengar, J.: Breaking Up the Transport Logjam. In: Proc. ACM HotNets-VII (October 2008)
6. D-5.2.0, Mechanisms for Generic Paths, public 4WARD deliverable (2009), http://www.4ward-project.eu/index.php?s=file_download&id=38
7. Day, J.: Patterns in network architecture: a return to fundamentals (Bd. 1). Prentice Hall International, Englewood Cliffs (2008)
8. OMNeT++ Discrete Event Simulation System Version 4.0

Using BGP-4 to Migrate to a Future Internet

Pedro A. Aranda Gutiérrez[1], Petteri Pöyhönen[2],
Luis Enrique Izaguirre Gamir[1], and Francisco Huertas Ferrer[1]

[1] Telefónica, Investigación y Desarrollo,
Emilio Vargas, 6, Madrid Spain
[2] Nokia Siemens Networks,
Linnoitustie 6, 00260 Espoo Finland

Abstract. The Internet has evolved to become one of the most critical communication infrastructures in the planet. And yet, some of its underlying concepts and protocols do not provide the adequate level of reliability for such an essential role in global communications. The inter-domain routing protocol of the Internet, Border Gateway Protocol (BGP-4), is being used with varying degree of success for tasks for which it was not originally designed, such as Traffic Engineering. This paper presents a rationalised view of the different functions implemented by routing nowadays and proposes the use of Autonomous System Compartments. The Autonomous System (AS) Compartments imply a new routing hierarchy over the traditional BGP-4 routing, where specific functionalities like Traffic Engineering can be better controlled and additional routing incentives can be introduced. The FP-7 project 4WARD is working on new communication paradigms for the Future Internet and AS Compartments are a choice to contain the Generic Path (GP) concept developed by it. In order to provide inter-domain capabilities and a migration tool to connect GP islands, the multiprotocol mechanism of the BGP-4 routing is used. This paper presents the AS Compartment concept and the integration of Generic Paths in it, as well as an implementation of the GP-BGP concept for the J-Sim simulator (JSIM) environment.

Keywords: Autonomous Systems, Inter-Domain Routing, Compartments, Traffic Engineering.

1 Introduction

The Internet is being perceived as a commodity nowadays. On the other hand it truly is a critical communication infrastructure. Many services are used and contents are distributed over the Internet. Users are stationary when using their PCs from home and/or office and moving users when using their mobile devices. Especially with the adoption of new 3G radio access networks like High-Speed Downlink Packet Access (HSDPA) and wider adoption of mobile broadband, the difference between these two types of users in terms of access bandwidth is vanishing. For service providers this means a potentially larger amount of users and for the Internet naturally this means higher traffic transfer demands. While operators and other service providers introduce new exciting value added services,

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 14–25, 2011.

this also requires a better communication infrastructure, especially in terms of availability and stability. When using the Internet to provide IP backbone connectivity between mobile operators, we face many challenges to ensure a similar stability compared to the GPRS Roaming Exchange (GRX) [1] service quality level. A global use of the Internet typically also involves the use of BGP-4 [2] based core routing. This core network provides a fairly resilient routing, but it is a well-known behaviour that it also could take relatively long time, i.e., tens or hundreds of seconds, until the routing system restores its stable state after a routing incident. Such an incident can occur due to a configuration error, network maintenance, (physical) link failure and so on. Routing system stability is perhaps one of the main challenges and is something that should be taken into account while considering "better than best effort" end-to-end services.

Despite the BGP-4 protocol and routing being well-defined, there are different deployment practices, which are derived from the need for traffic engineering in order to comply with peering agreements. Not all BGP-4 route attributes are used in all network domains in a consistent way. An example of this is the Multi-Exit Discriminator (MED) attribute [3], which provides a mechanism for an AS to indicate to adjacent ASs the optimal inbound link (e.g. in the case of multihoming). Another example is the AS Path (AS_PATH) attribute. [4] further explains differences on the routing policy deployment and how they affect to the BGP-4 routing due to the diversity in processing BGP-4 messages.

The FP-7 4WARD project [5] has taken a Clean Slate approach to the Future Internet, exploring new insights in multi-access and resource management. One of the solutions which have emerged from this effort is the GP concept [6], as a new paradigm to support rich and flexible communication schemes. As all new technologies, the adoption cannot be expected to be instantaneous and unanimous across the Internet. The most reasonable scenario is a gradual adoption by smaller user groups, resulting in networking "islands" which need to be interconnected. This has been the case of IPv6 with the 6bone [7] and other migration mechanisms [8,9].

This paper presents a framework to interconnect GP islands using the Autonomous System Compartment concept and multiprotocol extensions to BGP-4. The rest of paper is structured as follows. Section 2 describes a new concept called AS Compartment and explains how this concept is used in inter-domain networking environment. Section 3 discusses the roles of traffic engineerig in the AS Compartments and describes a high level logical architecture. Section 4 presents the simulation experiments carried out to evaluate how a BGP-4 multiprotocol extension functions and performs on top of BGP-4. Section 5 provides conclusions and finally, section 6 outlines the related work.

2 AS Compartments

In order to improve BGP-4 routing resilience and to support more flexible ways of doing Traffic Engineering (TE), a new concept called AS Compartment is proposed. The concept introduces a new routing "hierarchy" on top of the BGP-4

routing system and instead of relying on the standard BGP-4 convergence, a fast re-routing is supported at the AS Compartment level. The AS Compartment routing complements the BGP-4 routing by also taking into account different end-to-end incentives and could use for instance multi-path routing between AS Compartments. In case of a multi-path routing, the authors of [10,11] define new multi-path routing protocols designed for inter-domain environment that could be used to implement the multi-path routing support in AS Compartments. An AS Compartment could be either a single AS or it can include a set of ASs. And therefore, an AS Compartment could represent one or more Autonomous System Numbers (ASNs) and could be configured for instance to use AS Confederations [12]. If an AS Compartment consists of multiple ASs, then it is assumed that each border gateway hosting a *path* end point has connectivity to each other inside the AS Compartment. The AS Compartments provide means to introduce a limited control over the BGP-4 infra without modifying the basic BGP-4 protocol and to separate traffic engineering and routing functions from each other. So one of the main challenges is to ensure that for certain type of traffic that is passing through the BGP-4 routing system the perceived connection quality is sufficient only with introducing additional functionality in a selected set of Autonomous Systems as illustrated in Figure 1.

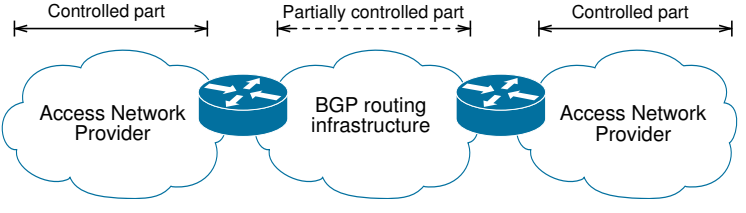


Fig. 1. An end-to-end connectivity over Internet core

An AS Compartment does not necessarily mean that QoS-aware routing is supported. Thus a simple form of AS Compartment could be a *best effort* network with some advanced routing and traffic engineering functionalities implemented. AS Compartments are interconnected by *paths* that are used to route traffic between AS Compartments. A *path* could be for instance IP tunnel having BGP-4 routable IP addresses as end points. This tunnel is terminated at border gateways located at the AS Compartments.

Figure 2 shows an example of two AS Compartments connected by 3 *paths* that are transported over normal BGP-4 routing implemented by AS topology. *Path 1* and *Path 2* are used to transfer traffic to the "remote" AS Compartment and *Path 3* is used to transfer traffic from the "remote" AS Compartment. The selection between *Path 1* and *Path 2* is done based on the underlying BGP-4 routing info. For instance, if *AS-n* indicates in its route advertisements with the MED attribute that one link should be preferred over another, then this is taken into account when selecting an active path. Additionally, this selection process could use any available BGP-4 routing information as long as BGP-4

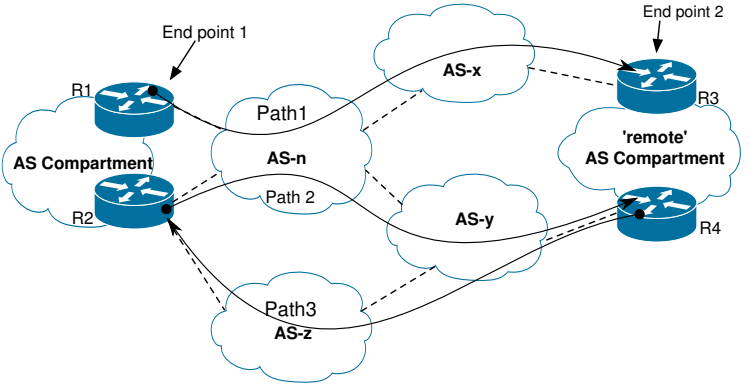


Fig. 2. An example of AS Compartments connected over the BGP routing

routing practices are honoured. In other words, AS Compartment routes are *paths* between Compartments and these routes exist only when corresponding BGP-4 routes are also present.

Concerning different types of relations between ASs, there are some exceptions we should consider. For instance, Tier1 and Tier2 ASs could have peering relations with other same level ASs. These peering agreements are not used for transit traffic due to the lack of economic incentives, thus the main motivation of their use is to minimize transit costs as well as minimize end-to-end latencies [13]. However, in some situations, it might be preferred to override this peering policy. For instance, to make it possible that two (or more) single homed ASs could create a multi-homed AS Compartment. This implies that the authorities creating an AS Compartment over a peering relation have common incentive(s) to allow the use of their peering link(s) for the selected transit traffic.

3 Traffic Engineering in AS Compartments

In order to implement new policies and to avoid contradicting routing policies between ASs, TE functions are separated from routing functions and a uniform TE process over the ASs is defined. Also, all ASs belonging into the same AS Compartment should contribute and implement "atomic" policies towards this AS Compartment. The objective of traffic engineering is to distribute traffic at AS peering points in such a way that they comply with the Service Level Agreements (SLAs) signed by each AS with its peers, assuming that SLAs are convertible from one AS to another one of the AS Compartment. Ideally, both input and output traffics should be controlled. In the current Internet, controlling the output traffic can be implemented internally in the AS, but controlling the input traffic can only be achieved by controlling the routing preferences in other ASs.

Basically, there are three kinds of attributes in the BGP-4 routing decision depending on their scope; 1) router local, 2) AS local, and 3) global [14]. The routing decision process during which the best path is computed is the well-defined

process taking also into account the local policies. So for the inbound traffic, an AS can tweak the route attributes to be announced in hope of influencing a neighbour AS best path selection. For the outbound traffic, there are more powerful means available like the attributes representing local policies like the *local preference* attribute. There are also other attributes, conditions and local policies influencing to the routing decision like route type ("customer", "provider" or "peer"), an internal (Interior Gateway Protocol (IGP)) topology, the BGP-4 community attribute, and so on. Since the AS Compartments are operating on top of BGP-4 routing, they can coordinate how to handle both inbound and outbound traffic in order to comply additional routing incentives without making this visible at the BGP-4 level.

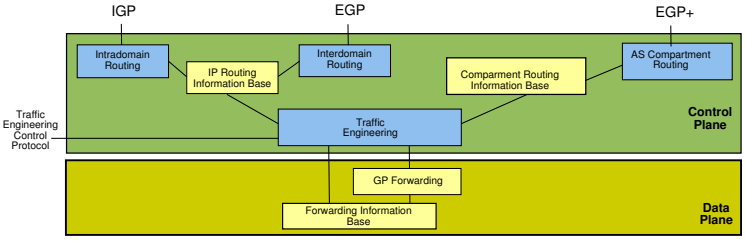


Fig. 3. Extracting the TE functionality from the extended routing framework

Figure 3 shows a high level architecture for node integrating traditional IP routing with augmented GP routing and TE functions. The control plane integrates today's IP intra- and inter-domain routing functions, GP routing functions and a separate TE block.

4 Towards a Practical Implementation

One way to interconnect different GP islands over traditional IP based infrastructures is to use IP tunnel [15] having BGP routable addresses as end points. This tunnel is terminated at border gateways located at GP islands. Inter-GP island routing is established by defining the GP Network Layer protocol that defines the so-called GP-BGP-4, a new routing hierarchy that enables GP island to exchange their routes over the traditional BGP-4.

To enable BGP-4 to support routing for multiple Network Layer protocols, Multiprotocol Extensions for BGP-4 [16] adds the ability to associate a particular Network Layer protocol (e.g., IPv6, IPX, L3VPN, etc.) with the next hop information and Network Layer Reachability Information (NLRI). GP-BGP uses the multiprotocol extensions capabilities to exchange GP islands routes and achieve, together with IP tunneling, the inter-GP islands routing.

In order to provide a proof of concept and a first evaluation of the applicability of the proposed solution, the Generic Path-BGP-4 extensions have been implemented for a proof of concept on the JSIM [17]. Figure 4 shows a basic simulation environment with three nodes, the two in the edges running a GP-BGP

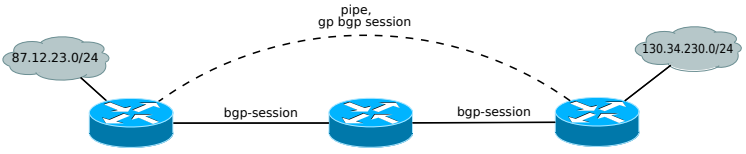


Fig. 4. GP-BGP Basic Setup

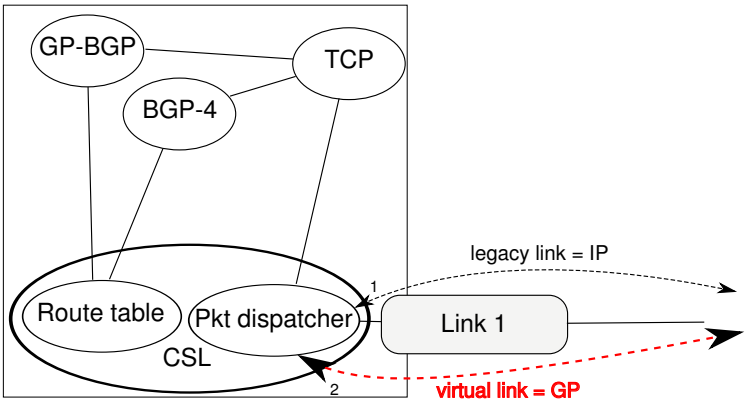


Fig. 5. Node structure in the simulation environment

session over an IP tunnel that basically follow the logical functionality decomposed in Figure 5. In order to obtain early results, routing compartments are simulated as IPv4 prefixes.

4.1 Fallback Scenario

There is however a situation where there is no loss of connectivity when a specific link or router failure happens, and this is when the AS Compartment concept is fully exploited. Figure 6 shows the simulation environment in that case. It shows a topology where the border gateways of the routing Compartments (CTs) - simulated by prefixes 1.0.0.0/8, 2.0.0.0/8, 6.0.0.0/8 and 7.0.0.0/8- run a full mesh of GP-BGP sessions. Dashed lines show GP-BGP sessions, while continuous lines show IP-BGP sessions. As mentioned before and in order to obtain early results, routing compartments were simulated as IPv4 prefixes.

The following listing shows the routing status for Router1 after the BGP sessions both IP and GP exchange their routes:

```
192.168.1.2 -> direct link
192.168.1.3 -> direct link
192.168.1.6 -> 192.168.1.3 (1003 1004 1006)
192.168.1.7 -> 192.168.1.3 (1003 1004 1007)
10.0.1.2 -> virtual connection (192.168.1.2)
10.0.1.6 -> virtual connection (192.168.1.6)
10.0.1.7 -> virtual connection (192.168.1.7)
```

where the virtual connection refers to the IP tunnel to be used for reaching the other CT routes. The IP tunnels generated are shown in Figure 7(a).

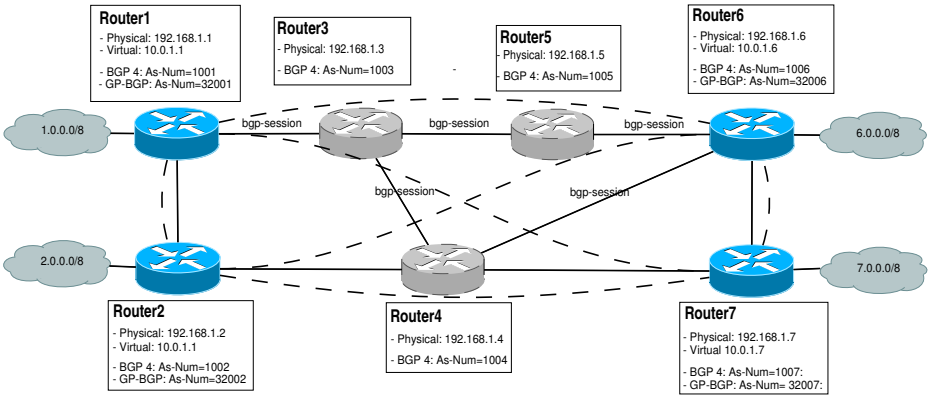
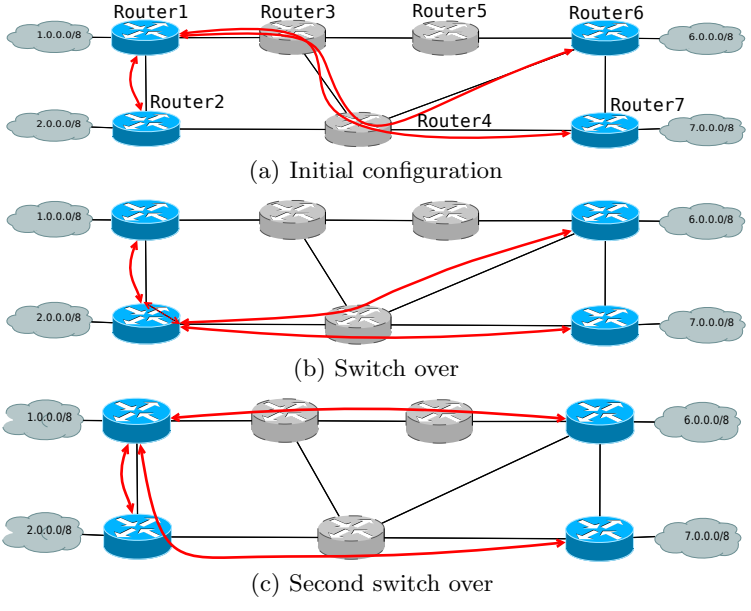


Fig. 6. Connection of AS Compartments



(a) Initial configuration

(b) Switch over

(c) Second switch over

Fig. 7. Evolution of tunnels in Router1

At instant t=1000s, the link between Router3 and Router4 fails so, when Router3 stops receiving the IP-BGP keep-alive messages, it resets its BGP Finite State Machine with Router4:

```
1084.8517681890637 bgp-id=3232235779, peer=(1004;192.168.1.4/32), ESTABLISHED=(HoldTimerExp)=>IDLE
```

Keep-alive message interval is set to T=30s while the time to tear down a session when not receiving keep-alives is 3*T for both GP-BGP and IP-BGP- Router3

withdraws the routes previously exchanged with Router4 announcing that withdraws back to Router1, which modifies its routing table accordingly:

```
1089.8523388557303 192.168.1.7 -> 192.168.1.2 (1002 1004 1007)
1089.8529095223969 192.168.1.6 -> 192.168.1.2 (1002 1004 1006)
1119.8523281890637 192.168.1.6 -> 192.168.1.3 (1003 1005 1006)
```

The route towards 192.168.1.6 and 192.168.1.7 traverses now Router2 instead of Router3. Note that the route towards 192.168.1.6 bounces back to Router3 once Router3 inform Router1 about the new path across AS 1005 instead of AS 1004.

On the meantime, even before the IP-BGP routing reorders its routing, the GP-BGP in Router1 reroutes its traffic through Router2 when it does not receive the appropriate keep-alive messages from Router6 and Router7:

```
1075.3373485086665 bgp-id=167772417, peer=(32006;10.0.1.6/32),ESTABLISHED=(HoldTimerExp)=>IDLE
1089.0377309912103 bgp-id=167772417, peer=(32007;10.0.1.7/32),ESTABLISHED=(HoldTimerExp)=>IDLE

1075.3373485086665 6.0.0.0/8 -> 10.0.1.2 (32002 32006)
1089.0377309912103 7.0.0.0/8 -> 10.0.1.2 (32002 32007)
```

With this routing change on the GP-BGP level, traffic towards 6.0.0.0/8 and 7.0.0.0/8 from 1.0.0.0/8 will be sent towards Router2 using the previously setup tunnel for encapsulating traffic towards 2.0.0.0/8 from Router1. Router2 then decapsulates the traffic and encapsulates it again using as well the previously setup tunnels for encapsulating traffic towards 6.0.0.0/8 and 7.0.0.0/8 from Router2. This behaviour is shown in Figure 7(b).

It is important to note that traffic is forwarded to 6.0.0.0/8 and 7.0.0.0/8 via Router2 using two consecutive tunnels before IP-BGP routing reacts from the link failure and updates its routes:

```
time routing recovers towards 6.0.0.0/8 is  setup again via GP-BGP = 1075.3373485086665 s
time routing recovers towards 7.0.0.0/8 is setup again via GP-BGP = 1089.0377309912103 s
time routing recovers towards 6.0.0.0/8 and 7.0.0.0/8 via IP-BGP = 1089.8523388557303 s
```

Once Router1 re-establishes the GP-BGP session with Router6 and Router7,

```
1099.0435123245438 bgp-id=167772417,peer=(32007;10.0.1.7/32),OPENCONFIRM=(RecvKeepAlive)=>ESTABLISHED
1106.342276508667 bgp-id=167772417,peer=(32006;10.0.1.6/32),OPENCONFIRM=(RecvKeepAlive)=>ESTABLISHED
```

and the routes towards 6.0.0.0/8 and 7.0.0.0/8 are announced again to Router1, Router1 selects these paths:

```
1104.0466803245442 7.0.0.0/8 -> 10.0.1.7 (32007)
1141.3457858420006 6.0.0.0/8 -> 10.0.1.6 (32006)
```

falling back to the previous situation where direct and unique tunnels where used to forward traffic between the different CT routes. This behaviour is shown in Figure 7(c).

It is important to note that even the tunnels are the same as in the initial situation, they traverse different ASs as specified by the underlying IP-BGP routing protocol.

Finally, as the link from Router3 and Router4 recovers at t=2000s the IP-BGP converges back to the initial situation of Figure 7(a). This recovery is totally transparent to GP-BGP protocol as no sessions states are changed and no routing announcements are done: the only changes are on the ASs the tunnels traverse but not on any of the GP-BGP signalling.

4.2 Simulation Results

In order to check the scenario shown in Figure 6, 1000 simulations were performed to see how the unpredictable conditions such as time delays, packet losses, etc. affect these results. Figure 8 compares the density functions of IP traffic recovery time in the IP-BGP case versus the GP-BGP case.

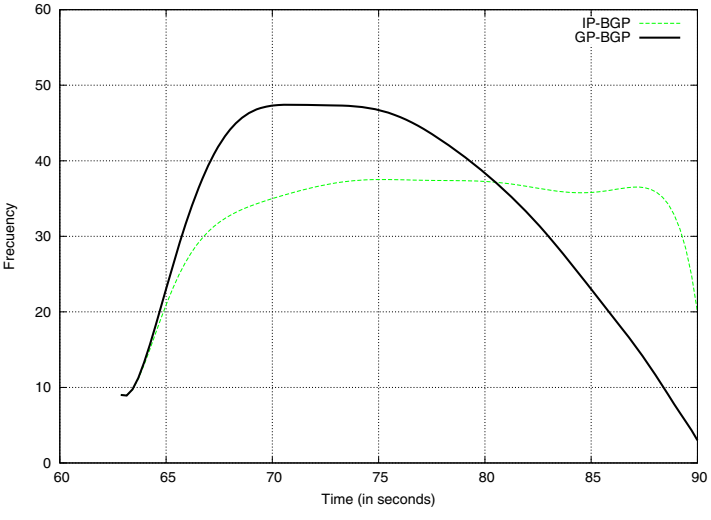


Fig. 8. Density function of traffic recovery time

It is important to note that IP traffic recovery time in the GP-BGP case is always slightly better than IP-BGP case. This is the case because the IP traffic for the GP-BGP routes can be recovered in two ways; the first one when GP-BGP decides to re-route the traffic through the new path, and the second one when BGP-4 detects the failure and chooses a new alternative route to forward the traffic. IP traffic for the GP-BGP case take advantage from these two detection and recovery algorithms: its recovery time is the minimum between the two. Also, this gain can be improved by decreasing the GP-BGP time between keep-alive messages, so failures can be detected earlier; however, there is a tradeoff with the extra signalling overhead it introduces, therefore extra care is needed when setting this parameter.

5 Conclusion

The 4WARD project has investigated different approaches to provide enhanced services which are not possible nowadays over the Internet. One of the cornerstones for such novel networking proposals is the Generic Path concept. But, in order for any Future Internet technology to be deployed, a migration path which takes into account pre-existing network technologies is needed. Experience shows

that adoption of a new technology starts with small isolated sections which need interconnection in order to flourish. In this paper we have presented the AS Compartment concept and supporting BGP extensions which will be helpful for such migration scenarios. We show through simulations, that the AS Compartment concept is not only applicable in Future Internet migration scenarios, but might also be used in the near future to enhance the resilience of the current Internet.

6 Related Work

One of the main challenges in the area of BGP-4 routing is scalability in terms of a size of Routing Information Base (RIB) and Forwarding Information Base (FIB) entries. There are many reasons why the sizes of these tables have increased, but maybe most significant reasons are address prefix de-aggregation and use of routing policies based on various reasons. There are many BGP-4 improvement proposals to improve the scalability. [18] proposes the method according to which a set of topologically co-located Internet Service Providers (ISPs) could agree to share a network prefix(es) and aggregate the common prefix(es). The authors of [19] discovered that many routable prefixes share same AS path and in order to optimize the space usage prefixes are divided into atoms that are then routed and advertised instead of prefixes. This would reduce both FIB and RIB sizes in the Default-Free Zone (DFZ) and therefore also potentially improve convergence. [20] analyzes the current BGP-4 routing including both interior and exterior BGP-4. Based on the analysis, the authors proposed a new enhanced BGP-4 protocol called the *atomic BGP* that could be also deployed incrementally. This protocol makes an AS to use non-contradicting routing policies, i.e., all routers inside the AS make route selection and dissemination in the same way. As a result of this, an AS can be seen as a single node to the outside. The *atomic BGP* can lead to a simpler network management which could mean less routing errors and misconfigurations resulting "false" BGP-4 updates and convergence.

Current practices to implement traffic engineering in BGP-4 routing [21] have to be re-examined. They are also relevant while designing how multi-path routing is setup and determine the kind of benefits which can be derived from their use. For instance, if the main motivation to use multi-pathing is to improve resilience, then it becomes essential to try minimizing a number of common BGP-4 links to be used by multi-path flows linked to a single end-to-end session to maximize resilience in case of BGP-4 routing failure.

Another issue of the current BGP-4 routing system is its relatively slow convergence from routing errors/updates. BGP-4 convergence can be divided into two phases, 1) failure detection and 2) path exploration. Once the routing failure has been detected in an AS, there is typically a predefined delay until BGP-4 updates are sent to notify other ASs. The default value of the Minimum Router Announcement Interval (MRAI) are 5 seconds for interior and 30 seconds for exterior routing. BGP-4 system behaviour has been widely analysed based

on BGP-4 traffic samples collected by Oregon Routeviews [22] and the RIPE RIS [23] projects. Thus, the analysis in [24] clearly shows that approximately 36% of monitored update sequences took longer than 60 seconds to complete. The authors of [25] have studied how Voice over IP (VoIP) calls and their quality degradation correlate with the BGP-4 updates in the core network. The study shows that BGP-4 has a similar negative impact on the VoIP quality as network congestion.

The impact, though, is not extremely severe for some application types and use scenarios. For instance, Delay Tolerant Network (DTN) based applications as well as elastic traffic could tolerate connection breaks quite well. On the other hand, in any real-time application the situation would suffer a greater impact, since typically relatively long connection breaks are not transparent to the end-users.

7 EU Disclaimer

This paper describes work undertaken in the 4WARD project, which is part of the EU IST programme. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the 4WARD project. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.

References

1. GSM Association. Official document: Ir.34 - inter-service provider ip backbone guidelines, v. 4.4 (June 2008), http://www.gsmworld.com/documents/ir3444.pdf
2. Rekhter, Y., Li, T., Hares, S.: A Border Gateway Protocol 4 (BGP-4). RFC 4271 (January 2006)
3. McPherson, D., Gill, V.: BGP MED Considerations. RFC4451 (March 2006)
4. Scholl, T.: Best Practices for Network Interconnections. Presentation on NANOG43 (June 2008)
5. The 4WARD Project, http://www.4ward-project.eu/ (last visit April 27, 2010)
6. The 4WARD Project. D-5.1 Architecture of a Generic Path (2009) (last visit April 27, 2010)
7. 6bone, http://en.wikipedia.org/wiki/6bone (last visit April 27, 2010)
8. Carpenter, B., Moore, K.: RFC3056 - Connection of IPv6 Domains via IPv4 Clouds (2001), http://www.ietf.org/rfc/rfc3056.txt (last visit April 27, 2010)
9. Making the Transition From IPv4 to IPv6 (Reference), http://docsun.cites.uiuc.edu/sun_docs/C/solaris_9/SUNWaadm/IPV6ADMIN/p21.html (last visit April 27, 2010)
10. Fujinoki, H.: Multi-path BGP (MBGP): A solution for improving network bandwidth utilization and defense against link failures in inter-domain routing. In: 16th IEEE International Conference on Networks, ICON (2008); ISSN: 1556-6463, Print ISBN: 978-1-4244-3805-1

11. Xu, W., Rexford, J.: MIRO: Multi-path Interdomain ROuting. ACM SIGCOMM Computer Communication Review archive 36, 171–182 (2006); ISSN:0146-4833
12. Traina, P., McPherson, D., Scudder, J.: Autonomous System Confederations for BGP. RFC5065 (August. 2007)
13. William, B.: Norton. Internet Service Providers and Peering (2000)
14. Caesar, M., Rexford, J.: BGP routing policies in ISP networks. IEEE Network 19, 5–11 (2005)
15. Perkins, C.: IP Encapsulation within IP. RFC 2003 (October 1996)
16. Bates, T., Chandra, R., Katz, D., Rekhter, Y.: Multiprotocol Extensions for BGP-4. RFC4760 (January 2007)
17. http://sites.google.com/site/jsimofficial
18. Internet Service Provider Address Coalitions (ISPACs). IETF Internet draft
19. Verkaik, P., Broido, A., Hyun, Y., Claffy, K.C.: Atomised Routing. Presentation in RIPE45 meeting,
http://www.nlnet.nl/project/atombr/20030512-atoms-ripe45.pdf
20. Zhang-Shen, R.: Atomic Routing Theory: Making an AS Route Like a Single Node. Invited talk in NSF FIND Routing Workshop (2008)
21. Halabi, B.: Internet routing architectures, 2nd edn. (2000)
22. http://www.routeviews.org/
23. http://www.ripe.net/projects/ris/rawdata.html
24. http://www.potaroo.net/ispcol/2007-06/dampbgp.html
25. Kushman, N., Kandula, S., Katabi, D.: Can You Hear Me Now?! It Must Be BGP. ACM SIGCOMM 37, 75–84 (2007)

Revisiting the Impact of Traffic Engineering Techniques on the Internet's Routing Table

Pedro A. Aranda Gutiérrez

University of Paderborn, Germany
paaguti@hotmail.com

Abstract. This paper studies the effect of simple Traffic Engineering techniques on the size of the Internet's default free routing table. Current best practises for traffic balancing in the Internet are based in disaggregating prefixes that cause an increase in size of the Internet's core routing table. An algorithm to show the impact of these techniques on the growth of the routing table is proposed. This algorithm is applied on routing tables between January 2001 and December 2009 and the results are discussed. Finally an alternative architecture is proposed, which allows Traffic Engineering while keeping the Internet routing table size optimised.

Keywords: Routing protocols, Network Operations, Network management, Network monitoring.

1 Introduction

IP routing protocols control the exchange of network layer reachability information between nodes in an IP network. This information, also known as routing information, is used to build the routing table. IP addresses in a router are grouped into ranges, which are known as prefixes. The packet forwarding process in IP nodes, which computes the outgoing port for an incoming IP packet, is controlled by the routing table and uses longest mask prefix matching on the destination address to compute the output port. The two basic concepts which have to be understood are longest mask prefix matching and the Route Decision Process. Longest mask prefix matching implies that a router will always prefer the most specific routing information installed in the routing table to reach a given IP address. The Route Decision Process (RDP) is specific for each routing protocol used by the router and defines the way routing information received from neighbours is treated and when routes are installed in the routing table.

The Internet is an IP infrastructure which is divided into different independent and interconnected domains, which are known as Autonomous Systems (ASes) [12]. Each AS is assigned addressing space in the form of one or more prefixes by the Internet Routing Registries (IRRs) of the region it belongs to. It is only allowed to advertise routing information for the addressing space it has been assigned. While there exist different routing protocols to control the routing information exchange within an AS, BGP-4 [21] is the only routing

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 26–37, 2011.

protocol which controls the routing information between ASes. ASes which are interconnected are said to be peering. Internet Service Providers (ISPs) control the traffic distribution on their peering links, because they need to assure that no traffic is lost when a link or border router fails, or that traffic levels are such, that the network meets Quality of Service (QoS) criteria, etc. Internet Service Providers have developed different strategies to cope with all these requirements. In order to assure resilient connectivity, most ISPs are *multi-homed*. i.e. have more than one upstream connection, sometimes to more than one upstream provider.

The rest of this paper is structured as follows: Section 2 discusses how ISPs or larger sites with more than one access to the Internet implement Traffic Engineering with BGP-4 and focuses on two techniques used to balance the inbound traffic of an AS, which can be considered current best practises. Section 3 presents the Internet routing table compression algorithm used to estimate the impact of the use of current best practises for Traffic Engineering in the Internet. The algorithm is then applied on data from the RIPE's Routing Repositories to quantify the impact of Traffic Engineering (TE) techniques on the current Internet routing table. Section 4 presents an alternative to control the growth of the Internet routing table while allowing for TE solutions and shows the impact on the Internet's core routing table if it had been applied between January 2001 and December 2009. Finally, Section 6 presents the conclusion.

2 BGP-4 and Traffic Engineering

Internet Service Providers organise their interconnection through peering agreements, which include the definition of technical and economical conditions under which they exchange traffic. The technical definitions include addressing space which is made mutually accessible, the mechanisms to route traffic through the interconnection links and Internet Protocol layer parameters like round trip delay and tolerated levels of packet loss and acceptable traffic levels for the in- and outbound links. Service Level Agreements (SLAs) introduce an additional incentive for mechanisms to control the in- and outbound traffic of a network and, thus, for the implementation of TE techniques. However, BGP-4 as the inter-domain routing protocol of the Internet lacks real TE capabilities. Despite this, routing configurations targeting simple load balancing between independent links to a major upstream ISP to load sharing between several upstream ISPs have been deployed from the early days of the commercial Internet. These are documented in vendor manuals [5], [20] and books about BGP-4 [14], [11]. In order to arrive as close as possible to the desired traffic distributions, attributes in the routing advertisements are manipulated in order to influence the routing decision process. Two examples are shown in Section 2.1. These configurations are considered current best practises [4], [14].

Controlling the inbound traffic of an AS implies influencing the routing decisions of other ASes. As Griffin and Wilfong have demonstrated in [10], predicting BGP-4 behaviour is impossible. BGP-4 is not always guaranteed to converge to

one single solution in the presence of policies. Since the effect of configuration changes cannot be predicted, arriving at traffic conditions that comply to the SLAs signed between an AS and its peers is an iterative process of Trial and Error based on deploying a certain routing configuration, assessing its quality by the traffic distribution it creates in the inter-provider links, refining the configuration and reassessing. This process aims at an ideal traffic distribution with respect to some objective, e.g. minimisation of peering costs, uniform traffic distribution, etc.

In order to achieve the best approximation to the ideal traffic distribution, ISPs fraction their addressing space. G. Huston [13] recognises the use of this technique and examines its impact on the routing tables of the Default Free Routing area of the Internet. But, as Section 3 shows, the impact is greater than Huston's graphics imply.

2.1 Current Best Practises to Control the Inbound Traffic

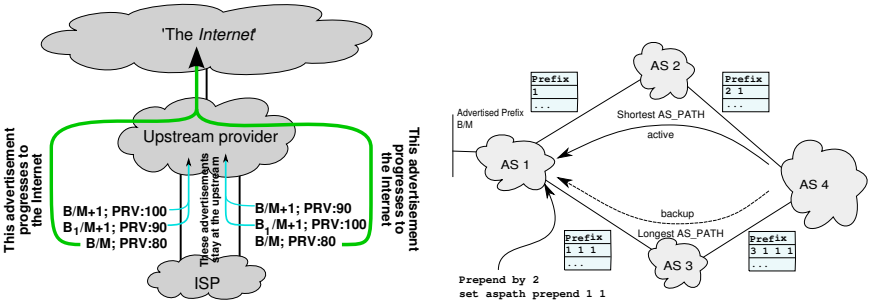
Prefixes are sets of contiguous IP addresses, designated by a base address (B) and a mask (B_M). In the case of IPv4 the base address and the mask are unsigned 32-bit integers and in the case of IPv6, they are 128-bit long integers. The mask B_M has its M most significant bits set to 1 and the rest set to 0.

Definition 1. *B/M represents a valid prefix $\mathcal{P}$ if and only if $B \wedge B_M = B$.*

Definition 2. *Let $\mathcal{P}$ be the set of addresses represented by prefix B/M and A be an IP address of the same family as B then:*
$A \in \mathcal{P} \Leftrightarrow A \wedge B_M = B$

Definition 3. ***Subnetting:****Let $\mathcal{P}$ be the set of addresses represented by prefix B/M. $\mathcal{P}$ can be divided in two subsets, known as sub-networks $\mathcal{P}_1, \mathcal{P}_2$ that contain the same amount of IP addresses. This implies that the mask length will be incremented by 1. The sub-networks will be denoted as $\mathcal{P}_1 = B/M + 1$ and $\mathcal{P}_2 = B_1/M + 1$ in the rest of this paper.*

Multihoming to one provider. Figure 1(a) shows an AS which is providing Internet access to another AS through two connections. The downstream ISP owns prefix B/M and has subnetted it into two subnetworks $B/M + 1$ and $B_1/M + 1$. It advertises the three prefixes and uses an agreed marker (a.k.a. community) to signal the preference. The upstream ISP translates this marker to his local preference and uses it to filter the advertisements to *the Internet*. Internally, due to longest prefix matching, the ISP will use the sub-networks $B/M+1$ and $B_1/M+1$ to direct the traffic towards its client. This configuration also assures redundancy: in case one link fails, the full routing information is available through the other. Figure 1(a) shows the ideal situation, where the sub-networks advertised for TE purposes do not progress to the Internet. In real life, as shown in Section 3, most of the TE sub-networks progress to the Internet.



(a) Multihoming to one provider (b) Multihoming over intermediate providers

Fig. 1. Current Practises in Multihoming: fractioning the assigned addressing space to balance the incoming traffic

Multihoming to more than one provider. The technique described in the previous section only works for direct client-provider peerings, because communities are not guaranteed to progress beyond the first AS and the local preference is only valid within the AS where it is defined. The only attribute which is guaranteed to progress through Autonomous Systems is the Autonomous System Path (AS_PATH). The number of hops it contains is one of the first variables used by the RDP to calculate the best path. On each interdomain border, BGP-4 speakers prepend their Autonomous System Number (ASN) at the beginning of the AS_PATH attribute when exchanging routing information with speakers outside their AS. Additionally, the AS_PATH attribute can be manipulated by AS_PATH Prepending, a technique which consists in prepending the ASN more than once. AS_PATH Prepending is the result of routing policies programmed in routers and cannot happen by protocol interaction due to loop protection mechanisms.

Figure 1(b) shows how the Autonomous System AS1 is signalling to AS4 to prefer the path through AS2 over the path through AS3 to send traffic to the prefix B/M. As in the case shown in Figure 1(a), this configuration assures a main path and an alternative in case the main path fails. And as in that case, the prefix B/M assigned to the ISP might well be one of the sub-networks $B/M+1$ or $B_1/M+1$. As shown further on, most downstream providers also use AS_PATH Prepending instead of communities when multi-homing to one provider, in order to gain some independence from the upstream providers' policies.

3 Assessing the Room for Optimisation in the Internet's Core Routing Table

The routing table in Internet core routers is a data structure which holds basically two types of data: reachability information and its attributes. The reachability information of an entry is the prefix. The attributes include the next-hop.

This data structure is also known as the Routing Information Base (RIB). The IP packet switching process in a router is controlled by the Forwarding Information Base (FIB). The FIB maps prefixes to their output interface and is generated by combining the RIBs of all active routing processes.

Prefixes are assigned in the public addressing space as per RFC 1930 [12] to Autonomous Systems needing public IP addresses by the different Regional Internet Registries. Prefixes from the public addressing space are biunivocally linked to their Autonomous System Number (ASN) at any given moment in time. In order to detect the configurations presented in Section 2.1 and eliminate sub-netting, the Internet routing table is modelled as a directed graph. The root of the graph is the router the routing table was extracted from and the Autonomous Systems (ASes) are the vertices of the graph. The leaves of the graph are $\{AS, Prefix\}$ pairs that represent the address allocations made by the different Internet Routing Registries (IRRs) to the ASes in their regions.

3.1 A Routing Table Compression Algorithm for the Internet

Algorithm 1 shows the proposed compression algorithm. It is not intended to be applied directly in routers, but rather helps assessing the overhead introduced by current best practises to implement multi-homing over different providers, load balancing, precaution against prefix-hijacking, etc. Therefore, optimisations were not sought and computational time analysis was not performed. As discussed in the conclusion, the aforementioned techniques can be ported to the new proposed architecture or implemented with alternative technologies. The algorithm is applied until no further optimisations can be introduced in the routing table. The concepts of sub- and supernetting are interpreted restrictively, in the sense that prefixes are associated to the AS that originated them and to the Autonomous System Path (AS_PATH) they are received through and sub- or supernetting is only allowed when both prefixes belong to the same AS and are reached through the same sequence of Autonomous Systems. The algorithm uses the following functions to check for possible optimisations:

- `nextAggregation(prefix)` decrements the prefix mask length by one
- `IsFeasible(prefix)` checks whether the prefix is correct as per Definition 1.
- `Contains(`$prefix_1$`,`$prefix_2$`)` checks that both prefixes are originated by the same Autonomous System and that $prefix_2$ is completely contained in $prefix_1$ and that both prefixes are reached following the same AS_PATH.

Algorithm 1 does not affect the reachibility of hosts in the Internet. The BGP-4 routing table of a router is a directed graph. The root of the graph is the router itself and each leaf contains a prefix that can be reached from the root paired with its AS. The other nodes of the graph represent the ASes traversed by a packet on his way to a given prefix. This graph has two types of edges; the regular edges connecting two nodes and the irregular edges connecting a node with itself, which are discarded by the algorithm. Algorithm 1 will only merge two paths of the directed graph if they share all nodes expect the leafs and if the

```
Data: An Internet routing table as an array of {prefix, AS_PATH} pairs
Result: The Internet table with one level of optimisation and a flag indicating
        whether the routing table was modified or not.
changed ← false;
foreach index = 0 to length(InetTable) − 2 do
    this_Prefix ← InetTable[index];
    next_Prefix ← InetTable[index + 1];
    aggregateThis = nextAggregation(this_Prefix);
    if IsFeasible(aggregateThis) then
        if Contains(aggregateThis,nextPrefix) then
            /* remove next_Prefix from the Internet table          */
            removeFromTable(InetTable[index + 1]);
            /* replace this_Prefix with the aggregation           */
            InetTable[index] ← aggregateThis;
            /* signal that the table has changed                  */
            changed ← true;
        end
    end
end
return changed
```

Algorithm 1. Routing table compression algorithm

leafs refer to prefixes that can be aggregated and are assigned to the same AS. The algorithm respects the address allocations made by the IRRs and the paths followed by packets at AS level and thus produces equivalent routing tables, in the sense that packets will arrive to their assigned destinations.

3.2 Status Quo

Algorithm 1 was applied on the routing table contributed by collecting device 203.119.76.3 to the RRC00 repository on the 5th of September, 2009. The initial routing table size was 287,414 routes. After 10 iterations, the algorithm was not able to reduce the table further. The resulting routing table had a size of 185,334 routes. Speaking in relative terms, the table could be reduced by 35.5% without losing connectivity. Figure 2 shows the reduction achieved on the first 6 iterations. It is worth to be noted that the first iteration reduces the routing table size by 22.4%. The improvement in the routing table size is mainly obtained in the first 4 iterations, where a reduction in size by approx. 35% is achieved. Figure 2 shows the compression ratio and the amount of routes suppressed after each iteration. The proposed algorithm suppresses more than 100.000 routes after the third iteration.

4 An Alternative Traffic Engineering Architecture for IP

The previous section shows that there is a significant amount of disaggregated routing information in the Internet's routing table, which is scoped. Thus, for

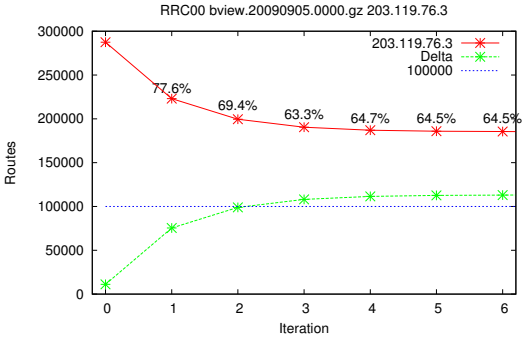


Fig. 2. Reduction achieved on the first 6 iterations for a default free routing table from September 2009

example, the target of the disaggregated information in the configuration shown in Figure 1(a) is the Provider, while in Figure 1(b), the target is a distant AS, which sometimes is not known to the provider. T. Li proposed a new BGP-4 attribute to limit the scope of an advertisement by the number of hops in 2007 in an Internet draft [17], which was later abandoned. In this section, a new architecture with techniques to enforce scopes to BGP-4 advertisements is proposed.

Since the Internet has become a basic infrastructure on which many critical applications rely, any proposal to enhance it must provide smooth migration mechanisms. This is one of the reasons, why the migration to IPv6 is taking so long. The architecture proposed in this paper is backward compatible, non-mandatory and might be adopted incrementally in different regions in the Internet. It is based on the principle that information which is essential for routing purposes should be kept in the Internet's routing table, while the remaining routing information can be migrated to a paralell routing table, which is managed by the Internet Service Providers (ISPs) involved in a certain Traffic Engineering (TE) configuration.

Figure 3 shows the relationship between the different routing tables and the forwarding table in a TE enabled router. The support for the best aggregations would be implemeted by the left side of the figure. These components are part of the current router architectures. Support for interdomain TE routes is added on the right side of the figure and is basically a replica of the current BGP-4 implementation. In order to keep the information exchange for the main routing table isolated from the information exchange for the TE routing table isolated, the use of Multiprotocol Extensions for BGP-4 (MP-BGP) is proposed. MP-BGP is widely used when routers exchange other routing information in addition to pure IPv4 routing information (e.g. IPv6 in RFC 2545 [18]), and when partial IPv4 routing information needs to be kept isolated from the main IPv4 routing table (e.g IPv4 VPN in RFC 2917 [19]).

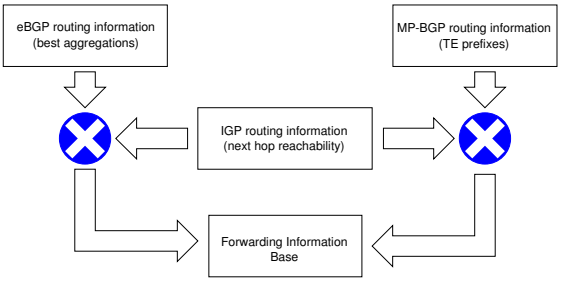


Fig. 3. A high level Traffic Engineering routing architecture

4.1 Advantages

One of the main advantages of the proposed architecture is that the routing setups needed for Traffic Engineering are kept local at the providers which are concerned by them and do not trickle into the Global Internet routing table. Figure 4 shows a routing configuration with unintended side effects which would be avoided with the proposed architecture. Additionally, as shown in Figure 5(a), it would mean going back in time to mid-2006 with regard to routing tables sizes, with a routing table size reduction of approx 33% or 100.000 routes (see Figure 5(b)).

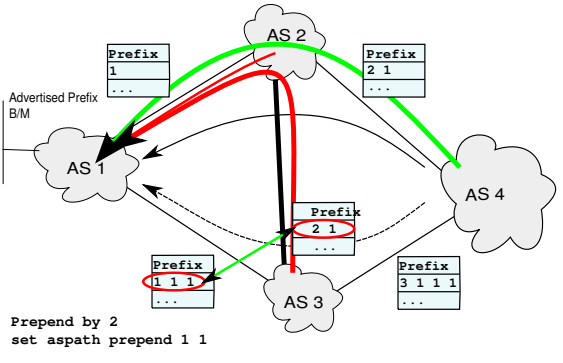


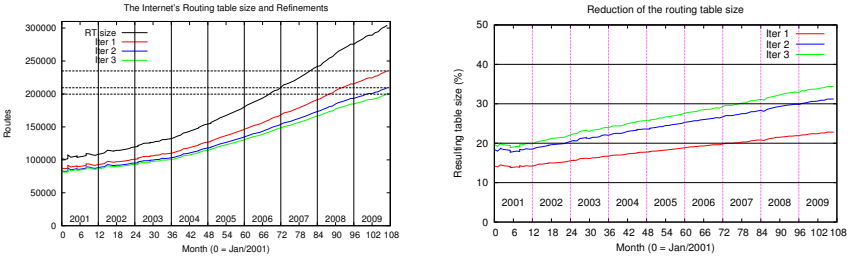
Fig. 4. Unintended side effects of Traffic Engineering using AS_PATH Prepending

There should not be a significant impact on the ability of transit ISPs to implement traffic balancing based on their clients' advertisements. When multi-homing to one provider, as presented in Figure 1(a), the upstream ISP is expected to have enough clients to arrive at a near optimum traffic distribution using the clients' best aggregations. For the case presented in Figure 1(b), AS1 and AS4 have to exchange the MP-BGP information. This can easily be done using multi-hop eBGP configurations (see [11], [14], etc.). AS2 and AS3 do not experience the hassle of having to take into account AS1's routing configurations. Figure 4

shows how, with current configurations, the traffic from a directly connected AS (AS3) can be diverted to another AS (AS2) by the TE routing configurations of AS1.

4.2 Historical Evolution

In order to study the historical evolution of the Internet's default free routing table, data from the RRC00 repository were used. The first routing table dump files of the months between January 2001 and December 2009 were used and the route collector (IP = 203.119.76.3) was selected because it contributed data to all analysed table dump files.



(a) Routing table size on the first 3 iterations

(b) Reduction achieved on the first 3 iterations

Fig. 5. Evolution between January of 2001 and December of 2009

Figure 5(a) shows the evolution of the collector's routing table size and of the resulting routing table after the first three iterations. The horizontal dotted lines mark the size of the December, 2009 routing table size after the first three iterations. They show that after the first iteration, the resulting routing table has the size of the unreduced Internet routing table of beginning of 2008. After the second and third iterations, the resulting size is that of the Internet's routing table of 2007.

One of the measures which can be used to study the trend in the use of subnetting is the efficiency of iteration i defined as $\rho_i = 1 - \frac{N_i}{N_0}$, where N_i is the number of routes in the resulting routing table and N_0 is the initial routing table size. ρ_i is the percentage of routes which could be effectively removed. Figure 5(b) shows that this percentage has been growing steadily since the end of 2001. The relationship of this measure with the "natural" fragmentation of the Internet's routing table is quite unsettling. The process by which the IRRs assign addresses to ISPs implies some level of fragmentation of the routing table. But if ISPs were not using subnetting for Traffic Engineering purposes, the reduction algorithm would not be able to reduce the routing table: Between successful addressing space allocations to a specific ISP, IRRs continue allocating addressing space to other ISPs. The probability that an AS gets adjacent and aggregatable addressing space in two consecutive applications is almost zero.

This confirms the effectiveness of the compression algorithm in suppressing subnetting introduced by the ASes with configurations akin to Figure 1(a) while preserving connectivity.

5 Related Work

The algorithm proposed in this paper shows that there is room for optimisation in the Internet's routing table, which would result in improved scalability and manageability. Fall et al. [7] have recently studied the impact of the size of the Internet's routing table on cost and CO_2 footprint and conclude that Moore's law will be able to cope with the growth of the routing table size. The architecture proposed in this paper reduces the complexity of the routing table and of the dynamics which can be linked to TE techniques, with the objective of reducing the OPEX of ISPs.

Other work related with the compression of the Internet's core routing table includes the virtual aggregation proposal ViAggre [3]. Its main drawback is , as the authors recognise, that they manipulate the routing tables and the results might divert the traffic to different paths, even at the Autonomous System level. Freedman et al. also study the aggregation level of the Internet's routing table in [8]. They conclude that geographic dispersion of IP prefixes reduces the level of compression which can be achieved when looking for the best aggregations in the Internet's routing table. The work described in the present paper shows that, limiting aggregation prefixes which share the same sequence of ASes produces significant savings. The aggregated prefixes fall under the multi-homing scenario depicted in Figure 1(a) and are likely to be geographically adjacent. This adjacency should be dealt with within the different Internet Service Providers and not affect the Internet's core routing table.

Suri et al. [22] study the compression of routing tables which take into account the source and destination address fields in IP packets. This kind of routing tables is significant for edge devices of the Internet like access routers. The approach of this paper differs in two main aspects from the approach proposed by Suri: firstly, this paper concentrates on core devices in the Internet, where routing is done based on the destination address only. Secondly, while Suri's approach can be implemented in routing devices, the algorithm presented in this paper is intended for assessing the amount of optimisation which is achievable.

Routers have evolved to complex systems with multiple routing protocols building concurrently the routing table [9,11,14]. These devices have a FIB and multiple RIBs. Draves et al. [6] propose an algorithm to compress the routing table which is better applicable to the FIB of a router than to the RIB. This is so, because they only take into account the destination address and the next hop. While their proposed algorithm provides a highly efficient tree to determine the next hop, essential information regarding the path followed by the packets is lost. Therefore, Draves' algorithm is incompatible with BGP-4.

Some level of de-aggregation in the Internet's routing table is used to protect ASes against prefix hijacking [15,1]. In this scope, the INTERSECTION project

[2] proposes alternative techniques to detect and provide remedy in case of prefix hijacking. In general, prefix hijacking should not be remedied by introducing additional, more specific routes to the Internet's table but by filtering: if an AS advertises a prefix it does not own, the upstream providers should it filter out and notify the offending AS. Systems like INTERSECTION allow for quick detection of suspicious prefixes. PHAS [16] is aligned with the INTERSECTION approach.

6 Conclusion and Further Work

This paper shows that ISPs massively use subnetting techniques as part of their Traffic Engineering (TE) implementations by using an algorithm to find best aggregations in the Internet's routing table, which achieves 33% optimisation rates on current routing tables. This algorithm, however, is not intended to be applied directly on routers. This paper also shows that the trend in the current Internet is to increase the use of TE techniques and therefore to decrease the optimisation of the routing table. To remedy this and render the infrastructure more controllable, an alternative TE architecture for IP networks is proposed.

Further work on this architecture includes modelling and simulating this architecture, and studying its impact on proposed IPv6 transition mechanisms and on the architecture of the IPv6 Internet in the long run.

Acknowledgement

This work was made possible by the extensive BGP-4 data collections of the RIPE's Routing Repository.

References

1. YouTube Hijacking: A RIPE NCC RIS case study, http://www.ripe.net/news/study-youtube-hijacking.html
2. INTERSECTION (INfrastructure for heTErogeneous, Resilient, SEcure, Complex, Tightly Inter-Operating Networks) (January 2008), http://www.intersection-project.eu/ (last visit June 25, 2010)
3. Ballani, H., Francis, P., Cao, T., Wang, J.: Making routers last longer with ViAggre. In: NSDI 2009: Proceedings of the 6th USENIX Symposium on Networked Systems Design and Implementation, pp. 453–466. USENIX Association, Berkeley (2009)
4. Green, B.R., Smith, P.: CISCO - ISP Essentials. Cisco Press (September 2002)
5. Cisco Systems Inc. Interworking technology handbook
6. Draves, R., King, C., Venkatachary, S., Zill, B.D.: Constructing optimal ip routing tables. In: Proc. IEEE INFOCOM, pp. 88–97 (1999)
7. Fall, K., Iannaccone, G., Ratnasamy, S., Godfrey, P.B.: Routing Tables: Is Smaller Really Much Better? In: Proceedings of Hotnets 2009. ACM, New York (2009)
8. Freedman, M.J., Vutukuru, M., Feamster, N., Balakrishnan, H.: Geographic locality of ip prefixes. In: IMC (2005)

9. Gredler, H., Goralski, W.: The Complete IS-IS Routing Protocol. In: Computer Science. Springer, London (2005)
10. Griffin, T.G., Wilfong, G.: An analysis of BGP convergence properties. In: Proc. of SIGCOMM 1999, pp. 277–288. ACM Press, New York (1999)
11. Halabi, S.: Internet Routing Architectures, 2nd edn. Cisco Press (2000)
12. Hawkinson, J., Bates, T.: Guidelines for creation, selection, and registration of an Autonomous System (AS). RFC 1930 (Best Current Practice) (March 1996)
13. Huston, G.: Analysing the Internet BGP Routing Table. The Internet Protocol Journal 4(1) (2001)
14. van Beijnum, I.: BGP - Building Reliable Networks with the Border Gateway Protocol. O'Reilly, Sebastopol (2002)
15. Kapela, A., Pilisov, A.: Stealing the Internet. DefCon August 16 (2008) (last visit, July 17, 2009)
16. Lad, M., Massey, D., Pei, D., Wu, Y., Zhang, B., Zhang, L.: Phas: A prefix hijack alert system (2006)
17. Li, T., Fernando, R., Abley, J.: The AS_PATHLIMIT Path Attribute (2001), http://tools.ietf.org/html/draft-ietf-idr-as-pathlimit-03 (last visit: January 17, 2010)
18. Marques, P., Dupont, F.: Use of BGP-4 Multiprotocol Extensions for IPv6 Inter-Domain Routing. RFC 2545 (Proposed Standard) (March 1999)
19. Muthukrishnan, K., Malis, A.: A Core MPLS IP VPN Architecture. RFC 2917 (Informational) (September 2000)
20. Networks, J.: Examine BGP Routes and Route Selection in Juniper routers (last visit December 12, 2009)
21. Rekhter, Y., Li, T., Hares, S.: A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard) (January 2006)
22. Suri, S., Sandholm, T., Warkhede, P.: Compressing two-dimensional routing tables. Algorithmica 25, 287–300 (2003)

End-to-End Performance Evaluation of Virtual Networks Using a Prototype Implementation

Asanga Udugama[1], Liang Zhao[1], Yasir Zaki[1], Carmelita Goerg[1], and Andreas Timm-Giel[2]

[1] Communications Networks, TZI, University of Bremen
Otto-Hahn-Allee, 28359 Bremen, Germany
{adu,zhaol,yzaki,cg}@comnets.uni-bremen.de
[2] Institute of Communication Networks, Hamburg University of Technology,
Schwarzenbergstr. 95E, 21073 Hamburg, Germany
timm-giel@tuhh.de

Abstract. Network virtualization is a concept where physical resources are used to create virtual resources that are combined to form virtual networks. As one of the key enablers of the future Internet, network virtualization solves a number of issues associated with today's networks. Concepts of network virtualization that are not restricted to virtualization technology, termed as overall concepts that include roles of parties and deployment aspects are currently being defined in different research activities. An area that lack attention is the evaluation of performance in prototypes that consider these overall concepts of network virtualization. The work presented here discusses and presents the performance in a network virtualization prototype that considers these overall concepts.

Keywords: Network Virtualization, Infrastructure Provider, Virtual Network Operator, Virtual Network Provider, Virtual Resources, FP7 4WARD Project.

1 Introduction

Network virtualization (Fig. 1) is the concept of using physical resources to create virtual resources that are formed into virtual networks. These networks can be brought up on-the-fly to serve different requirements in very short time frames. Transient networks for such purposes as conferences or emergency management can ideally utilize VNets to setup short lived networks that serve a required purpose. VNets have a number of advantageous properties such as isolation, dynamic provisioning and security.

As one of the most important enablers of the future Internet, network virtualization has received a higher level of research attention all over the world, e.g. VINI [1], GENI [2] and PLANETLAB [3] in US; AKARI [4] and AsiaFI [5] in Asia; 4WARD [6] in Europe.

The 4WARD project was formed to undertake research on the architecture of a future Internet adopting a "clean slate" research approach. This approach temporarily ignores the practical constraints of evolving from the existing TCP/IP-based network architecture in the interest of identifying a design that is appropriate to the present and

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 38–49, 2011.

expected future usage and is not forced to adapt to architectural decisions made some thirty years ago with quite different objectives and constraints. The main concepts of 4WARD include a "Generic Path" allowing inherent support of mobility, multipath connections and network coding, new addressing paradigms, referred to as "Network of Information" and network virtualization enabling coexistence of several legacy and new networking concepts on the same infrastructure. A prototype has been developed to demonstrate and evaluate the concepts developed for network virtualization in this research. This work presents the framework of this prototype and its performance.

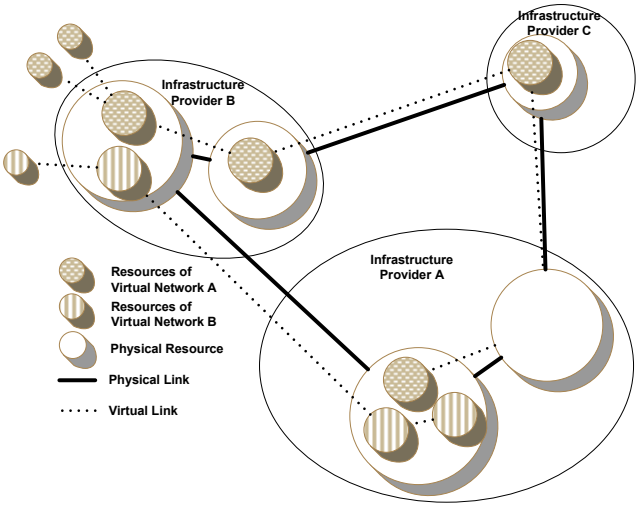


Fig. 1. Network Virtualization: Virtualized physical resources are connected together to form virtual networks

When considering prototyping and performance evaluation on network virtualization, especially on XEN-based [13] solutions, a number of activities can be found. A number of exemplary research works that have been recently published are taken and grouped them into the specific topic areas they address (I/O performance, resource scheduling, and virtual routing). [7] compares the streaming performance of open source hypervisors like XEN and OpenVZ in a Linux environment. [8] also gives a quantitative analysis on popular hypervisors, e.g. XEN, VMware and KVM, to compare their performance in terms of application test, disk test and I/O test, where the issue of scalability (meaning the ability of hosting multiple virtual machines in one physical machine) is also discussed since isolation among virtual networks is important to guarantee the performance of each. The authors of [9] and [10] discuss the performance degradation of network interfaces due to virtualization and the large receive offload (LRO) concept is employed as an effective method to improve the I/O performance. And similarly, in order to enhance the I/O performance, [11] proposes a command-line tool in the XEN environment to do the bandwidth reservation on XEN virtual machines.

When considering the above mentioned research, it is clear that they focus on specific aspects of network virtualization, leaving out the overall operations of virtual networks. That is the key difference in the work presented in this contribution. This prototype is able to present the creation and the operation of complete virtual networks. It includes the operation, management and visualization of virtual network related activities by the different parties involved in the whole network virtualization architecture. It is able to create wired as well as wireless QoS guaranteed virtual networks on the fly, initiated by the Infrastructure Providers (InPs) based on the requirements of the Virtual Network Operators (VNet Operators) who will finally make services available for their clients.

The sections that follows focuses on provisioning of VNets utilizing end-to-end infrastructure, reflecting on the concepts of VNets, revealing the bottlenecks when realizing multiple VNets and recommendations from the experiences. Specifically, the immediately following section explains the concepts and processes involved in network virtualization briefly. Then it moves on to explaining the architecture of the software framework that was developed. This is followed by an explanation of the scenarios considered together with the test-beds on which these scenarios are tested. The subsequent section explains and provides an analysis of the results taken from the above scenarios. The last section is a concluding summary that includes a look at the future direction of this work.

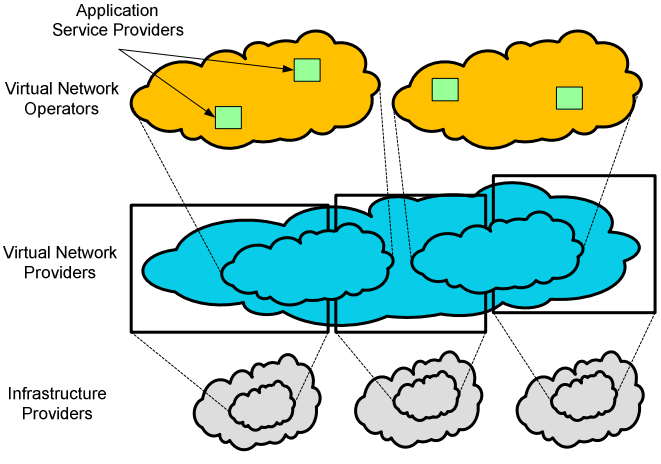


Fig. 2. Future Internet: Roles of the different parties associated with network virtualization

2 Network Virtualization Architecture of the Future Internet

Today, the Internet is operated by multiple Infrastructure Providers (InP). Similarly, the future Internet can be assumed to have a similar status-quo where large organizations that own and operate infrastructure enabling communications between different locations and providing connectivity for end users [12]. These InPs create virtual resources from their physical infrastructure for other parties to configure and use. Unless this aspect of inter-domain QoS guarantees is considered, intra-domain

QoS guarantees may seem meaningless due to bottlenecks of different InPs. Therefore in the future Internet, an intermediary called a VNet Provider will negotiate with different InPs and obtain virtual network slices that are in turn combined and/or sub-sliced to be made available for VNet Operators based on their requirements.

Fig. 2 shows the relationships of the different roles of the parties in the architecture of the Future Internet. Once the VNet is made available by the VNet Provider in the requested topology, the VNet Operator can setup and offer the network for its customers (Application Service Providers and end users). Setting up by the VNet Operator includes activities such as installing the different protocols in the VNet and configuring it to suit the requirements of its customers. The VNet Operator has full control over the VNet and the InP may not necessarily know what is in the VNet and how it is operated.

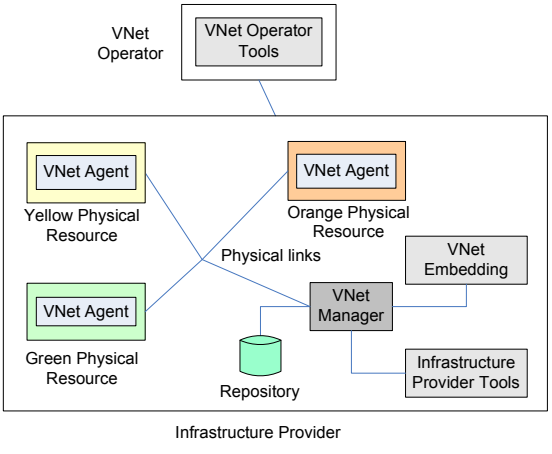


Fig. 3. VNet Management: Component connections of a VNet Management environment

3 VNet Management Prototype

The VNet Management (VNM) prototype is built to evaluate and demonstrate the VNet concepts and processes described in the previous section. This is built to operate on IP based networks and utilizes XEN [13] as the primary virtualization technology. The typical architecture of such a VNM environment consists of agents that reside in different physical resources and control the resource to create, remove or modify virtual resources. These agents are controlled by managers located centrally or distributed, which accept commands from an InP to manage the physical resources. The managers hold repositories that are continuously updated based on the current status of the VNM environment. The interactions of agents and managers are similar to the operations of SNMP, but differ from SNMP due to its independence from specific networking technologies (such as IP).

Fig. 3 shows an example of a VNet management environment consisting of 3 physical resources (colored boxes). Each physical resource has a VNet Agent running in it to manage the virtual resources instantiated within the physical resource. The

VNet Manager controls all the VNet Agents and the repository. Using a set of tools, the InP initiates management requests based on the requests of the VNet Operators.

3.1 VNet Manager

VNet Manager (Fig. 4, left) coordinates all the activities of the whole VNet environment. It is a multi-threaded daemon that can reside in the hardware of the InP and communicate with the other components of the environment. It has the following functional activities:

- Handling of requests and information communicated with VNet Agents and InP management tools
- Handling of synchronized updates to the VNet Repository which serves as the storage for the configurations of the VNets

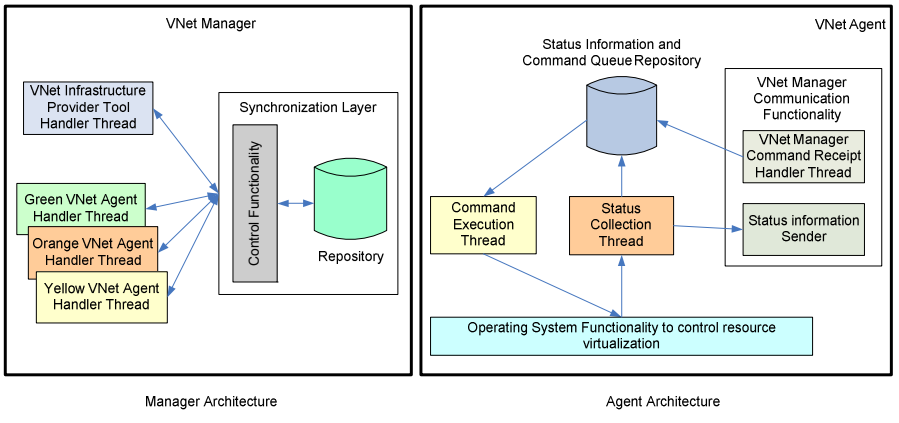


Fig. 4. VNet Component Architecture: The modular architecture of a Manager and an Agent

All the InP management tools use the same interface in this prototype to send and receive data to the VNet Manager. Each tool instance started will result in a separate process being created in the VNet Manager to handle the requests and return information. The communication between the user tools and the VNet Manager is based on TCP sockets in this prototype, to maintain reliable communications.

The connections of VNet Agents also have a similar concept where each VNet Agent instance will result in a process being created to handle the sending and receiving of information to the VNet Agent. The interfaces to the agents have the same format for every agent and uses TCP sockets for communications.

The control functionality which is protected (made thread safe) for synchronized access provides the following functionality to the different processes created:

- Handle command actions received from the InP tool processes
- Command VNet Agents to perform activities
- Handle information received by the VNet Agents through the processes
- Send information to the InP tools
- Update the repository

3.2 VNet Agent

VNet Agents manage the actual virtualization of physical resources. A VNet Agent is executed on each physical resource. Each VNet Agent is specific in its functionality to the resource under its control. Fig. 4 (right) shows the generic architecture of a VNet Agent.

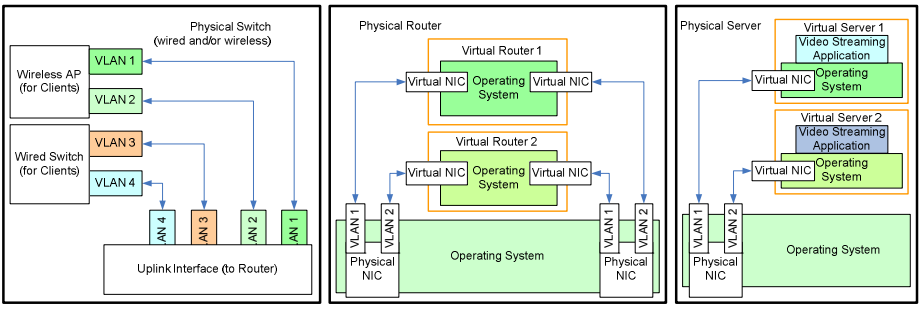


Fig. 5. Physical Device Configurations: Virtual resources and their connections in physical resources to form virtual networks

Every VNet Agent has an interface with the VNet Manager to get commands to create, remove, modify, bring up or shut down virtual resources associated with a given Agent. There are a number of processes in an instance of an Agent which perform the following tasks:

- Queue incoming virtual resource management instructions
- Execute these instructions in a FIFO manner
- Retrieve current status information
- Notify current status information

The operating system functionality component in the Agent architecture is the component that holds functionality that is unique to each of the resources. That means, an Agent for a Wireless Access Point (WAP) of a particular hardware manufacturer differs from the Agent for a WAP of another hardware manufacturer in its operating system functionality component. There are a number of different VNet Agent types supported by the VNet environment.

The VNet Agent for Servers manages the virtual servers that can be created, removed, brought up, etc. on a physical server with a single network attachment. Fig. 5 (right) shows a VNet Agent for a physical server that has created two virtual images of servers.

The virtualization at the link level is achieved through the use of Virtual Local Area Networks (VLANs) where each VNet is carried with a separate VLAN ID.

The VNet Agent for Routers controls physical routers to manage the virtual environment. A physical router consists of multiple network interfaces to perform the routing. These interfaces are bridged to the virtual router instances to perform the routing. VNets are realized using VLANs. Fig. 5 (middle) shows a router that has two virtual routers instantiated. VNet Agents for Switches/Wireless Access Points use

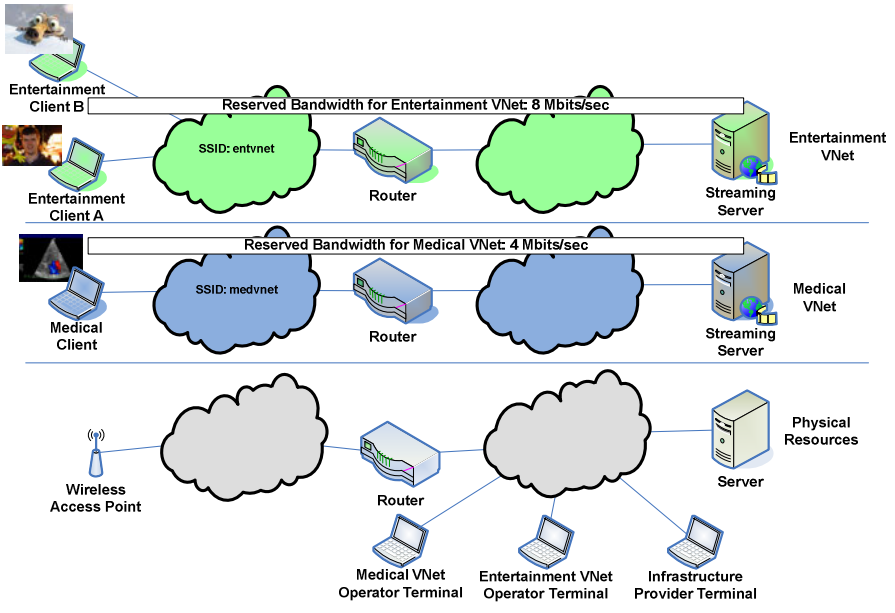


Fig. 6. Wireless Scenario: Test-bed setup for the scenario with 2 simultaneous VNets

VLAN and the capability of multiple virtual SSIDs to distribute traffic to different virtual network. Fig. 5 (left) shows a physical Wireless Access Point handling network traffic related to four VNets.

4 Scenarios and Test-Beds

The evaluation of the VNM prototype is done using 2 test-beds that focus on 2 different evaluation scenarios. The hardware used in these test-beds is installed with the VNM prototype.

4.1 Wireless Scenario

The first scenario, referred to as the Wireless Scenario (Fig. 6) considers an InP who offers hardware for VNet Operators to create virtual networks. Two such operators use this service and request the InP to create 2 specific virtual networks with differing QoS characteristics. In this case the role of the VNet Provider is left out due to the use of only one InP. The first of these networks is created for an operator who serves a medical organization's requirements by providing a secure network with a guaranteed bandwidth to carry video and vital signs data of remote surgical operations.

The second virtual network operator maintains an entertainment network that hosts an Application Service Provider providing movie streaming services to its customers. The customers of both of these VNet Operators connect over wireless networks (WLAN).

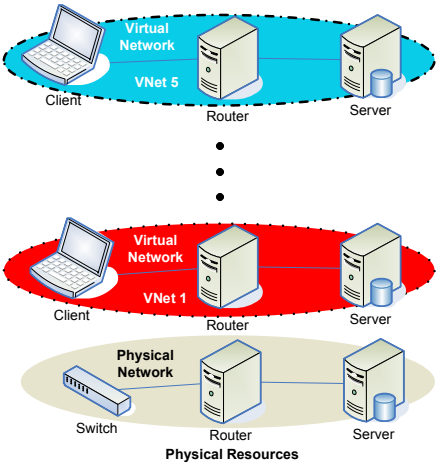


Fig. 7. Wired Scenario: Test-bed setup for the scenario with multiple simultaneous VNets

The Medical VNet Operator requests a VNet with a 4 Mbit/s guarantee while the Entertainment VNet Operator requests a VNet with an 8 Mbit/s guarantee. The management terminals of the VNet Operators which are connected to the InP's network requests InP to create the required VNets in the configuration they require. In this case, both operators request for a network with a server, a router and a wireless access point. Each of the clients that connect request for 5 Mbit/s UDP streams from the respective server of the VNet, which hosts the service. In this scenario, the 2 VNets are created sequentially, where the Medical VNet is started before the Entertainment VNet, with a time difference of about 100 seconds. The 5 Mbit/s UDP streams, one in the Medical VNet (MedVNet) and 2 in the Entertainment VNet (EntVNet 1 and EntVNet 2) are started in around 150 second intervals.

The main purpose of this scenario is to show the overall operation of the Network Virtualization architecture with the different activities involved in establishing and using the VNets.

4.2 Wired Scenario

The second scenario called the Wired Scenario (Fig. 7) setup on the second test-bed is similar to the first scenario but is intended to evaluate the performance of multiple VNets with an emphasis on increasing the VNets until the performance degrades. This test-bed is equipped with commercially available off-the-shelf hardware. The equipment consists of Quad-core hardware and Giga-bit links connecting them. To leave out the effects of wireless networks, a completely wired setup is considered in this scenario. In this scenario, the VNets are created in 100 second intervals and each of the clients associated with a VNet is attempting to download 150 Mbit/s stream from its server.

5 Performance Results and Analysis

Fig. 8 shows an example of the throughput performance of the 3 VNet streams on each of the clients. It also shows how the physical hardware performs. Since the Medical VNet has a bandwidth limitation of 4 Mbit/sec over the Ethernet link, the 5 Mbit/s input stream is reduced to around 4 Mbit/s. In addition, it will not reach the exact 4 Mbit/sec level due to packet losses. Since a lost packet means a loss of a number of bytes, the stream will not be able to reach the 4 Mbit/s level. The packet losses graph (Fig. 9, left) shows how the packets for each VNet stream are lost. What could be noticed is that when the second stream of the Entertainment VNet (EntVNet 2) is started, both of the two other streams became very unstable. This is due to the insufficient bandwidth available for both streams. But, the performance of the MedVNet continues proceed unaffected due to the given bandwidth guarantee.

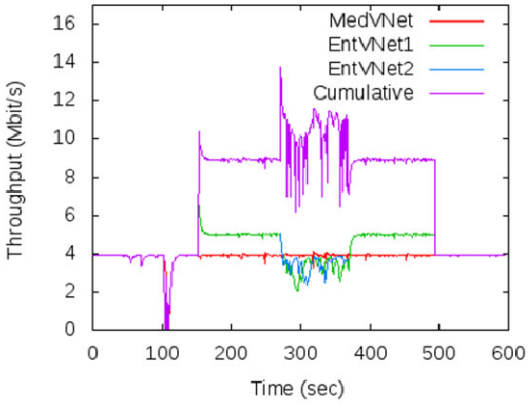


Fig. 8. Wireless Scenario: Throughput performance of VNets with UDP streams

The drastic throughput drop that is visible around 100 seconds in Fig. 8 is due to the creation of the second Entertainment VNet. This is specifically due to the creation of the second WLAN SSID which makes the WAP halt the data flows on other active SSIDs until the new SSID is configured.

This experiment was done in an environment where other WLAN APs not related to the experiment were in operation. Hence, all transmissions are influenced by the activities of the other APs. The constant fluctuations that are visible in graphs are a sign of that.

Fig. 9 (right) shows how bandwidth limitations can influence the jitter of the streams. As can be seen, the jitter is higher for the MedVNet while the EntVNet streams have lower jitter values. This is due to the operation of the queues of traffic shaping. It was seen that when the bandwidth limit (e.g. 4 Mbit/s in the MedVNet) is close to the utilization of the bandwidth (e.g. 5 Mbit/s in MedVNet), a higher jitter is shown compared to the opposite.

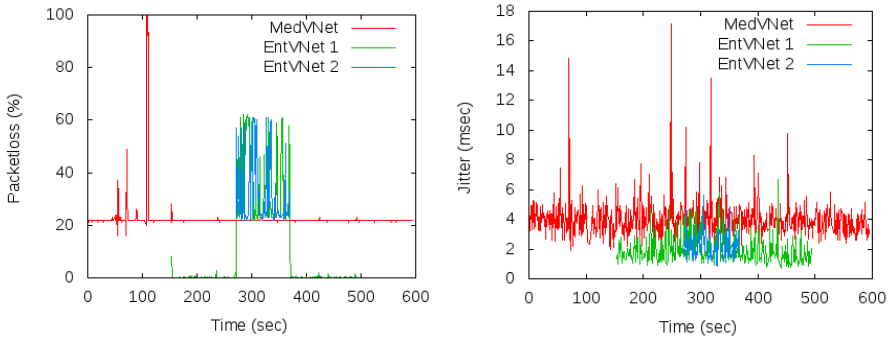


Fig. 9. Wireless Scenario: Packet loss and Jitter with UDP streams

The second set of experiments (based on the Wired Scenario) puts the emphasis on stress testing the test-bed, which is made out of commercially available hardware. The stress test mainly looks at how much of VNets can be created in this hardware and looks at what the performance statistics show.

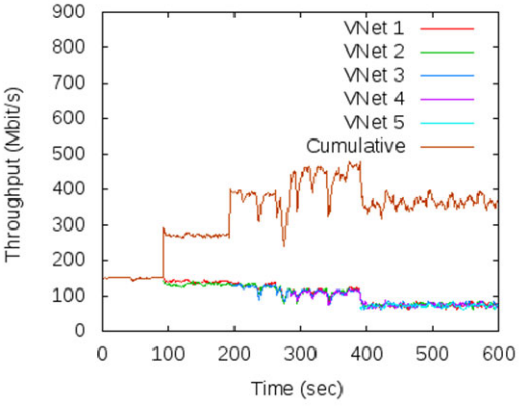


Fig. 10. Wired Scenario: Throughput performance of 150 Mbit/sec UDP streams

Throughput performance of this setup (Fig. 10) shows the effects of creating multiple VNets using the same infrastructure. Up to 2 VNets can be operated without any degradation of throughput but when the 3rd VNet is started, the throughput of all the active VNets starts fluctuating. This is in spite of using Giga-bit Ethernet links between them. When the 5th VNet is created, the throughput of all the active VNets simply start to decrease and experiences close to 50% packet losses (Fig. 11, left). The jitter values of these streams (Fig. 11, right) show the same tendency where fluctuations become more visible starting from the 3rd VNet.

A look at the resource utilization of all the hardware used to create the VNets show that CPU utilization is not an issue with the start of the 3rd VNet and the corresponding 150 Mbit/sec stream. Since other resources such as disk space and link capacities are

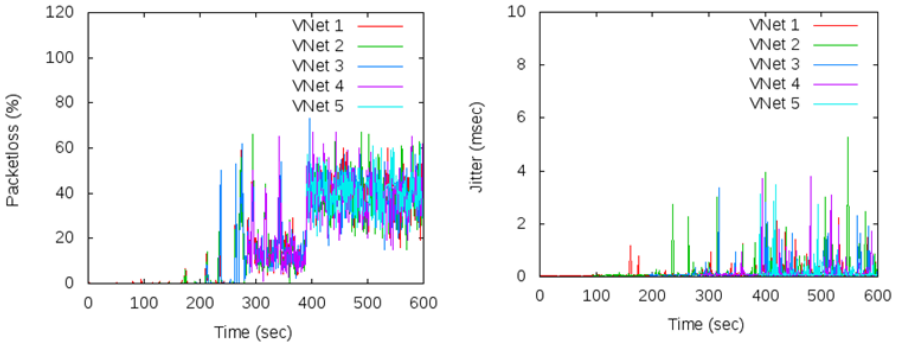


Fig. 11. Wired Scenario: Packet loss and Jitter with 150 Mbit/s UDP streams

sufficiently available, the degraded performance points to inefficiencies in the way the Hypervisor schedules the data in different queues related to communications.

6 Conclusions

Network virtualization is considered as a key enabler of the Internet of the future. A number of projects are developing concepts and processes related to network virtualization. A noticeable drawback of those works is that they attempt a piecemeal evaluation of network virtualization and thereby leaves out certain aspects when evaluating the others. The prototype that was built in this work attempts at focusing on all the aspects starting from the roles of the parties involved to the technology used. The experiences gained through this prototype, points to the following conclusions.

- Network virtualization appears to be feasible even in an overall setting where all aspects (from roles of the parties to virtualization technologies) are considered
- Use of commodity hardware in the scenarios point to the feasibility of using them for creating multiple virtual networks (does not necessarily require high end purpose-built hardware)
- Virtualization environments used (e.g. XEN in this prototype) needs to be enhanced further to improve on scheduling related to networking traffic
- Sufficient resource margins must be considered from a deployment perspective to ensure optimum performance

The work done with this prototype will be extended further to consider scenarios that increase the number of physical resources used.

Acknowledgments

The work presented here was done during the FP7 4WARD project funded by the European Commission. The authors wish to thank all the partners of the FP7 4WARD project, especially the partners of the network virtualization work package who have contributed in terms of discussions and ideas to this work. Further, the authors wish to thank Nikola Zahariev for his assistance during this work.

References

1. Bavier, A., Feamster, N., Huang, M., Peterson, L., Rexford, J.: VINI Veritas: Realistic and Controlled Network Experimentation. In: Proc. ACM SIGCOMM 2006 (September 2006)
2. GENI: Global Environment for Network Innovations, http://www.geni.net
3. Bavier, A., Bowman, M., Culler, D., Chun, B., Karlin, S., Muir, S., Peterson, L., Roscoe, T., Spalink, T., Wawrzoniak, M.: Operating System Support for Planetary-Scale Network Services (March 2004)
4. Architecture Conceptual Design for New Generation Network (akari-project.nict.go.jp)
5. Asia Future Internet (AsiaFI), http://www.asiafi.net
6. Niebert, N., et al.: The Way 4WARD to the Creation of a Future Internet. In: PIMRC 2008, Cannes, France (September 2008)
7. Sukaridhoto, S., Funabiki, N., Nakanishi, T., Pramadihanto, D.: A comparative study of open source softwares for virtualization with streaming server applications. In: ISCE 2009, Kyoto, Japan (May 2009)
8. Xu, X., Zhou, F., Wan, J., Jiang, Y.: Quantifying Performance Properties of Virtual Machine. In: ISISE 2008, Shanghai, China (December 2008)
9. Fumio, N., Hitoshi, O.: Optimizations of Large Receive Offload in Xen. In: NCA 2009, Cambridge, MA (August 2009)
10. Hitoshi, O., Fumio, N.: Performance Analysis of Large Receive Offload in a Xen Virtualized System. In: ICCET 2008, Singapore (February 2008)
11. Zhang, J., Li, X., Guan, H.: The Optimization of Xen Network Virtualization. In: CSSE 2008, Wuhan, China (December 2008)
12. 4WARD, http://www.4ward-project.eu
13. Williams, D.E.: Virtualization with Xen: Including XenEnterprise, XenServer, and XenExpress, Syngress (2007)

Addressing Stability in Future Autonomic Networking

Timotheos Kastrinogiannis[1], Nikolay Tcholtchev[2], Arun Prakash[2], Ranganai Chaparadza[2], Vassilios Kaldanis[3], Hakan Coskun[2], and Symeon Papavassiliou[1]

[1] Institute of Communications and Computer Systems (ICCS), School of Electrical and Computer Engineering, National Technical University of Athens, Athens 15780, Greece
timothe@netmode.ntua.gr, papavass@mail.ntua.gr
[2] Fraunhofer FOKUS Institute for Open Communication Systems, Berlin, Germany
{nikolay.tcholtchev,arun.prakash,ranganai.chaparadza, hakan.coskun}@fokus.fraunhofer.de
[3] VELTI S.A. – Mobile Marketing & Advertising, Athens, Greece
vkaldanis@velti.com

Abstract. When considering autonomic networking, where multiple self-* functionalities, in terms of node-wide or network-wide control loops, must operate, interact and proficiently collaborate, stability problems inherently arise due to the distributed nature of the decision making process and autonomic nodes interactions towards enabling various self-* functionalities, along with the stochastic nature of the networking environment. This article provides a systematic, concrete view of stability in autonomic networks design. It aims at identifying and categorizing fundamental autonomic networks' architectural and designing issues that cause or affect stability, highlighting and discussing corresponding solutions and thus, providing theoretic tools for analyzing and treating them. As a reference model we adopt Generic Autonomic Network Architecture (GANA), a holistic framework for autonomic networks engineering.

Keywords: Autonomic Networks, Stability, Control Loops.

1 Introduction

The vision of Future Internet, is of a self-managing network whose nodes/devices are designed/engineered in such a way that all the so-called traditional network management functions, i.e. FCAPS framework (Fault, Configuration, Accounting, Performance and Security) [1], as well as the fundamental network functions (e.g. routing, forwarding, monitoring, mobility, resource management, e.t.c.) are enhanced with autonomic attributes. In such an evolving environment, it is envisioned the network itself to detect, diagnose and repair failures, as well as to constantly adapt its configuration and operations aiming at optimizing its performance.

In general, towards realizing networks' autonomic behaviors, the presence of control-loops in the system is essential. Inputs to a control loop consist of various status signals, information and views continuously exposed from the system, component(s) or resource(s) being controlled (e.g. protocols, nodes, functionalities, etc.), along with (usually policy-driven) management rules that orchestrate the behaviour of the system or component(s). Outputs are commands to the system or component(s) to adjust its operation, along with status to other autonomic systems.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 50–61, 2011.

Practically, control loops consist of iterative and (most of the time) distributed algorithms, that enable various node's self-* behaviours and hence, guide them to act in line with their own optimization goals or towards achieving global optimal network objectives. Henceforth, future autonomics envisions the aggregation of node-scoped control loops, i.e. within a node, in terms of interacting intra/inter-node control loops or triggered/managed low level control loops by higher level control loops within the node or the network. Intuitively, the above view leads to a hierarchal control loops paradigm that enables the efficient design of autonomic nodes and architectures [2].

As the designers of a network's architecture increase its autonomic attributes, in terms of introducing various self-* functionalities (i.e. control loops) at node or network level, the inherent issue of stability becomes more and more complex and crucial. In general, the stability of dynamic systems has been extensively studied since the beginning of *Control Theory*, but when viewing it through a dynamic networking environment, additional drawbacks and challenges emerge [3].

In most cases, stability is defined as a property of a (dynamic) system or element through which it is reassured (it reassures) that its output will ultimately attain a steady state (i.e. the state when the recently observed behaviour of the system will continue into the future). When considering an autonomic networking environment, where multiple self-* functionalities must operate, interact and collaborate, in terms of node or network wide control loops, its stability implies reaching an equilibrium point over a finite time frame. In other words, the system/network should be stable in the sense that parameter values change, however remaining bounded in a small neighbourhood of a final value. This is especially important since self-* algorithms mainly run completely automatic and without the possibility of manual intervention.

Despite the vital role of stability in future autonomic networking, there is a lack of a concrete framework and corresponding theoretic and designing tools for addressing and treating autonomic networks' stability. Recent attempts to address the latter are either closely coupled to the studied networking environment [4] or the autonomic functionality that is engineered [5]. Towards enlightening the evolutionary path of future autonomics via highlighting the significance and role of stability in such dynamic networking paradigms, this paper makes the subsequent contributions:

- Identifies, categorizes and discusses the key factors that affect autonomic networks' stability attributes from both theoretic and designing point of view.
- For each one, corresponding theoretic tools and concrete methodologies for studying and treating them are provided.
- Proposes a concrete framework for addressing both off-line (network designing phase) and on-line (network runtime phase) aspects of stability.

The rest of this paper is organized as follows. In section 2, fundamental designing principles and the GANA model are highlighted. Section 3 highlights the key principles of the proposed methodology for addressing stability in autonomic networks. Section 4, identifies, analyses and provides solutions to fundamental autonomic networks' stability issues, concerning both theoretical and architectural aspects at design time. Then, section 5, illustrates the concepts of action synchronization functions and self-stabilization monitoring towards tackling the issue of runtime autonomic networks' stability. Finally, section 6 concludes the paper.

2 An Overview of the Generic Autonomic Network Architecture

The concept of "autonomicity" – realized through control-loop structures operating within network nodes/devices and the network as a whole, is an enabler for advanced and enriched self-manageability of network devices and networks. To address the lack of such autonomic notions formalism, in both theoretic as well as pragmatic designing levels, an evolvable, standardizable architectural Reference Model for Autonomic Networking and Self-Management within node/device and network architectures, dubbed the Generic Autonomic Network Architecture (GANA) [2], has recently emerged. The central concept of GANA is that of an autonomic Decision-Making-Element ("DME" or simply "DE" in short—for Decision Element). A Decision Element (DE) implements the logic that drives a control-loop over the "management interfaces" of its assigned Managed Entities (MEs). Therefore, in GANA, self-* functionalities such as self-configuration, self-healing, self-optimization, etc, are implemented via a Decision Element(s).

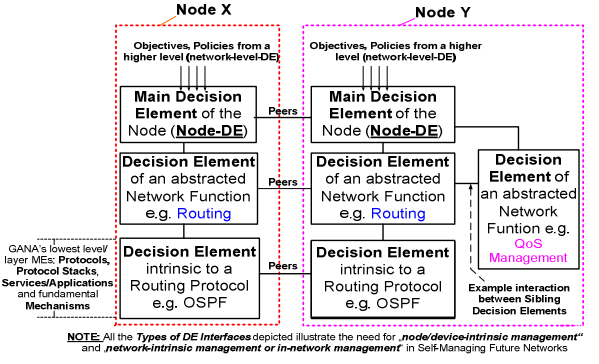


Fig. 1. Illustration of hierarchical, peering, sibling relations and interfaces of DEs in GANA

Specifically, GANA introduces "autonomic manager components" i.e. Decision-Elements (DEs), meant to operate at four different abstraction levels of functionality. These "autonomic manager components" are designed following the principles of "hierarchical", "peering", and "sibling" relationships among each other within a node or network. GANA defines four hierarchical levels of abstractions for which DEs, MEs, Control-Loops and their associated dynamic adaptive behaviours can be designed, capturing a "holistic view of interworking autonomic and self-management levels". Thus, the levels of (DEs) abstractions are as follows:

Level-1: protocol-level (the lowest level) by which self-management is intrinsically designed and associated with a network protocol itself (whether monolithic or modular) [6], [7].
Level-2: the abstracted function-level (directly above the protocol(s)-level) that abstracts some protocols and mechanisms associated with a particular function e.g. "routing function", "forwarding function", "mobility management function", etc.
Level-3: the level of the node/device's overall functionality and behaviour i.e. a node or system as a whole, is also considered as level of self-management functionality.

Level-4: the level of the network's overall functionality and behaviour (the highest level). Network-Level-DEs are characterized by the following attributes: (1) they are the ones with wider network-wide view to perform sophisticated decisions e.g. network optimization; (2) they are logically centralized to avoid processing overhead in managed nodes with having distributed decision logic in network elements; (3) they are the ones that provide an interface for a human to define goals and objectives or policies e.g. business goals to the autonomic network.

Fig. 1 illustrates a GANA aware autonomic node, whereby lower level Decision-Making-Elements, operating at the level of abstracted networking functions, become the Managed-Entity of the upper layer Decision-Making-Element of the system (node), exposing "views" to the upper level DME(s), which uses its overall knowledge to influence (enforce) its lower level DMEs to take certain desired decisions, formalizing the notion of DEs hierarchy.

In the following sections, without loss in the generality of the proposed methodology, we address and study stability in autonomic networks based on the concepts, design principles and attributes of the GANA framework.

3 Stability Issues in Autonomic Networking

There are two key complementary notions of stability when designing autonomic network architectures namely, autonomic node's stability and autonomic network's stability. Node stability concerns the interactions of control loops (self-* functionalities) that exist within the same node, either in the same or different layers of its protocols' stack or concerning various networking functionalities. Network stability copes with the interactions among control loops located in different components of the network; therefore, a two steps approach is required, i.e.

- Initially, the proficient collaboration of control loops towards reaching a stable outcome must be reassured without considering the drawbacks imposed by the networking environment (i.e. assuming a deterministic environment).
- Then, the same study should be repeated when the proposed autonomic approach (e.g. mechanism, architecture, etc) functions/operates over the actual network. At that point, additional challenges emerge that may affect or prevent the efficient collaboration of control loops, like the presence of churn, the loss of communication between nodes (and their corresponding control loops) or even, the existence of miss-behavioural or malicious autonomic nodes.

In either of the previous cases, a methodological approach is required when addressing and studying stability in an autonomic networking paradigm, placing special emphasis on the communication and collaboration among the introduced control loops, regarding their managing, managed, peering and sibling relationships. Specifically, the following key issues must be considered:

Issue I. Self-* functionalities (i.e. control loops) interactions and their convergence at a stable outcome (i.e. equilibrium point). Thus, it is vital to reassure that only important changes cause the triggering of actions or the recalculation of parameters (in order to reduce the causes of instability). Towards this direction, we introduce and

exploit well established guidelines that stem from concepts in the fields of Game Theory and Optimization Theory towards:

- Establishing well-defined valid operating regions of particular control loops.
- Decoupling control systems by ensuring that they control different independent outputs based on independent inputs and if this is not possible, then tuning them so that they impose control at very different timescales. Such an approach will allow us to decouple systems that would otherwise be strongly coupled.

Issue II. Conflicts resolution, not only between self-* functionalities and behaviours (i.e. control loops) belonging to different network components but also among control loops within the same node. Conflicts may occur when more than one control loops manage or affect the same functionalities or resources, especially when they are placed in a hierarchical manner.

Issue III. Time scaling issues among collaborating control loops. In this case, the main difficulty is that timing varies significantly when considering various self-* functionalities which are interacting, while residing on different layers of a node's/network's protocols stack. Thus, it must be reassured that the following two dynamic and interacting sets of "events", that is {signalling, monitoring information and decision making} and {changing environment and triggering events}, are changing on a similar timescale.

In the following we analyze each of the above critical issues and we provide concrete methodologies and key theoretic tools for addressing and treating them.

4 Designing for Stability

In this section we place emphasis on revealing vital aspects related to stability (i.e. control loops stability) that need to be addressed while still being at the design time of an autonomic network and its corresponding autonomic elements, such as the GANA DEs. Hence, we propose solutions and corresponding theoretic tools to address the latter, in contrast to aspects of stability that must be handled at run-time (i.e. as the autonomic system evolves) which are highlighted in the next section.

4.1 Stable Autonomic Behaviours Design through Game Theory – *From Theory to Theory*

Despite the variety of alternative autonomic architectures that may emerge when obtaining the solution of the corresponding initial system's optimization problem, one common designing attribute characterizes them, the decentralized nature of the consequent autonomic algorithms/mechanisms. This not only necessitates the collaboration of various network components to achieve different layering objectives, but also implies the distribution of the decision making procedures of the network among its components, instead of traditional centralized approaches, which eventually increase the role of individual network components.

Aiming at composing efficient distributed and iterative autonomic algorithms, appropriate designing theoretic tools need to be adopted, which are promoting

network's autonomic nature. To that end, Game Theory and Network Utility Maximization (NUM) theory via its decomposition methods [8] (i.e. a theoretical tool for designing optimal distributed algorithms over layered architectures) consist of some of the most widely adopted mathematical frameworks applied in autonomic networking. On one hand, these analytical tools can be used to derive distributed algorithms and determine their efficient collaboration (i.e. autonomic node's DEs (i.e. control loops) signalling), providing theoretically founded answers to the question: *who does what in the autonomic network and how to connect them.* On the other hand, they inherently provide methodologies not only for investigating and reassuring that the corresponding derived distributed algorithms are reaching a stable outcome, in terms of equilibrium points, but also for enabling the following desirable designing attributes: a) limited amount of overhead among nodes (and control lops), b) fully asynchronous updates and c) robustness to arbitrary signaling information [9].

In general, the following thread of analysis holds. An autonomic network's/ functionality's stability is assured via (a) DEs cooperation stability, which can be further reassured via (b) their corresponding control loops collaboration stability and thus, via (c) proving the stability of the distributed iterative algorithms that steer autonomic node's control loops, obtained via the solution of the corresponding optimization problems.

In the following section, we place emphasis and analytically present a concrete methodology for studying autonomic mechanisms' stability via game theory.

How to Treat Stability via Analytical Methods? - A Game Theoretic Approach

Game theory is currently widely explored in autonomic networking due to the following two main inherent attributes:

i) Autonomic nodes' selfish non-cooperative nature can be properly defined and formulated as a non-cooperative game, where nodes act as individual players aiming at maximizing their own interests. Then, game theory provides the foundations and mathematical tools for studying and solving such problems.

ii) The distributed nature of the produced algorithms, which seek non-cooperative game's solution and steer corresponding autonomic nodes DEs (i.e. are their control loops), facilitate the efficient design of a network's autonomic mechanisms.

A non-cooperative game consists of three basic components, (i) a set of players S, consisting of N autonomic nodes, (ii) a set of actions A_i, i.e. the feasible strategy space of i autonomic node and (iii) a set of preferences, which can be expressed via appropriately defined utilities $U_i(a)$ of each autonomic node, where $a_i \in A_i$. Thus, the corresponding non-cooperative game is formulated as follows:

$$\max_{a_i \in A_i} U_i(a_i) \quad s.t. \ a_i \in A_i \ . \tag{1}$$

Furthermore, regardless of the designed problem's formulation, a non-cooperative game's investigation requires the consideration and analysis of the following key steps:

A. Game's steady state via the existence of equilibriums (e.g. ***Existence of Nash equilibrium***).
B. Equilibrium's properties (e.g. ***Nash Equilibrium's properties***).
C. Optimality of equilibrium (e.g. ***Pareto optimality***)

D. Convergence of equilibrium via studying the convergence of users' corresponding (best) response towards selfishly maximizing their utilities, thus reaching game's equilibrium. (e.g. ***Convergence of Nash equilibrium***).

How to address stability via Game Theory? Steps A and D can provide the answer to the previous question. Specifically, upon setting a non-cooperative game, with respect to the required behaviour of the autonomic nodes in the network, and upon deriving distributed algorithms that steer nodes to have the expected behaviour (i.e. autonomic node's control loops), the ability of the corresponding autonomic network to reach a stable outcome (i.e. an equilibrium point) can be analytically proved:

1. By showing the existence (or even the uniqueness) of such a stable point (**step A**);
2. Via proving that the derived distributed (and often iterative) algorithms (i.e. control loops) will always reach such a point **(step D**).

Intuitively, that latter suggests that autonomic nodes' interactions, via their corresponding DEs (control loops), will always reach a stable outcome; thus, leading to autonomic network stability, under the assumption that DEs communication synchronization is always achieved. Game theoretic tools that allow treating issues A and D are super-modular games, utility functions' properties (e.g. quasi-concavity), potential games, coalition games e.t.c. Moving one step forward, the stochastic nature of the networking environment should be considered (i.e. dropping the assumption of reassured DEs communication and synchronization). This makes the analysis of the corresponding games much more difficult. To that end, stochastic games formulations can be applied. A deeper analysis on game theory in autonomic networks is out of this paper's scope; interested readers are referred to [10] and [11].

4.2 Addressing Stability in an Architectural Level – From Theory to Practice

Apart from treating stability via analytical theoretic means, in the following we identify a plethora of vital autonomic network architectures' designing aspects that play a crucial role in determining their stability attributes. Throughout our analysis, the benefits of adopting GANA [2] for devising stable autonomic networks are also revealed, since GANA inherently adopts (i.e. inherent architecture's attributes) the key prerequisites illustrated below.

4.2.1 Hierarchy of Control-Loops (DEs)

An important feature of an autonomic architecture is to maintain a hierarchy of control-loops (as defined in GANA). The benefits from introducing hierarchy to manage complex autonomic systems are extensively justified in [3]. From an autonomic network's stability point of view, the introduction of hierarchy allows the horizontal (among network nodes) and/or vertical (different functional levels) partition of the decision making process. Thus, the failure of a single entity or DE will not result in the total failure of the network under control. DEs with independent goals and policies will continue to operate and manage their corresponding protocols, allowing the failed DE to restart or "correct" itself in the mean time. This feature of GANA allows the network (or at least part of functionalities) to be stable in spite of the failure of a single DE or few DEs.

In addition, some of the major drawbacks of traditional hierarchical systems are large convergence time, sensitivity to failures, large computational power requirement and communication overhead. GANA on the other hand does not follow the traditional hierarchical systems architecture; i.e. it's distributed hierarchical system architecture. Thus, it allows communication between sibling and peer DEs, providing a distributed solution to control its MEs.

To that end, GANA incorporates the following key design principles: **(1)** Lower level DEs expose "views" up the Decision Plane, allowing the upper (slower) control loops to control the lower level (faster) control-loops driven by lower level DEs); **(2)** Changes computed in the upper DEs implementing slower (i.e. within larger time frames) Control-Loops are propagated down the DE hierarchy to the Functions-Level DE(s) implementing the faster control-loops that then arbitrate and enforce the changes to the lowest level Managed Entities (protocols and mechanisms). This operation reduces the drawbacks of large convergence and sensitivity times, traditionally found in conventional hierarchical systems. Finally, the nature of the distribution of the tasks to DEs inside the node and Network-Level DEs further ensures that the communication and computation power requirements are kept to a bare minimum inside the node.

4.2.2 Concept of "Ownership"

The "Concept of Ownership" is another feature of the intrinsic stability attributes that must be considered for an autonomic network architecture (as is also defined in GANA). This concept requires that every ME is managed by a single DE, i.e. no two DEs (i.e. control loops) can control the same ME (i.e. functionality, resource, e.t.c.) at any given point of time in the network. This is important from system's stability point of view since it relieves the burden of "*conflicts resolution*". Specifically, if an ME is controlled by two or more DEs at the same time, contrasting, conflicting and at times repetitive policies, objectives and reconfiguration requests, etc, originating from different DEs lead to an unstable ME and thus, to an unstable autonomic network. Through the "Concept of Ownership", GANA ensures that this instability is avoided.

4.2.3 Separation of "Operating Regions"

Another prerequisite of an autonomic architecture is the efficient separation of the "operating regions" for the control-loops as advocated in [12]. This can be achieved by decoupling control systems and ensuring that they control different independent outputs based on independent inputs, as defined in GANA. If it is impossible to decouple certain outputs and inputs from affecting each other, it is important to reassure that they impose control at different timescales on their MEs.

4.3 Model-Based Techniques

Model-driven approaches for design DEs and their inter-relationships should also be exploited to efficiently address aspects related to stability of control-loops. Specifically modelling and validation of DEs autonomic behaviours using Formal Description Techniques (FDTs), such as the well-known and successful ITU-T SDL language, can be explored to address certain aspects of stability. Such an approach would enable the design, model-checking and verification of DEs, as well as the validation, simulation and some partial code-generation of autonomic behaviours of

DEs. Finally, the aspects (i.e. methodology) [13] that need to be taken into account when following such an alternative are: (1) A Meta-Model that enforces constraints in the design models for DEs and their interactions with assigned MEs and with other DEs, that should be embedded in the Modelling Tools e.g. a "model editor", is required; (2) A Model-Walker that can be designed for walking over a design model in order to detect conflicting control-loops and overlaps in the so-called "valid operating regions" of control-loops specific to DEs; (3) Simulations for detecting behaviour conflicts between interacting control-loops designed to operate within a node/device and in the network.

5 Addressing Stability at Runtime

After applying the methodologies and guidelines presented in the previous sections, an autonomic network is intrinsically equipped with interacting control loops having some degree of self-stabilization. However, it is possible that due to the stochastic nature of the networking environment, situations may occur, which have not explicitly been considered during the design time of the Decision Elements. We denote such situations as *"emergent situations/behaviours"*. Intuitively, "emergent behaviours" can be considered as an indirect interference between a set of DEs, whereby each DE is optimizing its own goal based on its embedded logic, but the overall set of executed actions is leading the system to an unstable state of oscillations and continuous responses of diverse control loops. In order to avoid such emergent behaviours, the concept of Action Synchronization Functions (ASFs) has been proposed in [14] (which is also part of a GANA's Decision Elements). In simple words, [14] introduces ASFs in the GANA hierarchical structure, in order to allow DEs on a lower level to resolve potential conflicts via requesting DEs belonging to a higher level (in the DE's hierarchy) for taking over their synchronization and coordination. Specifically, after a number of DEs have referred to a higher level DE in order to be informed whether particular tentative action(s) are allowed or not, the higher level DE selects the optimal subset of tentative actions reported by the corresponding lower level DEs. The tentative actions are gathered over a pre-defined time window. Consequently, the higher level DE responds back to the requesting lower level DEs on whether they are allowed to proceed with executing tentative actions or not.

There are various aspects of stability which are addressed by the architectural role of the ASF namely, 1) acting as an arbiter for conflict resolution among DEs with potentially interfering actions, 2) exploiting the GANA hierarchical structure in order to realize the notion of hierarchical optimal control – autonomic entities on a higher level have access to more global information, and 3) enabling the gathering of actions for synchronization over a pre-defined time period has a smoothing effect on the rate at which control loops operate, thereby avoiding sudden oscillations and chaotic situations in the autonomic network.

In [14], the issue of conflict resolution is tackled via deriving a binary integer program, which aims at the optimization of a set of key performance indicators (KPI) by selecting an optimal subset from the $m \in N^+$ actions waiting to be synchronized:

$$\max_{p \in \{0,1\}^m} w^T I_m p \quad \text{s.t.} \quad D_m p \leq c \tag{2}$$

where, *w,* is a vector containing a weight value for each considered KPI. These weights specify the importance of each KPI for the network and are specified by the network operator; *P,* is a binary (0-1) vector being optimized. This vector reflects the actions that have to be synchronized. A "1" at a particular position means that the corresponding action has to be executed. A "0" stands correspondingly for the case when an action must be stopped. I_m is an impact matrix (similar to a payoff matrix in the context of game theory) that specifies the impact of the tentative actions on the KPIs considered by the ASF.

The constraints of the previous optimization problem are specifying the number of actions that are allowed to simultaneously influence (as specified in I_m) a particular KPI. The matrix multiplication on the left side of the inequality gives the number of actions which influence the KPIs. In order to achieve that, D_m is defined as a binary (0-1) matrix, in which a "1" is set in case $(I_m)_{i,j}>0$ and a "0" in case $(I_m)_{i,j}=0$.The components of the vector $c \in N^m$ stand for the number of actions which can simultaneously influence the corresponding KPI.

We can interpret the above optimization problem as *"selecting the most appropriate subset of tentative actions such that the change in the state of the system is positively maximized"*. For a bottom-up definition of this optimization problem, starting with the current state (as defined by the current KPI values) of the system, we refer the reader to [14]. Unfortunately, binary program (2) belongs to a class of NP-hard problems. For that purpose, in this work, we propose to relax the binary constraint imposed on the vector *p,* and turn it into an inequality: $0 \leq p_i \leq 1,\ i \in \{1, \ldots, m\}$. Thus, the new optimization problem (contrasted to the one in [14]) is given by:

$$\max_{p} \ w^T I_m p \quad \text{s.t.} \ D_m p \leq c,\ 0 \leq p_i \leq 1,\ i \in \{1, \ldots, m\}. \tag{3}$$

This optimization problem is a linear program which constitutes a convex optimization problem. Hence, there is only one optimum and every local optimal solution is also a global one, which means that the diverse optimization techniques will always improve iteratively the quality of the obtained solution. The above formulation is non-NP-hard. The result of this optimization is a vector containing values from the interval *[0, 1]*. If *ith* component of this vector is 0 then the corresponding action is disallowed. Correspondingly, if it is 1 then the action should be issued. If a DE, requesting for synchronization, receives as response a value from the interval *(0, 1)*, then it can either execute or drop the action, based on an internal threshold θ_i. These thresholds can be supplied by the human experts tweaking the autonomic network. For instance, the history of requests for synchronization can be analyzed offline (e.g. by employing statistical and/or machine learning methods) and appropriate thresholds can be extracted using some statistical or optimization tools.

5.1 Autonomic-Aware Metrics to Infer and Self-assess Stability

Since an autonomic network is expected to be self-adaptive to changes and challenges in its environment during its operation, it must be able to self-assess and infer stability related problems experienced at various levels of the Decision Plane Hierarchy. The assessment of stability at different levels where a DE implements an autonomic functionality (e.g. autonomic routing, autonomic QoS-Management, autonomic Mobility-Management, etc), requires that every DE must implement some monitoring

functionalities towards self-awareness (i.e. implement *Counters* for storing statistics e.g. i) the number of times a DE enforced a change the behaviour of its assigned Managed Entities (MEs) in reaction to specific events over a period of time, or ii) the number of times a DE received information indicative of instability problem from its MEs). What is also required, are *Timing Variables* (e.g. timers) for storing time durations that measure some DE activities. Diverse types of Timing Variables are required for each type of input that flows into a DE—according to the "valid operating region" of its associated control-loop that was captured and defined at design time.

Thus, every DE must expose the "views" captured by the *Counters* and the *Timing Variables* to its upper DE, which then assesses the "*degree of stability*" of the autonomic functionality realized by the particular lower level DE and decides to enable an appropriate response strategy towards reassuring stability. The upper DE may further aggregate some statistics and "views" and expose then further up the Decision Plane (possibly up to the level of network-level DEs) where more sophisticated decisions can be taken based on the wider knowledge about the network that is gathered by network-level DEs. Finally, an autonomic behaviour triggered by a DE reacting to a change may inductively span a number of DEs, and their associated control-loops, which are enabling (via collaborations and interactions) *an ultimate goal* (i.e. an autonomic behaviour). Such an *ultimate goal* could be the (re)-setting of a parameter value on a single ME or multiple parameters on an ME(s) managed by the first triggering DE in the DE interaction chain, or could be the setting of multiple parameters on ME(s) managed by one or more other DEs involved in the DE interaction chain. A DE e.g. a Node-DE, belonging to a higher level than the DEs involved (which could be Function-Level DEs), provided it knows the *causality graph for actions/policies employed by different DEs* to achieve the *ultimate goal*, could then use information stored in *Timer* variables stored by lower-level DEs in the interaction chain, in order to infer stabilization time and challenges, via using this knowledge to improve cognitive response strategies over time.

6 Concluding Remarks

In this paper we address the issue of autonomic networks stability. Despite the existence of recent elegant analytic results in specific networking paradigms, the issue of stability in autonomic networking still lacks of concreteness, generality and mainly extensive validation; hardening the wide applicability of autonomic solutions in realistic environments. Following a pragmatic and concrete thread of analysis we identify, categorize and discuss all the key aspect that affect or characterize autonomic architecture stability, from theoretic analysis and network design point of view, to practical implementations and runtime solutions. Moving one step further, we propose concrete methodologies and highlight corresponding theoretic tool for addressing and studying stability problems in autonomic networks [15].

Acknowledgement

This work has been partially supported by EC FP7 EFIPSANS project (INFSO-ICT-215549).

References

1. ITU: The FCAPS management framework, ITU-T Recommendation M. 3400, Telecommunications management network, Copyright ITU (2001)
2. Chaparadza, R., Papavassiliou, S., Kastrinogiannis, T., Vigoureux, M., Dotaro, E., Davy, A., Quinn, K., Wodczak, M., Toth, A.: Creating a viable Evolution Path towards Self-Managing Future Internet via a Standardizable Reference Model for Autonomic Network Engineering. In: Future Internet Assembly (FIA) book in Europe, pp. 136–147. IOS Press, Amsterdam (2009)
3. Diao, Y., Hellerstein, J.L., Parekh, S., Griffith, R., Kaiser, G.E., Phung, D.: A control theory foundation for self-managing computing systems. IEEE Journal on Selected Areas in Communications 23(12), 2213–2222 (2005)
4. Jiang, T., Baras, J.S.: Fundamental Tradeoffs and Constrained Coalitional Games in Autonomic Wireless Networks. In: Proc. of 5th Int. Symp. on Modeling and Opt. in Mobile, Ad Hoc and Wireless Networks, WiOpt 2007, pp. 1–8 (2007)
5. Boutaba, R., Xiao, J., Zhang, Q.: Toward Autonomic Networks: Knowledge Management and Self-Stabilization, Autonomic Computing and Networking, pp. 239–260. Springer, Heidelberg (2009)
6. Greenberg, A., Hjalmtysson, G., Maltz, D.A., Myers, A., Rexford, J., Xie, G., Yan, H., Zhan, J., Zhang, H.: A clean slate 4D approach to network control and management. ACM SIGCOMM Comp. Com. Review 35(5), 41–54 (2005)
7. Ballani, H., Francis, P.: CONMan: A Step Towards Network Manageability. ACM SIGCOMM Computer Com.Review 37(4), 205–216 (2007)
8. Fu, F., Van der Schaar, M.: A New Systematic Framework for Autonomous Cross-Layer Optimization. IEEE Trans. on Veh. Tech. 58(4), 1887–1903 (2009)
9. Rad, H.M., Huang, J., Chiang, M., Wong, V.: Utility-optimal random access: Reduced complexity, fast convergence, and robust performance. IEEE Tran. on Wireless Com. 8(2), 898–911 (2009)
10. Felegyhazi, M., Hubaux, J.P.: Game Theory in Wireless Networks: A Tutorial. In: Proc. of Infocom 2006, Barcelona, Spain (April 2006)
11. Saad, W., Han, Z., Debbah, M., Hjorungnes, A., Basar, T.: Coalitional Game Theory for Communication Networks: A Tutorial. IEEE Signal Processing Magazine, Special Issue on Game Theory 26(5), 77–97 (2009)
12. Mortier, R., Kiciman, E.: Autonomic Network Management: Some Pragmatic Considerations. In: Proc. of 2006 ACM SIGCOMM, NY, USA, pp. 89–93 (2006)
13. Prakash, A., Chaparadza, R., Theiz, Z.: Requirements of a Model-Driven Methodology and Tool-Chain for the Design and Verification of Hierarchical Controllers of an Autonomic Network. In: First International Conference on Models and Ontology-based Design of Protocols, MOPAS 2010 (June 2010) (to appear)
14. Tcholtchev, N., Chaparadza, R., Prakash, A.: Addressing Stability of Control-Loops in the context of the GANA architecture: Synchronization of Actions and Policies. In: Intl. Workshop on Self-Organizing Sys., IWSOS, Zurich (2009)
15. EC funded- FP7-EFIPSANS Project, http://efipsans.org/

An Empirical Evaluation of a Shim6 Implementation

John Ronan and John McLaughlin

Telecommunications Software & Systems Group,
Waterford Institute of Technology,
Cork Road, Waterford, Ireland
{jronan,jmclaughlin}@tssg.org

Abstract. Several solutions are proposed to enable scalable multihoming over IPv6. One of these proposals is Shim6, a host-based multihoming solution based on the modification of the Internet Protocol stack of the host. This modification adds a layer below the transport protocols but above the forwarding layer. As this approach makes the modifications to the network stack transparent, existing applications automatically benefit from Shim6 functionality.

In this paper we investigated aspects of the performance of the LinShim6 implementation from Université Catholique de Louvain. We also outline our modifications of the LinShim6 implementation to allow external software to control the locators used between hosts.

Keywords: Shim6, multihoming, ECN.

1 Introduction

For a number of years now, the IETF has been working on IPv6, a successor to IPv4. More recently, work has been done to address the scalability of the current internet architecture. The Internet Architecture Board (IAB) has identified several limitations of the current internet architecture[1]. These issues impact the scalability of inter-domain routing systems, which is reflected in the growth of Border Gateway Protocol (BGP)[2] routing tables, and also in the number of routes in the Default Free Zone (DFZ) Routing Information Base (RIB) processed by BGP routers. Several factors which influence the growth of BGP routing tables include, multihoming, traffic engineering, IP address allocation policy, and business events, such as large mergers and acquisitions. All of these factors can lead to an increased number of unique routing prefixes that cannot be aggregated within the DFZ RIB and hence cause routing table growth.

Several years ago, after examining the multihoming issue[3], the IETF chartered the Site Multihoming by IPv6 Intermediation (Shim6) working group to develop a host-based IPv6 multihoming solution, [4] presents a good overview of the requirements, constraints, and the process that led to the emergence of

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 62–70, 2011.

Shim6 as a multihoming solution. The Shim6 specification documents are now published[5,6,7], and in this paper we report on our experiences with Université Catholique de Louvain's (UCL) publicly available Shim6 implementation LinShim6[8,9] for the Linux kernel.

This paper is organised as follows: Section 2 provides a brief description of the capabilities of Shim6 for the benefit of those readers unfamiliar with the protocol. Section 3 describes both our overall goal and experiences in creating a Shim6 testbed. Section 4 presents the baseline performance measurements, with a brief description of the results. In closing, section 5 discusses the work presented here within the broader context of the EU FP7 EFIPSANS[1] research project.

2 Shim6 Host-Based IPv6 Multihoming

The Shim6 protocol[6], has been designed to add multihoming capabilities to IPv6 end-hosts. Potentially, this allows for far more IPv6 enabled sites to protect their upstream connections, without having to go to the trouble of implementing BGP peering. This means that entities can retain control within their own site without incurring the overhead of deploying BGP.

Along with the Shim6 protocol, the IETF Site Multihoming by IPv6 Intermediation (Shim6) working group designed a failure detection and repair mechanism, called the REAchability protocol (REAP)[7] which allows hosts to detect and recover from failures.

Today, in the current (IPv4) Internet, a multihomed site is obliged to have a network connection with each of its upstream providers, and the site has to use IP addresses independent from those providers. These addresses come from what is called Provider Independent or PI address space (as opposed to Provider Aggregatable (PA) address space).

With Shim6 however, multihoming functionality is made available to the end host using Provider Aggregatable addresses — removing the need to involve BGP or any other protocol. At present, the default IPv6 address selection algorithm [10] defines how the address pair for a communication session is selected, this address pair does not change for the duration of the session. Shim6 offers the ability to change the address pair used (and thus the path) during the session, transparent to the application. The Shim6 approach uses routable IP addresses (locators) as the identifiers visible to the transport layer. This also provides the facility to change the locator pair in use should REAP detect that the currently used pair of addresses (or interfaces) between two communication nodes has failed. REAP will search for a working pair of locators and pick another working pair (if available) when this occurs[7]. This change is performed at the network layer, which means that applications and transport protocols do not need any changes to benefit from this new capability.

[1] **E**xposing the **F**eatures in **IP** version **S**ix protocols that can be exploited/extended for the purposes of designing/building **A**utonomic **N**etworks and **S**ervices.

3 Testbed

Our primary goal in this work was to get a baseline performance metric for an existing Shim6 implementation, and then to integrate Shim6 into the overall EFIPSANS architecture[11]. While Shim6 is already somewhat autonomous, in that it can detect and recover from link failures, we augmented LinShim6 with functionality to allow third party code to directly inform the Shim6 implementation which locators it should use. This facility could be used in the case of a scheduled downtime, for example. As a proof-of-concept for the EFIPSANS project, to demonstrate monitoring functionality, we developed a small daemon coupled with a Linux Netfilter[12] module to detect congestion or loss in a network through the Explicit Congestion Notification (ECN)[13] mechanism. This information could then be acted upon by third-party code to instruct LinShim6 to change the locator set in use, based on congestion detected and other variables such as, network load, jitter, delay etc.

3.1 Shim6 Testbed

This testbed (figure 1) was set up to replicate the scenario in "Performance Analysis of REAchability Protocol for IPv6 Multihoming"[14]. We installed and configured the UCL Shim6 implementation, LinShim6, and proceeded to generate a set of results in order to ascertain what differences (if any) were present in the behaviour of this implementation versus the simulated results available in[14]. The testbed consisted of:

- Two Dual PII blade servers, each with 3 Network Interface Cards
- One Juniper M10i running JunOS

From our initial work on constructing this testbed, a number of issues arose. Initially we uncovered various bugs in the LinShim6 implementation, that became apparent due to our deployment being on real hardware. This led to much work being done both by ourselves and Sébastien Barré, the LinShim6 author, to diagnose and fix these issues.

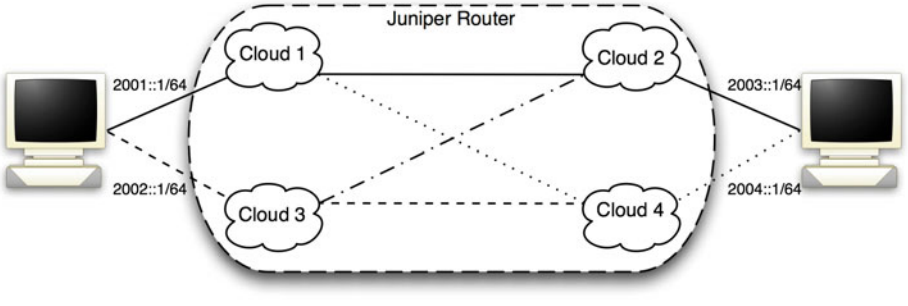


Fig. 1. Shim6 testbed in TSSG

When deploying the testbed, it quickly became clear that the authors in [14] did not use any routing protocols in the simulation. As our testbed was deployed on real equipment we felt that is was important to be as realistic as possible so we used the Open Shortest Path First (OSPF) [15] protocol on the inter virtual-router links[2]. Obviously this meant that OSPF was able to recover from any single link failure by itself. Consequently, in order that a path failure could be simulated, the link between *Cloud 2* and *Cloud 3*, was manually disabled, thus restricting the available redundant paths between the two hosts.

When generating traffic, we originated all sessions from 2003::1 to 2001::1 through *Cloud 1* and *Cloud 2*. Then, to simulate failure, at a certain point in time, the link between *Cloud 1* and *Cloud 2* is disabled, this failure is detected by REAP, and, after path exploration, the session should continue between the locator pair of 2004::1 and 2002::1.

The tests performed involved the TCP [16,17,13] and UDP [18] protocols. For TCP tests, we evaluated the TCP behaviour using the FTP [19] protocol. The traffic used to evaluate the UDP behaviour corresponded, as close as possible to a Voice over IP (VoIP) application using a G.729 [20] codec both with a uni-directional and bi-directional packet flow. To emulate G.729 the Iperf [21] tool was configured to generate 8 kilobits of data per second (50 packets per second, 20 bytes per packet).

The Round Trip Time (RTT) in both paths was configured to be identical. The Netem tool [22] was used to implement a normal distribution with a mean of 80ms and a 20ms variance. The "failure" event occured at a random interval between 75 and 125 seconds after the test run commences. All test runs were terminated 60 seconds after the "failure" event. These choices were dictated by those used in [14].

To run the tests, scripts were written to automate every run. For each value of T_{send} from 1 to 15, 45 test runs were completed. This gave a total of 675 test runs. This was done for each of TCP (the FTP protocol), bidirectional UDP, and unidirectional UDP. Giving over 2000 total runs or over 4000 unique log files.

3.2 Explicit Congestion Notification

Congestion is a perpetual problem in networks and can have a detrimental effect on user experience in situations where a high QoS is required (video streaming, VOIP etc). The Explicit Congestion Notification (ECN) protocol provides a means to detect congestion in IPv4 and IPv6 networks. Although it has been standardised for over a decade, it has suffered from slow uptake. This appears to be as a result of packet loss from intermediate routers rigidly enforcing earlier RFCs, and hence dropping packets as "invalid".

Briefly, when two endpoints have negotiated use of ECN, the sender of data packets will mark the outgoing packets with an ECN code point (2 bits in the IPv6 Traffic Class octet). An intermediate router approaching the point of congestion which comes across one of these packets will update it to signal that it is about to become congested, by setting the Congestion Encountered (CE) code

[2] OSPF is used internally in our site, and was a logical choice.

point. Upon receipt of such a packet, the receiver will notify the sender via the ECN Echo (ECE) bit in the TCP header of the next TCP ACK, that the data packet experienced congestion. The sender then will take steps to "back-off" in an attempt to alleviate the congestion problem. Also, when an ECN-Capable TCP sender reduces its congestion window for any reason, the TCP sender sets the CWR bit in the TCP header of the first new data packet sent after the window reduction. This means there are two indicators available for use to use as congestion (or loss) indicators.

In order to facilitate some control over the chosen Shim6 network pair, we have extended the LinShim6 user space daemon *shim6d* with a "put" command. This command allows one to request Shim6 to use a specific locator pair at any time. However, the selected pair is still subject to the normal Shim6 rules in that if it should fail for whatever reason, Shim6 will automatically start the process to select a valid locator pair. For the EFIPSANS project, we have added functionality to LinShim6 to use the information presented by the ECN implementation when congestion (or loss) is detected and this information could potentially be used to migrate any affected Shim6 sessions to a clear path. The decision to migrate could be based on information such as prior knowledge of clear paths or other knowledge that could be supplied to the host from another service[23,24].

The code consists of a Netfilter module and a user space daemon. The Netfilter module intercepts any ECE or CWR marked packets for the user space daemon to examine. If the daemon determines that the packet is, indeed of interest. The daemon will output the source and destination addresses of the effected stream via a network socket such that a listening application could make decisions based on this data. The output from the network socket is shown in listing 1.

```
labadmin@sam:~$ telnet localhost 2223
Trying ::1...
Connected to localhost.
Escape character is '^]'.
<?xml version="1.0" encoding="UTF-8"?>
<ecn src="2001:770:20:4::2" dest="2001:770:20:e::2" congested="true"/>
<?xml version="1.0" encoding="UTF-8"?>
<ecn src="2001:770:20:4::3" dest="2001:770:20:e::2" congested="true"/>
```

Listing 1.1. Output from network socket

4 Results

4.1 LinShim6

As mentioned already, our Shim6 testbed was set up to replicate the scenario depicted in[14]. In that paper the metric the authors used was "Application Recovery Time". This is defined as the difference in time between the last packet arriving through the old locator set (addresses), and the first packet arriving through the new one, after the the path between the locator set has failed. This metric accurately measures the time taken to recover from a path failure when there is a continuous flow of traffic. The same metric was used for our measurements.

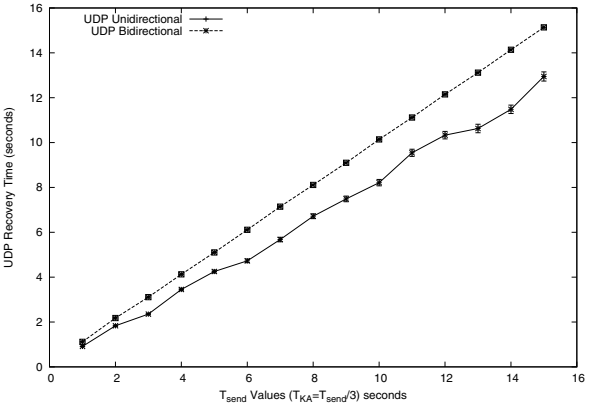


Fig. 2. UDP Recovery Time

UDP behaviour. Figure 2 shows both bidirectional and unidirectional UDP recovery times. Two traffic profiles have been emulated. The Iperf network testing tool was used to generate a bidirectional UDP stream (VoIP conversation) and a unidirectional UDP stream (audio stream), with similar characteristics. Comparison with [14] reveals a marked similarity in results.

TCP behaviour. TCP has several characteristics that UDP does not, such as reliability and congestion control. The authors used the FTP protocol to reliably generate high-bandwidth traffic. Also, as the LinShim6 implementation is capable of resetting TCPs retransmission timeout (RTO), we also performed this test. Figure 3 shows TCP recovery times for TCP both with and without the retransmission timers reset. Figure 4 compares bidirectional UDP and TCP. As can be seen from figure 3 and figure 4. The results obtained validates the proposals in §4.2 of [14]. In this work, the authors proposed that after a new path is chosen for a communication, that the TCP retransmission timer value should be reset. They argue that this is both more efficient and more appropriate as the timer values are dependent on the path in use now, not previously. Their simulation results showed that the relation between the TCP recovery time and T_{send} was was linear, and the the modified TCP behaviour was also very similar to that of UDP.

4.2 Comments on ECN

We were unable to successfully enable IPv6 ECN in the router equipment in our testbed. It appears that none of our routers (several Cisco, Juniper, Linux, FreeBSD) supported full IPv6 ECN functionality. However, we were able to test that our code functioned correctly. We were able to generate packet loss, and thus detect when the TCP sender set the CWR bit of the TCP header, as shown listing 2.

Consequently, we are confident that should we get to work with an IPv6 capable, ECN enabled router that we will be able to detect congestion and feed that information up to a management entity.

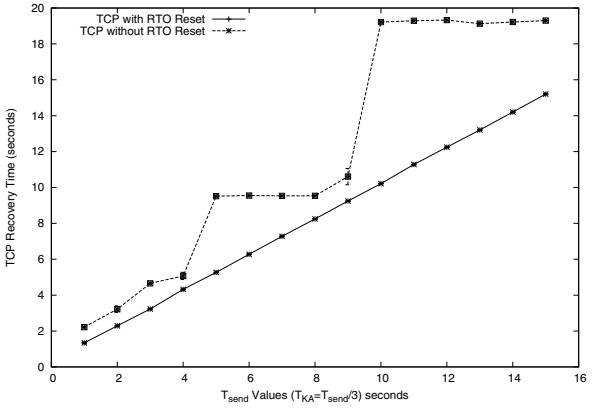


Fig. 3. TCP Recovery Time

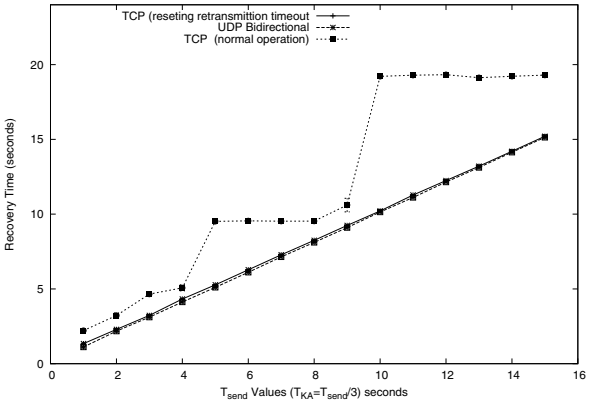


Fig. 4. Here we can see that for TCP with RTO reset enabled, the behaviour is almost identical to that of bi-directional UDP. While unmodified TCP behaviour is clearly visible, showing the staircase like graph.

```
root@sam:~/src/tmp/ecnd# ./ecnd
[Signal] Creating signal handlers
[[ECN Queue]] Binding nfnetlink_queue as nf_queue handler for AF_INET6
[[ECN Queue]] Binding this socket to queue '0'
[Client] (client socket handle=5)
[ECN Monitor:parse()]
(seq=916108003, ack_seq=-1562873524) CWR = 1, ECE=0, SYN=0, ACK=1
[Process] Added path :
(src='2001:770:20:84::2',
dest='2001:770:20:84:250:c2ff:fe07:92db:) ECN=Yes
[ECN Monitor:parse()]
(seq=575124451, ack_seq=-1562873524) CWR = 1, ECE=0, SYN=0, ACK=1
[ECN Monitor:parse()]
(seq=-234932923, ack_seq=-1011743955) CWR = 1, ECE=0, SYN=0, ACK=1
```

Listing 1.2. Output from ecnd daemon

5 Conclusion

This paper presents details of work done in evaluating Université Catholique de Louvain's (UCL) publicly available Shim6 implementation. Its performance in our test network was compared against prior work, where the the behaviour of the Shim6 was simulated. The actual results obtained compare favorably with the simulation results. We then proceeded to implement a feedback mechanism based on the the Explicit Congestion Notification (ECN) protocol. This could allow for the re-balancing of traffic between hosts on clear (not experiencing congestion) paths should the hosts desire this functionality. Or indeed, in our case just act as a mechanism for reporting network congestion or loss to an EFIPSANS Managed Entity, which, monitors network performance.

We are also interested in testing our work across the Internet itself and gaining more relevant information as to the behaviour of Shim6 in larger deployments.

Acknowledgements

This work was partly funded by the European Commission via the 7th Framework Programme Integrated Project EFIPSANS (grant no. 215549). Many thanks to Sébastien Barré, the LinShim6 author, for his assistance and swift response to innumerable questions.

References

1. Meyer, D., Zhang, L., Fall, K.: Report from the IAB Workshop on Routing and Addressing. RFC 4984 (Informational) (September 2007)
2. Rekhter, Y., Li, T., Hares, S.: A Border Gateway Protocol 4 (BGP-4). RFC 4271 (Draft Standard) (January 2006)
3. Huston, G.: Architectural Approaches to Multi-homing for IPv6. RFC 4177 (Informational) (September 2005)
4. de Launois, C., Bagnulo, M.: The paths towards IPv6 multihoming. IEEE Communications Surveys and Tutorials 8(2) (2006)
5. Bagnulo, M.: Hash-Based Addresses (HBA). RFC 5535 (Proposed Standard) (June 2009)
6. Nordmark, E., Bagnulo, M.: Shim6: Level 3 Multihoming Shim Protocol for IPv6. RFC 5533 (Proposed Standard) (June 2009)
7. Arkko, J., van Beijnum, I.: Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming. RFC 5534 (Proposed Standard) (June 2009)
8. Barré, S.: Linshim6 - implementation of the shim6 protocol. Technical report, Université catholique de Louvain (February 2008)
9. Barré, S., Bonaventure, O.: Shim6 implementation report: Linshim6. Internet draft, draft-barre-shim6-impl-03.txt, work in progress (September 2009)
10. Draves, R.: Default Address Selection for Internet Protocol version 6 (IPv6). RFC 3484 (Proposed Standard) (February 2003)
11. Tcholtchev, N., Grajzer, M., Vidalenc, B.: Towards a unified architecture for resilience, survivability and autonomic fault-management for self-managing networks. In: 2nd Workshop on Monitoring, Adaptation and Beyond (MONA+), Stockholm, Sweden, November 23-24 (2009)

12. The Netfilter Project. Netfilter - firewalling, nat and packet mangling for linux, http://www.netfilter.org
13. Ramakrishnan, K., Floyd, S., Black, D.: The Addition of Explicit Congestion Notification (ECN) to IP. RFC 3168 (Proposed Standard) (September 2001)
14. de la Oliva, A., Bagnulo, M., García-Martínez, A., Soto, I.: Performance Analysis of the REAchability Protocol for IPv6 Multihoming. In: Koucheryavy, Y., Harju, J., Sayenko, A. (eds.) NEW2AN 2007. LNCS, vol. 4712, pp. 443–454. Springer, Heidelberg (2007)
15. Moy, J.: OSPF Standardization Report. RFC 2329 (Informational) (April 1998)
16. Postel, J.: Transmission Control Protocol. RFC 793 (Standard), Updated by RFCs 1122, 3168 (September 1981)
17. Braden, R.: Requirements for Internet Hosts - Communication Layers. RFC 1122 (Standard), Updated by RFCs 1349, 4379 (October 1989)
18. Postel, J.: User Datagram Protocol. RFC 768 (Standard) (August 1980)
19. Postel, J., Reynolds, J.: File Transfer Protocol. RFC 959 (Standard), Updated by RFCs 2228, 2640, 2773, 3659 (October 1985)
20. International Telecommunication Union. G.729 (2007), http://www.itu.int/rec/T-REC-G.729/e
21. National Laboratory for Applied Network Research. Iperf, http://iperf.sourceforge.net/
22. Stephen Hemminger. Network emulator, http://www.linuxfoundation.org/collaborate/workgroups/networking/netem
23. Saucez, D., Donnet, B., Bonaventure, O.: Idips: Isp-driven informed path selection. IETF Draft (February 2008)
24. Bonaventure, O., Saucez, D., Donnet, B.: The case for an informed path selection service. IETF Draft (February 2008)

Future Autonomic Cooperative Networks

Michał Wódczak

Telcordia Poland Sp. z o.o.
ul. Umultowska 85
61–614 Poznań, Poland
mwodczak@telcordia.com

Abstract. Both cooperative transmission and autonomic networking have emerged recently as very promising technologies ready to become the key components of the concept referred to as the Future Internet. Cooperative transmission has been one of the hottest research topics lately capitalizing on the exploitation of relay nodes, while autonomic networking is promoting a very desirable vision that networked systems should be able to act as a living organisms and self-configure without any external intervention. This paper promotes the idea of joint approach to both technologies so the end users are benefited in terms of the quality of services they are provisioned.

Keywords: Cooperative transmission, autonomic networking, service provision, Future Internet.

1 Introduction

Both cooperative transmission [7] and autonomic networking have emerged recently as very promising technologies ready to become the key components of the concept referred to as the Future Internet. On the one hand, there is the cooperative transmission being one of the hottest research topics lately which capitalizes on the exploitation of radio channel diversity provided by additional relay nodes. On the other hand, there exists the paradigm of autonomic networking based on the idea that a given networked system should be able to act as a living organism so that it is able to self-configure without any external intervention. Now, as the topic of Future Internet has become the focal area of interest investigated thoroughly be the research community, both approaches seem to be coming to a common point. In particular it is of prime importance to guarantee seamless service provision so end users can observe actual benefits resulting from the latest advances in modern communications systems. There is then an urgent need to devise networked systems that would be able to self-manage so they offer the best possible service to end users while strictly following the policies imposed by network operators. Consequently, such a system would have to express the ability to analyse the current situation and then make an attempt to benefit from the gains offered by cooperative transmission by means of the concept of autonomic networking in terms of configuring network devices properly so they are able to expose certain cooperative behaviours. The concept presented in this paper is developed as a part of the Generic Autonomic Network Architecture (GANA) [2].

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 71–78, 2011.

The paper is organised in the following way. First the concept of virtual antenna array aided space-time block coded cooperative transmission is outlined together with the considerations about its applications and specific topics of research. Then the section about autonomic networking follows, where the paradigm of autonomicity is presented along with a discussion about the facilitation of certain network behaviours such as an ability to instantiate cooperation among nodes. The discussion is carried out in the context of service provision for the Future Internet. Conclusion outlines the main points of the analysis and brings the closing remarks.

2 Cooperative Transmission

The idea of cooperative transmission aims at improving the reliability of wireless mobile communications through the use of diversity that can be provided thanks to additional network nodes assisting in the transmission between the source node and the destination node. This is in fact possible because the rationale behind space-time processing can be easily mapped onto networking as long as tight synchronization is guaranteed. In other words, network nodes can be perceived as the elements of a virtual antenna array [5] and they can act as a distributed space-time encoder [9]. The basic approach is described in Figure 1 and the transmission takes two stages.

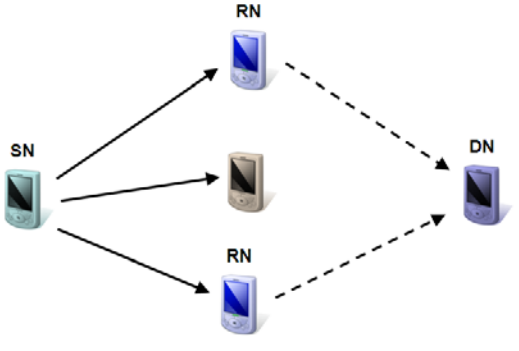


Fig. 1. Generalised virtual antenna array concept

First, the source node (SN) broadcasts the data to be delivered to the destination node (DN) and the relevant information is received by the intermediary node(s). Each of these nodes might become a relay node (RN) and during the second phase only the intermediary nodes selected as relays are entitled to resend the received data towards the destination. This process requires proper synchronisation and application of a spatial-temporal processing technique to orthogonalise the wireless radio channel.

Different spatial-temporal processing techniques can be taken into account and in this paper the emphasis is laid on space-time block coding while the

other techniques are generally equally well applicable to the presented concepts. The simplest space-time block code G_2 is defined below (1) to outline the basic operation of space-time processing aiming at wireless channel orthogonalisation, as well as to better illustrate its relation to the concept of virtual antenna arrays.

$$G_2 = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \tag{1}$$

The operation of a space-time block encoder (or its distributed version in the form of a virtual antenna array) for the presented matrix of the G_2 code can be described as follows: in the first time slot the x_1 and x_2 symbols are sent by the first and second transmit antenna respectively and then, in the second time slot, the $-x_2^*$ and x_1^* symbols are transmitted alike. As it has been already mentioned this operation can be translated to virtual antenna array aided cooperative transmission by simple substitution of antennas with network nodes and by the addition of certain network logic able to provide the notion of cooperative processing.

Cooperative transmission is not only applicable to mobile ad hoc networks, mesh networks or sensor networks but there already exist interesting applications to the next generation cellular systems as well [6] (see also the evaluation results contributed by the author to [8]). Regardless the environment, however, in most of the cases there is a need to answer the question of the selection of relay nodes to be included in a virtual antenna array. The question of transmit antenna selection was similarlyl applicable to plain space-time coding systems where one could have observed benefits from proper antenna selection [12]. For networked systems, however, this issue becomes even more substantial and complex as the network topology may be changing rapidly. One of the proposed approaches assumes the employment of specific existing routing layer mechanisms for the purposes of gaining access to and capitalising on topology information readily available at the network layer. In particular the Optimised Link State Routing protocol (OLSR) [4] is used which belongs to the proactive group of routing protocols tailored to Mobile Ad hoc Networks (MANET). The obvious advantages of the OLSR protocol include its ability of proactive network topology discovery and its inherent optimised broadcasting mechanism in the form of the multi-point relay (MPR) station selection heuristic. As described in [14] and [13] the modification of this mechanism allows for a seamless integration of the concept of virtual antenna aided and space-time block coding based cooperative transmission into the existing protocol. In other words, thanks to careful extensions to the OLSR protocol ensuring its backward compatibility one is able to capitalise on the routing mechanisms and additional information readily available at the network layer for the purposes of optimising the performance of a link layer system employing the aforementioned virtual antenna arrays. The applicability of multi-point relay selection heuristic is outlined in Figure 2.

Analysing Figure 1 and Figure 2 one can see that both the concept of virtual antenna arrays and the multi-point selection heuristic overlap in the sense that multi-point relay stations can cooperatively constitute virtual antenna arrays,

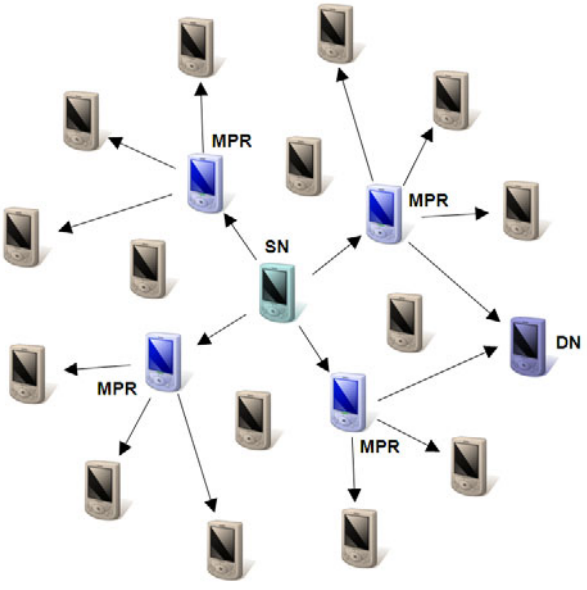


Fig. 2. Multi-Point Relay Stations

i.e. the nodes denoted ad MPRs can take the role of RNs. An important aspect here is that both approaches can be seamlessly integrated into an operating protocol which is continually further developed [3] and will be certainly deployed in the future. The reader is referred to [14] for additional details and analysis of this approach while more information about incorporating this idea as one of the key building blocks of the future autonomic networking will be outlined in the remainder of this paper. Especially on the verge of introducing the Future Internet envisioning wide use of the end user mobile communications devices one can expect the presented issue of cooperative transmission to become one of the key elements of the systems to be devised. The concept of assigning nodes to a virtual antenna arrays is particularly vital when viewed from the network perspective. In that case numerous concurrent cooperative transmissions may occur at the same time and the network itself will have to handle the provision of certain quality of service not only according to user demands but also keeping in mind the policies imposed by a network operator. Autonomic networking seems to be a straightforward approach to addressing this question.

3 Autonomic Cooperative Networking

Autonomic networking has emerged as one of the most promising approaches towards the instantiation of the self-managing Future Internet [2]. Although it attracts a lot of attention, the notion of autonomicity should not be, however, mixed with cognition or the ability of being autonomous. An autonomic system is simply characterised by the ability to self-configure without a need for

any external intervention. On the contrary an autonomous network is expected to display certain dose of cognition which is a very interesting add-on to autonomicity and is in the scope of this work. The inherent feature of autonomic networking is a need for continuous monitoring so the network is able to self-configure according to the imposed policies and taking into account additional information about incidents that through proper fault-management can improve the service resilience [1]. The aforementioned factors are particularly important for mobile ad hoc networks which depending on the scenario might be characterized by a very dynamically changing topology affecting the ability of nodes to cooperate efficiently.

In general, such an autonomic network should behave like a living organism which is continually driven by a very large number of processes running on their own but remaining in close correlation without any specific need for intervention form a central entity for most of the time of operation. To map such a concept onto networked systems one needs to apply specific network engineering mechanisms which are currently undergoing pre-standardisation [11]. Particularly, the idea of control loops is applicable to such systems, where a decision element (DE) is controlling a managed entity (ME) based on a closed information flow and with the use of external monitoring and policies related data (Figure 3).

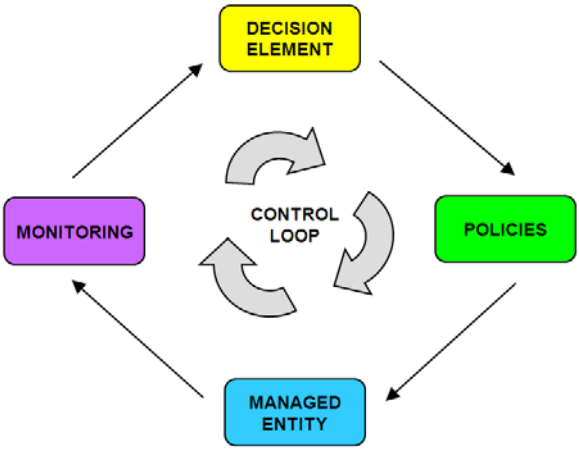


Fig. 3. Generic control loop

It is crucial to note that not only distinct nodes should express autonomic behaviours but the network should be autonomic as a whole. Consequently, it is expected that the network should continuously monitor itself and be able to allign its current operation with the requirements arising from monitoring data or simply imposed by (changing) policies. It means that distinct nodes should express autonomic behaviours but in specific cases decisions might need to be taken at the network level. As a result in certain circumstances the freedom of a decision element to make decisions based just on the available information can be

limited by directions given by a higher level decision element. The decisions might need to be taken at the network level in order to make it feasible for the network to perform tasks of a global scope. At the same time different operations, such as cooperation between or among nodes can be carried out without any interruption as long as this does not result in a violation of the rules. Whether to exercise autonomic behaviours at the node and/or at the network level depends on the character of the environment. Full ability to exercise autonomicity at nodes might be of prime importance for mobile ad-hoc networks formed by the end use where it is the user to cover the costs. On the contrary, one would probably emphasise the autonomic network features when network operator devices are concerned. Ideally, the network should be autonomic and at the same time the autonomic behaviours of distinct nodes should not be affected. It means that in case there is a need for a mobile ad-hoc network to request a node to act as a router, the node should be offered some benefits (e.g. free-of-charge access to services, et.) that could compensate, e.g., for quicker battery drainage. One can indeed think about certain mechanisms for motivating the nodes to autonomously trade their level of autonomicity for some benefits in order to make it possible for the whole group to meet predefined goals. An autonomic network formed of autonomic components should be then able to carry out certain internal negotiations in order to reach a common agreement. This is done through the interactions among decision elements and as a result service is more likely to be guaranteed in situations where typically it would be very difficult or even impossible to be handled.

The question of service provision in future autonomic cooperative networks has many very interesting aspects and the cooperation between/among nodes is one of them [10]. Similarly to the aforementioned situation, where specific users might not be willing to agree to become routers, also nodes that could guarantee the required level of service through mutual cooperation at the link layer might not want to do so. There are many degrees of freedom and mechanisms for sorting similar problems out need to be devised. Again the notion of autonomicity can be instantiated here through the interaction between the decision element and its managed entity, which might be e.g. a routing protocol such as OLSR. For the OLSR protocol it is possible to define willingness of a node to carry and forward traffic. This is one of the links to introducing autonomicity. Another aspect is the multi-point relay station selection heuristic which uses this parameter and can be aligned with the concept of virtual antenna array aided cooperation through distributed space-time block coding. The issue of willingness has an implication on the type of cooperation. Actually, cooperative transmission might not always be possible and there might be an urgent need to schedule different cooperative (Figure 4a and Figure 4b) or even non-cooperative transmissions (Figure 4c) at the same time.

The presentes approach to autonomic cooperative networking in the context of service provision brings up some open questions. First of all, assuming node cooperation to be one of the components thanks to which system performance and coverage can be improved in mobile ad-hoc networks, one ends up with a question regarding the feasibility of encouraging end users carrying their battery

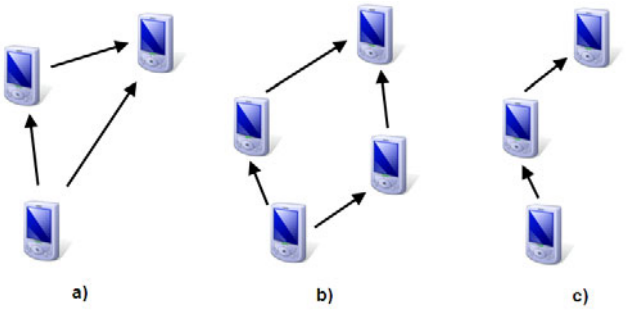


Fig. 4. Potential transmission modes

power limited devices to enter into cooperation. Secondly, assuming this can be done through incentives an optimisation problem arises on how to schedule multiple concurrent cooperative and non-cooperative transmissions, especially when different types of cooperation and relaying are allowed. Finally, making an attempt to capitalise on the paradigms of autonomic networking one comes across the issue of stability and scalability of such an overlay system composed of numerous interacting control loops.

4 Conclusion

Both the idea of cooperative transmission and autonomic networking seem to be the crucial building blocks for the Future Internet. It is envisaged that depending on the situation and policies a network of the future should be able to self-organize in an autonomic way so the expected service provision criteria are met. These criteria may be imposed e.g. by routing, i.e. the network is expected to be able to self-organize so the quality of service requirements for different flows are fulfilled. To this end, the network is expected to schedule autonomically an optimum combination of the aforementioned concurrent cooperative and non-cooperative transmissions and make any necessary updates on the fly. This is crucial for networks composed of battery constrained user devices such as laptops which might not necessarily be willing to offer their resources to be shared through cooperation because of the risk of draining the battery power more quickly. This might have a negative influence on the services offered and special arrangements are necessary so users are encouraged to participate by being granted special incentives. It is then also expected that a network will be able to observe the behaviors of different nodes and learn from this knowledge through cognition.

Acknowledgments. This work has been partially supported by EC FP7 EFIP-SANS project (INFSO-ICT-215549).

References

1. Liakopoulos, A., Zafeiropoulos, A., Polyrakis, A., Grammatikou, M., Gonzlez, J.M., Wódczak, M., Chaparadza, R.: Monitoring Issues for Autonomic Networks: The EFIPSANS Vision. In: European Workshop on Mechanisms for the Future Internet (2008)
2. Chaparadza, R., Papavassiliou, S., Kastrinogiannis, T., Vigoureux, M., Dotaro, E., Davy, K.A., Quinn, M., Wódczak, M., Toth, A.: Towards the Future Internet - A European Research Perspective. In: Tselentis, G., Domingue, J., Galis, A., Gavras, A., Hausheer, D., Krco, S., Lotz, V., Zahariadis, T. (eds.) Creating a viable Evolution Path towards Self-Managing Future Internet via a Standardizable Reference Model for Autonomic Network Engineering, IOS Press, Amsterdam (2009) ISBN: 978-1-60750-007-0
3. Clausen, T., Dearlove, C., Jacquet, P.: The Optimized Link State Routing Protocol version 2. draft-ietf-manet-olsrv2-10 (September 2009)
4. Clausen, T., Jacquet, P.: Optimised Link State Routing Protocol (OLSR). RFC 3626 (October 2003)
5. Dohler, M., Gkelias, A., Aghvami, H.: A resource allocation strategy for distributed MIMO multi-hop communication systems. IEEE Communications Letters 8(2), 99–101 (2004)
6. Doppler, K., Osseiran, A., Wódczak, M., Rost, P.: On the Integration of Cooperative Relaying into the WINNER System Concept. In: 16th IST Mobile & Wireless Communications Summit (July 2007)
7. Doppler, K., Redana, S., Wódczak, M., Rost, P., Wichman, R.: Dynamic resource assignment and cooperative relaying in cellular networks: Concept and performance assessment. EURASIP Journal on Wireless Communications and Networking (July 2007)
8. Dottling, M., Irmer, R., Kalliojarvi, K., Rouquette-Leveil, S.: Radio Technologies and Concepts for IMT-Advanced. In: Dottling, M., Mohr, W., Osseiran, A. (eds.) System Model, Test Scenarios, and Performance Evaluation. Wiley, Chichester (2009); ISBN: 978-0-470-74763-6
9. Laneman, J.N., Wornell, G.W.: Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks. IEEE Transactions on Information Theory 49(10), 2415–2425 (2003)
10. Chaparadza, R., Rebahi, Y., Wódczak, M., et al.: Demystifying Self-awareness of Autonomic Systems. In: ICT Mobile Summit (2009)
11. Wódczak, M., Chaparadza, R., Ciavaglia, L., et al.: ETSI Industry Specification Group on Autonomic network engineering for self-managing Future Internet (ETSI ISG AFI). In: 10th International Conference on Web Information Systems Engineering (September 2009)
12. Wódczak, M.: On the Adaptive Approach to Antenna Selection and Space-Time Coding in Context of the Relay Based Mobile Ad-hoc Networks. In: XI National Symposium of Radio Science URSI, pp. 138–142 (April 2005)
13. Wódczak, M.: On Routing information Enhanced Algorithm for space-time coded Cooperative Transmission in wireless mobile networks. PhD thesis, Faculty of Electrical Engineering, Poznan University of Technology, Poland (September 2006)
14. Wódczak, M.: Extended REACT Routing information Enhanced Algorithm for Cooperative Transmission. IST Mobile and Wireless Communications Summit (June 2007)

An Autonomic Monitoring Framework for QoS Management in Multi-service Networks

Constantinos Marinos, Christos Argyropoulos,
Mary Grammatikou, and Vasilis Maglaris

Network Management & Optimal Design Laboratory (NETMODE)
School of Electrical & Computer Enigineering
National Technical University of Athens (NTUA)
Tel.: +30 210.772.1451, Fax: +30 210.772.1452
{cmarinos,cargious,mary,maglaris}@netmode.ntua.gr

Abstract. Autonomic monitoring procedures in multi-service networks provide not only feedback to end users, but also self-handling monitoring events to network operators. In this work, we present an autonomic monitoring framework for Quality of Service (QoS) management in multi-service networks. Our framework introduces aggregation mechanisms to deal with the excessive number of alarms, triggered in an autonomic networking environment. The proposed framework was assessed via an early prototype, deployed to IPv6 end-sites, distributed across Europe and interconnected via the Internet.

Keywords: autonomic monitoring, QoS management, multi-service networks.

1 Introduction

End-to-end Quality of Service (QoS) for multi-domain service calls remains a challenge for next generation networks, as various real-time and bandwidth-sensitive applications are all migrating to an IP based architecture. Popular applications such as Video on Demand, Voice over IP, IPTV, online gaming require strict QoS and underlying Quality of Experience.

Increasing complexity and size of computer systems and networks lead to more complicated and distributed management systems. Moreover, multiple end-node services probe network elements or the operating system, in many cases simultaneously, with no coordination. Hence unreliable measurements may be taken causing unpredictable system behaviors or service degradation due to tuned measurements. Autonomic computing and autonomic networking have been arisen as the key concept in the effort of complexity confrontation. Autonomic computing and networking exhibit several self-management properties that will characterize autonomic systems like: self-configuration, self-healing, self-optimization and self-protection [1], [2], [3]. One of the main pillars of the autonomic networking is the self-monitoring mechanism. Self-monitoring is a fundamental operation in the autonomous systems providing not only a feedback to the other entities of the autonomic framework for the changing conditions, but also providing the capability of self-handling health measurements, network measurements and alarms. When an incident occurs in an

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 79–86, 2011.

autonomic environment self-management entities accomplish their functions by taking the appropriate actions, without human intervention. Following the design principles of an emerging Generic Autonomic Network Architecture (GANA) [4] for autonomic networking we present the architecture of a framework that orchestrates several monitoring entities, while handling and controlling the monitoring processes.

Advanced monitoring tools that have been developed, aiming to achieve more accurate monitoring alarms and notifications, take into account the network topology and services dependencies. Mechanisms that have as an input host and network dependencies use this information to diagnose fast and accurately, where and what problem occurred within the network. When an incident is classified as a host or service event, according to specific thresholds, the monitoring mechanism activates multiple checks. The self-monitoring mechanism defines if the specific service is actually the root cause of the emerged event or another one exists that produce collateral effects.

This paper focuses on the design and development of an autonomic monitoring framework[1] to maximize the experienced QoS of the end user during an inter-domain application. In this work we propose a dynamic self-monitoring mechanism introducing the Orchestration Engine entity with self-adapting characteristics. The main goal is to design a mechanism that it could have the ability to be self-adjusted in a varying network topology and different services. Furthermore the proposed entity will have the ability to monitor the overall node health status by collecting information, keeping monitoring records and aggregating information in different time instances, collected by different sensor interfaces.

The modular architecture of the Autonomic Monitoring Framework is based on the separation of monitoring tools and business logic mechanism. This permits the harvest of different monitoring data from different protocols and mechanisms. The different interfaces act as middleware between monitoring architectures of heterogeneous networks, regardless of the protocols and data models used by each network. In addition, the monitoring framework should have the ability to set a hierarchical map of dependencies for the hosts and the services that are under surveillance according to different policies.

The rest of the paper is organized as follows: Section 2 presents related work; Section 3 describes the autonomic framework along with its core components and its architecture for monitoring procedures. In section 4 the evaluation and testing of the framework is performed through a multi-domain application scenario. Finally in section 5 we provide some concluding remarks, including challenging issues that are part of our current and future work.

2 Related Work

To date several frameworks have been proposed to support the monitoring for specific end-to-end (e2e) network metrics like packet loss, RTT, delay, bandwidth, jitter. Each framework has its own methodology based on installing measurement points either at

[1] According to [14] "Frameworks are a special case of software libraries in that they are reusable abstractions of code wrapped in a well-defined Application programming interface (API)".

the end nodes (sites), i.e. the receiver or the sender, or by setting a measuring point at an intermediate node along the path. Notice that some of the methods make assumptions for a part or the whole e2e path, based on intermediate measurement points.

Monitoring specific QoS metrics between client and server is a well known problem. The authors of [11] outline a method that measures round-trip-time in the three-way handshake at the beginning of every TCP connection. In [12] two methods for measuring RTT from an intermediate node are examined.

With respect to the above proposals, our framework differentiates from the others in the sense that it introduces an autonomic mechanism for the detection and reconfiguration of the measurement mechanisms. Additionally, our monitoring framework has been deployed in IPv6 sites, hosted by four European partners of the FP7 European Commission Project EFIPSANS [1].

3 Framework Architecture

In this section, we describe the main principles that must be followed by a multi-domain performance monitoring tool for accurate measurements. We then describe the logical view of our architecture schema and present the development of a framework prototype.

Our monitoring framework should be supported in each end-site of a service path. In the multi-domain environment of the Internet, support of the framework in intermediate points of the path would enhance the accuracy of e2e measurements. Well defined APIs carry out the communication between frameworks of the involved sites. When an alarm or notification occurs the monitoring framework collects and handles the event without administrators interfere. In this way, an autonomic approach is achieved with the integration of the corresponding entities.

3.1 Design Objectives

Three different methods of monitoring are being used to aggregate network level performance assessment: active, passive and piggybacking. Modern network infrastructures use these main techniques in order to collect and analyze measurements.

Active measurements require test-packet generation into the network. Traditionally, active measurements include *ping* and *traceroute* utilities. More sophisticated tools like Multicast Beacons [5] have been developed, which emulate application specific traffic and use the obtained results of performance to estimate the end-user Quality of Service. More popular tools for active measurements are iperf [6], bwctl [7], owamp [8] that use sophisticated packet probing techniques.

In comparison to the active measurements, the passive measurements do not produce test packets. They require capturing of packets and their corresponding timestamps transmitted by applications running on network attached devices over various network paths. Some of the popular passive measurement techniques include collecting Simple Network Management Protocol (SNMP) and NetFlow data from network switches and routers. Piggybacking is at the moment mostly a research-oriented approach with time stamping and sequence numbering information being used.

Our Monitoring Framework uses specific monitoring tools that should follow some fundamentals architectural guidelines in order to be in conformance with monitoring

techniques. A monitoring tool that will be installed in an end-to-end path must not interfere with the application. If the monitoring tool generates traffic in order to measure the end-to-end path, that traffic must be low enough so that we will not have any interjections with the running applications. In addition, the tool should support the on-demand measurements at any time. It should be simple and easy to be installed in any node across the network path without consuming too much hardware or network resources. Finally, it should be IPv6 enabled in order to call IPv6 addresses.

Scalability is another critical design goal. At any given time, a varying number of users can make a varying number of measurements. The monitoring framework should be able to support multi-service optimization. This means that the framework should not maximize one service, if this leads to discrimination of another.

3.2 Architecture Overview

In order to build our monitoring framework and manage the Quality of Service along a requested service path, we have applied the three-tier architecture [9]. The core elements of the Monitoring Framework are the Orchestration Engine, the Policy Handling Module and the Event Correlation Module.

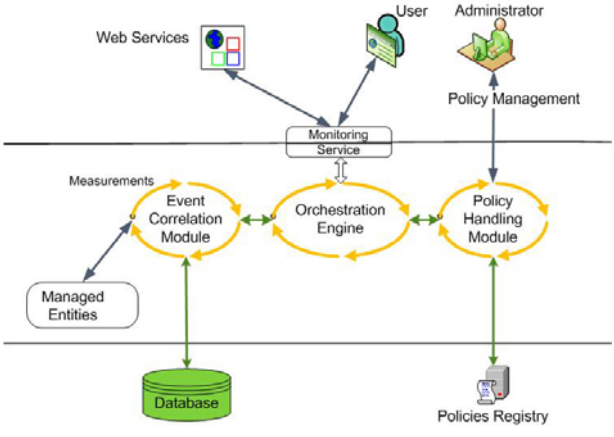


Fig. 1. Framework Architecture

Today's practice of segregated monitoring mechanisms leads to multiple service requests for QoS-related measurements generating redundant measurement data. For this reason the lowest tier, referred to as the Persistence Layer [13], of our Monitoring Framework consists of a database that stores all measurements collected from the different monitoring entities. This database also contains the different policies derived from the Policy Handler and related statistical analysis results. The Policy Repository contains data for the network and system inventory that every service can access. By continuously collecting this information we build up knowledge about the performance of a service like *ftp* or *video on demand* between two specific end-sites. The monitoring entities belong to the Middleware Layer feed the Event Correlator with periodic values for all the network metrics.

The Policy Handler implements the manipulation mechanism of the monitoring architecture. Policies originate either from the Administrator or the entities of the autonomic system with the authority to edit policy profiles. These entities are referred to as *decision elements* in [4].

The Policy Handler is responsible for informing the Orchestration Engine for the policies that govern the different services: It retrieves the different parameters and thresholds that must be fulfilled depending on the requested service, from the database and informs the Orchestration Engine. If, for example, there is a request from the Orchestration Engine for an *IPTV* stream servicing an H.264 video movie file with 24 frames/sec in standard definition, the Policy Handler will respond with a minimum requirement of 15 Mbps bandwidth. The administrator can have access to the policies – add, edit, delete - that are stored in the database through an administrative interface, if he needs to change a specific policy. An additional interface of the Policy Handler permits policy manipulation from an upper layer of self-manageability (autonomicity), such as the *node level* or *network level* in [4]. The existence of this second interface permits the adaptation of the Monitoring Framework to a more generic Reference Model for Self-Managed Networks.

The Orchestration Engine is the core entity that is responsible for the alarm propagation outside the monitoring framework and to respond to any quality control request coming from a user service request. It handles the role of the logic entity which is responsible for the smooth monitoring operation inside the Autonomic Architecture. It is responsible to satisfy any monitoring request providing system and network measurements with an efficient non-redundant way.

Moreover the Orchestration Engine takes care of active measurements scheduling and handling. QoS-related measurements in many cases are CPU and bandwidth demanding, in order to reveal bandwidth bottlenecks and packet forwarding prioritization mechanisms [10]. It schedules active measurements based on information from already stored data; new measurements are executed if data are out-to-date according to thresholds defined by the Policy Handler.

The Policy Handler defines thresholds of measurements concerning specific service policies. It accordingly initiates notifications and alarms. The Orchestration Engine is in charge of deciding for: (i) the scheduling of measurements, (ii) the severity of an alarm, (iii) the impact of the event that triggered an alarm, (iv) notifications that need to be propagated outside the monitoring entity and their verbosity, (v) the abortion of an alarm in case of flapping event state, (vi) the handling of an event without further interactions with other entities.

The Orchestration Engine's Logic combines input from the Event Correlator and Policy Handler to make decisions related to its aforementioned six tasks. The required logic rules manage the measured data from the Event Correlator according to policies taken from the Policy Handler.

The Event Correlator is responsible for: (i) generating messages, (ii) sending them to the Orchestration Engine, (iii) managing the suppression and aggregation of the measurements, (iv) correlating multiple measurements and statistics from different sources (monitoring entities), that form a logical set, escalate the severity of a notification and create a new notifications, (v) binding repeated measurements, (vi) correlating measurements based on service dependencies, (vii) suppressing transient measurements. It may combine inputs directly from the different Monitoring Entities, but also from monitoring data stored to a local repository.

Requests that could trigger the Orchestration Engine could be a query from a user that requests a custom measurement between two nodes during a streaming service; or it could be a notification from the Event Correlator about a service threshold violation. In the first case, the Orchestration Engine would send a message to the Event Correlator in order to notify the responsible monitoring entities for the custom measurement. In the second case, if a violation of a threshold service would take place, a notification would be sent to the Orchestration Engine from the Event Correlator. Then the Orchestration Engine could re-initiate the service and inform the administrator for the violation, in order to change the service settings, reconfigure the service, or start troubleshooting process, via the notification interface of the Monitoring Framework.

4 Framework Evaluation

To evaluate our framework, we used a test-bed consisting of four IPv6 enabled nodes, distributed across Europe in partner sites of the EFIPSANS project [1]. A large file was transferred over the Internet; the end-nodes were acting as measurement points with Monitoring Frameworks installed on them. Measurements were collected in a main node, located at the NETMODE Laboratory, National Technical University of Athens (NTUA). The other nodes were hosted by Telefónica in Madrid, the Telecommunications Software & Systems Group (TSSG), Waterford Institute of Technology, Dublin and the Greek Research & Technology Network (GRNET) in Athens.

An Apache HTTP server running at NTUA was playing the role of the framework User Interface (UI). Through this UI, a user could take on-demand measurements and see graphical representations of the various network paths through the four nodes. Different scenarios with respect to packet loss, one-way delay and jitter were tested between different node pairs. For each network pair we ran periodically measurements continuously for 30 days. In the following figures we present some sample measurements obtained by using our Monitoring Framework.

In Figure 2, we show the Bandwidth (Mbps) measurements between two IPv6 nodes. As we can see, it exhibits significant variations ranging from 10Mbps to 70Mbps.

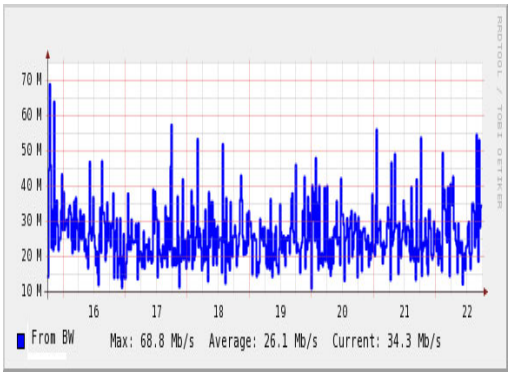


Fig. 2. Bandwidth (Mbps) representation

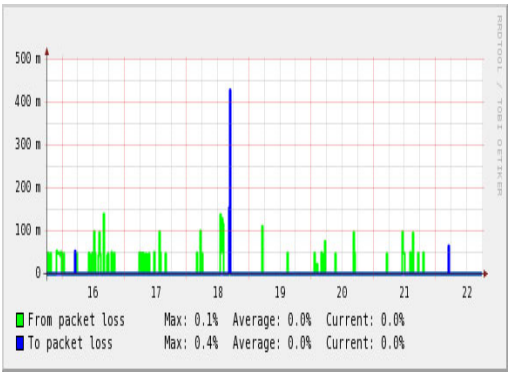


Fig. 3. Packet loss (%) representation

In Figure 3 we show Packet Loss during a week in both directions. It is mostly 0%, but frequently exhibit peaks in both directions. The difference in the two curves may be due to the asymmetric nature of the two directions between the two nodes.

Finally, the jitter (ms) representation is shown in Figure 4. In one direction we can notice that we have almost no jitter with some peaks in some cases, while in the opposite direction we have an average jitter of 1ms.

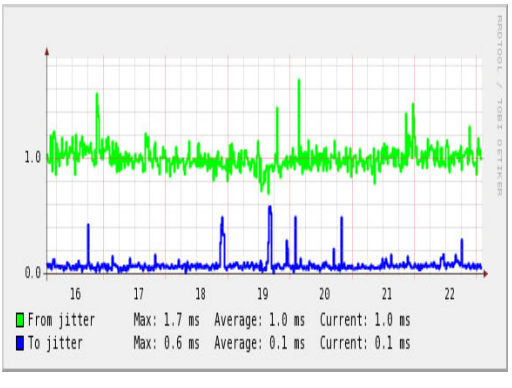


Fig. 4. Jitter (ms) representation

5 Conclusions

In this paper we propose an autonomic Monitoring Framework to assess end-to-end QoS in multi-service networks. As a proof of concept, we designed and developed an autonomic monitoring framework introducing the Orchestration Engine entity with self-adapting and self-monitoring capabilities.

We deployed our prototype in four IPv6 enabled sites across Europe. These sites were interconnected by multi-domain networks, i.e. the commercial Internet and over-provisioned high speed communication networks (NRENs and GÉANT). We

subsequently ran measurements on e2e QoS (packet loss, bandwidth, one-way delay and jitter). Our results are reported in this paper but require further analysis. We are currently considering several extensions, e.g. assessing Quality of Experience (QoE) from QoS metrics for demanding services (streaming of HD video, IPTV, VoIP and Online Games).

Acknowledgments. This paper was partially supported by EFIPSANS project (INFSO-ICT-215549) of the European Union's 7th Framework Programme for Research & Technological Development.

The authors wish to express their thanks to their EFIPSANS collaborators in Madrid, Dublin and Athens that contributed in the test-bed deployment.

References

1. EFIPSANS project, http://www.efipsans.org
2. Kephart, J.O., Chess, D.M.: The vision of autonomic computing. Computer 36(1), 41–50 (2003)
3. Dobson, S., Denazis, S., Fernandez, A., Gaiti, D., Gelenbe, E.: A Survey of Autonomic Communications. ACM Transactions on Autonomous and Adaptive Systems (TAAS) 1(2) (2006)
4. Chaparadza, R.: Requirements for a Generic Autonomic Network Architecture (GANA). Suitable for Standardizable Autonomic Behavior Specifications for Diverse Networking Environments. IEC Annual Review of Communications 61 (2008)
5. Type of Service field in IP packets, http://en.wikipedia.org/wiki/Type_of_Service
6. IPerf, http://www.noc.ucf.edu/Tools/Iperf/
7. Bandwidth Controller, http://www.internet2.edu/performance/bwctl/
8. One-Way Active Measurement Protocol, http://www.internet2.edu/performance/owamp/
9. Three-tier Architecture, http://en.wikipedia.org/wiki/Multitier_architecture
10. Lu, G., Chen, Y., Birrer, S., Bustamante, F.E., Li, X.: POPI: A User-Level Tool for Inferring Router Packet Forwarding Priority. IEEE/ACM Transactions on Networking 18(1), 1–14 (2010)
11. Jiang, H., Dovrolis, C.: Passive estimation of tcp round-trip times (2002)
12. Veal, B., Li, K., Lowenthal, D.: New methods for passive estimation of TCP round-trip times. In: Dovrolis, C. (ed.) PAM 2005. LNCS, vol. 3431, pp. 121–134. Springer, Heidelberg (2005)
13. Persistent Data Structure, http://en.wikipedia.org/wiki/Persistent_data_structure
14. Software Framework, http://en.wikipedia.org/wiki/Software_framework

Safetynet Version 2, a Packet Error Recovery Architecture for Vertical Handoffs

Henrik Petander[1] and Emmanuel Lochin[2,3]

[1] NICTA, Sydney, Australia
[2] CNRS, LAAS, 7 avenue du colonel Roche, F-31077 Toulouse, France
[3] Université de Toulouse, UPS, INSA, INP, ISAE, LAAS, F-31077 Toulouse, France

Abstract. Mobile devices are connecting to the Internet through an increasingly heterogeneous network environment. This connectivity via multiple types of wireless networks allows the mobile devices to take advantage of the high speed and the low cost of wireless local area networks and the large coverage of wireless wide area networks. To maximize the benefits from these complementing characteristics, the mobile devices need to be able to switch seamlessly between the different network types. However, the switch between the technologies, also known as a vertical handoff, often results in significant packet loss and degradation of connectivity due to handoff delay and also increased packet loss rate on the border of the coverage area of the networks. In our previous work, we have proposed an inter technology mobility management architecture which addresses the packet losses using selective resending of packets lost during the handoff period. In this paper, we extend the architecture to address packet losses due to wireless errors more efficiently by taking advantage of erasure codes to form redundancy packets. We propose to send these redundancy packets over both links. We show that this proposal reduces both the chances of packet loss and the buffering requirements of the original SafetyNet scheme.

1 Introduction

With the proliferation of new wireless access network technologies, mobile users can now access the Internet using multiple types of access network technologies. This heterogeneous network environment provides access through a varying range of network technologies. The characteristics of these access networks vary greatly; Wireless Local Area Networks (WLANs) provide high speed access with a network latency of tens of milliseconds, often at the price of fixed Internet access but with a very limited coverage. Wireless Wide Area Networks (WWANs) on the other hand provide wide coverage but have a significantly lower data rate, higher latencies up to several hundreds of milliseconds and a cost which may several magnitudes larger than that of WLAN networks. Therefore, it is beneficial for a mobile user to be able to switch seamlessly between the different technologies.

Seamless switching between heterogeneous access networks requires carefully managed vertical (inter-technology) handoffs. Protocols, such as Mobile IP [4], can be used for ensuring the handoff does not break the on-going connections of a

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 87–95, 2011.

mobile node and that the mobile node remains reachable in spite of the handoff. However, the vertical handoff performance of Mobile IP often leads to significant disruption of on-going traffic [1]. Earlier work in the field of mobility management has mostly focused on minimizing the impact of horizontal handoffs, i.e. moving between two networks of the same technology and does not provide optimal performance in vertical handoffs.

In our previous work, we proposed a localized mobility management protocol, *SafetyNet* [2] for minimizing the impact of vertical handoffs. The SafetyNet protocol utilises make-before-break handoffs, in which the mobile node breaks its connectivity with the previous access router (PAR) only after connecting to the new access router (NAR) which makes it possible to perform lossless handoffs. However, upward vertical handoffs, i.e. handoffs from WLAN to WWAN networks are typically performed only when the signal to noise ratio for the WLAN has degraded to nearly unusable. This poor signal strength of the WLAN may result in packet losses due to wireless errors or complete loss of connection with the PAR during the time it takes to prepare the WWAN interface and link layer connection to the NAR. To address this issue, SafetyNet combines the make-before-break handoffs with buffering at the NAR with selective delivery of packets from the buffer, so that any packets lost on the previous link (between the mobile node and PAR) are delivered from the buffer of the NAR at the new link. This allows for the recovery of the packets lost during the handoff period due to wireless errors or loss of connection with the previous access router. This mechanism is described in more detail in the next section.

The SafetyNet protocol showed a significant performance improvement over Fast handovers for Mobile IPv6 [5] in empirical measurements in favorable conditions for a vertical handoff scenario both in terms of TCP performance and in terms of over-the-air overhead. However, the worst case performance of Safetynet is close to that of Fast Handovers for Mobile IPv6 protocol. In this paper, we target the average and worst case performance of the protocol by replacing resending of packets with use of packets based forward error coding (FEC) [6] both on the link of the previous router and on the link of the new router. We show that by using an adaptive coding scheme we can reduce the cost of recovering lost packets significantly while improving the Safetynet architecture in terms of processing.

In the next section, we firstly present an overview of the SafetyNet protocol and then discuss its limitations in section 3. Our contribution is given in section 4 where we detail an improved solution and provide a preliminary performance evaluation. Finally we present relevant related work in section 5 and conclude this work in section 6.

2 Overview of the SafetyNet Protocol

In the SafetyNet protocol [2], a Mobile Node (MN) moving from a link connected to the Previous Access Router (PAR) to a link connected to a New Access Router (NAR), initiates the vertical handoff from PAR to NAR with PAR when it senses

that it is about to lose connectivity with the PAR. The handoff process shown in Figure 1 is described briefly below.

- A mobile node which performs a handoff from an old network to a new network will first initiate the handoff with the access router of the previous network (PAR). In a vertical handoff, the mobile node would start activating the network card and preparing a link layer connection with the router of the new network link (NAR) at the same time.
- After a negotiation with NAR, PAR starts delivering copies of packets destined to the old location (care-of address) of the mobile node to the NAR by using tunneling. The PAR labels the packets using a sequence number, so that both the original and the tunneled copy of each packet have the same sequence number. The sequence number is incremented for each new packet. NAR stores these packets in a buffer.
- Mobile node observes the sequence numbers of packets it receives during the handoff at the old location on the link of the PAR. When the new link is ready, it signals the NAR to deliver all the packets from its buffer that the mobile node did not receive during the handoff at the link of the PAR. In Fig. 1, this would be packets starting from sequence number 2.

The details of the protocol are described in more detail in the original article [2].

3 Problem Statement

The original SafetyNet architecture has four main issues. One of them deals with the rate at which the next AR (NAR in Fig. 1) buffer is filled. In the case where a mobile node moves from a WLAN to a WWAN, the rate at which packets arrive may often be higher than the rate at which they can be sent in the new network and as a result, the rate at which the buffer is emptied. Indeed, the new WWAN network may often combine a higher delay and a lower bandwidth than the old WLAN network, and this can lead to a significant loss of packets due to buffer overflow. The buffering has a second problem: The buffer space required is the product of the data rate for a user at the old link times the handoff duration. For vertical handoffs from WLAN to WWAN requiring activating a WWAN interface from a sleep state, this buffer space can amount to several megabits which may be a problem, if a large number of users are moving simultaneously. The third issue is the cost of recovering lost packets. The cost of retransmitting packets over a wide area network, such as a UMTS network, is typically significantly higher than the cost of (re)transmitting the same packets over WLAN networks. Additionally, the retransmission over the WWAN link increases the latency of recovering lost packets. Thus, it is beneficial both from a cost and performance point of view to retransmit less packets over the expensive WWAN networks. A possible solution to this is to enforce the reliability of the transmitted packets before the handoff and during the packet duplication process (See Fig. 1).

Looking at Fig. 1, a naive solution would be to retransmit packets during the bicast process from PAR to MN. Although this could be a solution over a WLAN network, it would be a problem over WWAN networks usually characterized by

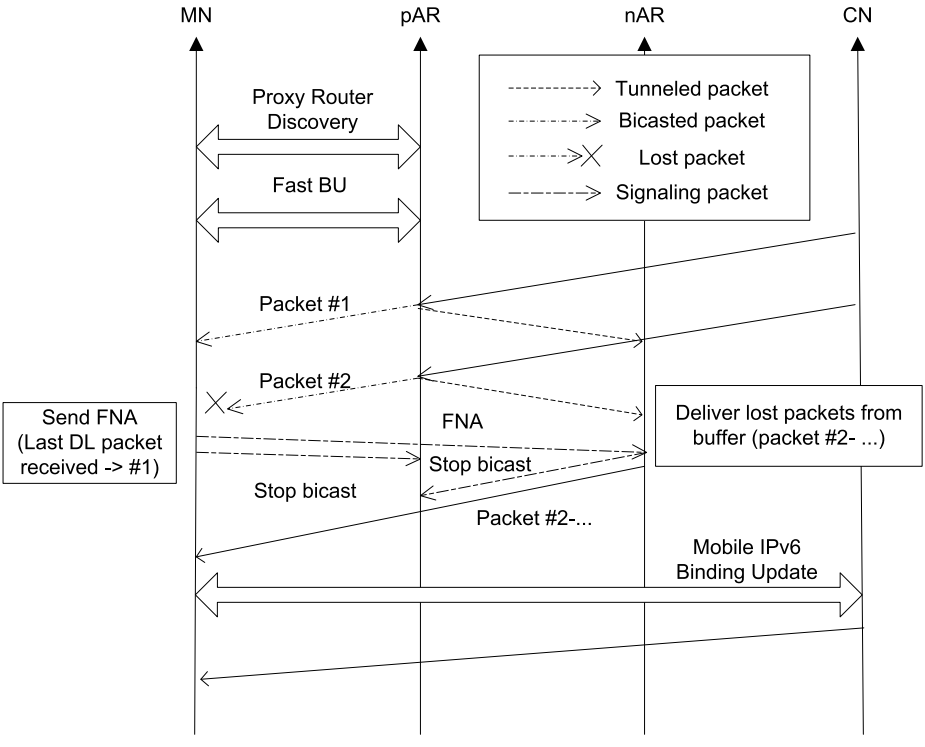


Fig. 1. The SafetyNet protocol operation

a high delay, low bandwidth and a higher cost of transmission due to the long delays of the retransmissions. A mobile node performing a vertical handoff from a WLAN would only do so when the connectivity of the WLAN became marginal, i.e. the bit error rate (BER) increased. Therefore, the mobile node would receive the packets with less resending via the NAR and a new link with a lower BER. This is illustrated in Fig. 2 which shows that even with a very high frame error rate (FER) of 25% on the new link, use of SafetyNet is less costly in terms of bandwidth when compared with resending over the IEEE 802.11 WLAN link. Even with a higher cost of communications over the WWAN link, it would still make sense in most cases to use SafetyNet to recover the packets. Therefore, we decided to not tackle this possibility.

To briefly summarize, the main points we would like to improve are:

- to minimize the buffer occupancy at the NAR side in order to avoid dropped and delayed packets;
- when possible, to enforce the reliability of the transmitted packets during the handoff process in order to decrease the number of the retransmitted packets by the NAR;
- and to achieve this, without introducing overhead in terms of number of packets sent over both networks.

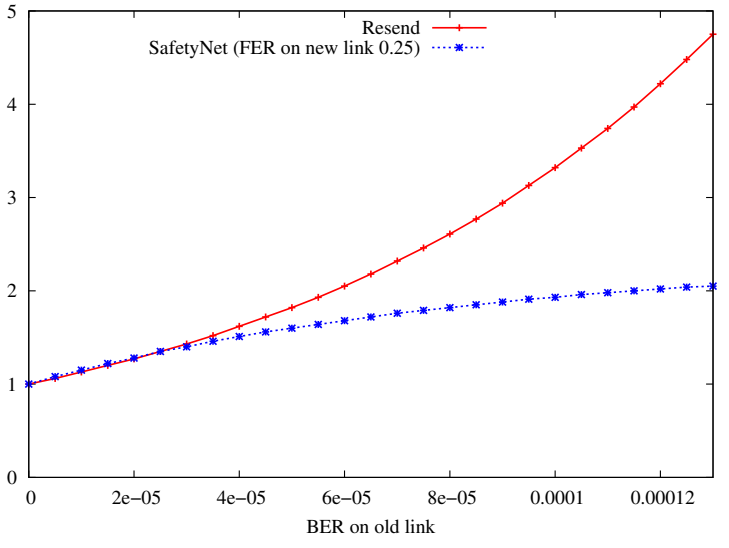


Fig. 2. Comparison of bandwidth cost of resending lost packets over the WLAN link with the use of SafetyNet

In this paper, we propose to enhance the existing SafetyNet protocol by using an erasure code scheme. The use of such a mechanism leads to potential performance improvements both in terms of reliability during the handoff process and minimization of buffer occupancy. In this context, we propose to use such code both by the PAR and NAR. Indeed, we propose to introduce some redundancy packets on the PAR side during the handoff and the bicast process to decrease the number of lost packets transmitted before handoff and to send only redundancy packets to the NAR. Obviously, by sending only redundancy packets to the NAR, the buffer occupancy will decrease markedly.

4 Redundancy Packet Bicasting Scheme

Most of forwarding error codes used over packet erasure channels are block codes [3]. This means that at the encoder side, a set of repair packets (R) is built from a given set of source data (SD) packets and at the decoder side, these repair packets can only be used to recover SD packets from their corresponding set. We denote k the number of SD and $n-k$ the number of R. If too many packets (among the SD and repair packets) are lost during the transmission, the recovery of the missing SD packets is then not possible (i.e $n-k$ redundancy packets allow to recover k packets among n where the k recovered packets correspond to the original SD packets). On the opposite, if only few packets are lost, some of the repair packets become useless. A solution to this problem, known as Hybrid FEC-ARQ (or H-ARQ) mechanism [3], is to use receiver's feedback to send

additional repair packets or to adjust the redundancy level of the FEC to the observed packet loss rate. In our context, retransmissions of redundancy packets are done, if necessary, by the NAR after receiving the FNA message. When the NAR receives the FNA message, containing sequence numbers of any lost packets, it can determine which redundancy packets need to be sent to the mobile node. Before the handover, the mobile node receives the traffic from PAR and moves towards the second network. During the handover, the bicasting procedure duplicates different linear combination of redundancy packets towards both networks. Finally, when the MN arrives in the new network and after sending the FNA feedback packet, the bicasting process ends and the communication is re-routed through the NAR.

To assess the right configuration of the FEC block code, an estimation of the PER is needed. However, including a mechanism able to estimate the exact PER during a handover procedure is not realistic and prevent from any real deployment. Furthermore, estimating a PER in WLANs is hard due to the unpredictability especially in indoor environments. Following our experiments [2] and experience, a PER above 25% makes it impossible to maintain TCP connectivity whatever the protection method used as the large amount of lost packets to retransmit or rebuild result inevitably in TCP timeouts. However, the use of a fixed PER allows getting around this problem and dealing with a wide range of PER with little additional overhead. Thus, we propose to set a default peak PER value on the AR device that can be tuned by the wireless network administrator and propose as a default value 20%. For instance, we use a $(k, n) = (4, 5)$ FEC code, which means that one R packet is built from a linear combination of four SD packets, these five packets are sent through the previous link and the bicasting procedure builds and sends another redundancy R', which is a second linear combination of the fourth SD, to the NAR. Basically, this means that we bicast supplementary linear combinations of redundancy packets to the NAR. As a result, we double the number of redundancy packets and half of them are used to correct losses on the wireless link (the estimated 20%) while the other part is simply stored on the NAR for retransmission purpose.

We previously said that a mobile node can move from either a slower to a faster network or the reverse. Usually, faster networks are WLAN networks which have very low delay compared to WWAN. In the Safetynet/FEC scheme, the rationale to bicast redundancy packets is: 1) to enforce the reliability of the flow on the previous link during the handover in order to reduce the amount of retransmission by the NAR; 2) to ensure all lost packets during the handoff will be rebuilt thanks to supplementary redundancy packets as soon as MN advertises to the NAR the sequence numbers of any missing packets (through the FNA feedback).

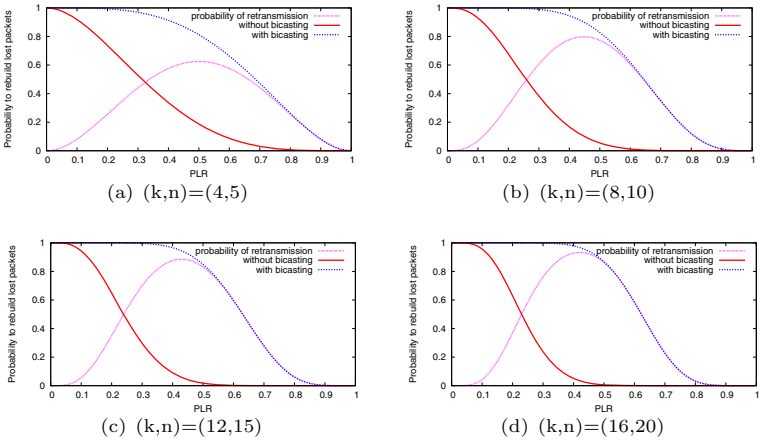
Without bicasting, the probability P to rebuild all packets is given by:

$$P = \sum_{i=k}^{n} \binom{n}{i} (1-p)^i p^{n-i} \tag{1}$$

This protects the flow to a PLR p when $p < \frac{(n-k)}{n}$. For a given FEC code (k, n), the bicasting process actually offers (k, n') protection where $n' = 2*n-k$. With the bicasting process, this probability, denoted P_{FEC} becomes:

$$P_{FEC} = \sum_{i=2(n-k)}^{n} \binom{n}{i} (1-p)^i p^{n-i} \tag{2}$$

and protects $p < \frac{2(n-k)}{n}$ with potential redundancy packet retransmissions. The theoretical results are given figure 3. In this figure, we draw the probability to rebuild a packet as a function of the PLR for a FEC code able to theoretically correct a PLR of 20% with and without bicasting following (1) and (2). The difference between these two curves allows us to assess the probability of redundancy packet retransmissions which is given by the third curve.



(a) (k,n)=(4,5) (b) (k,n)=(8,10)

(c) (k,n)=(12,15) (d) (k,n)=(16,20)

Fig. 3. Probability to rebuild a block as a function of the PLR for various FEC codes

As a matter of fact, the possibility to retransmit supplementary linear combination of packets allows to better protect the flow during the handover when the signal becomes poor while decreasing the buffer size of the NAR. Compared to the previous Safetynet proposal, this FEC scheme decreases the number of possible retransmissions as some packets are rebuilt thanks to the FEC redundancy packets. This combination allows to ensure that, even if FEC code is not able to correct all losses inside a block, a retransmission of redundancy packets via the NAR will allow to correct up to the double the PLR (i.e. up to $\frac{2(n-k)}{n}$) with the price of redundancy packets retransmission delay. The success of this mechanism is obviously linked to the configuration of the FEC code. Indeed, if a full block of packets is lost on the PAR link, the retransmission of supplementary redundancy packets from the NAR will not allow to rebuild the missing block (i.e. if $p > \frac{2(n-k)}{n}$). In this case, a retransmission from the source will be inevitable and would correspond to a wrong setting of the FEC parameters by

the administrator. However, Figure 3 shows that in our example, the probability to not recover more than k packets among n with a retransmission (which corresponds to $1 - P_{FEC}$) is nil when the PLR is below 20%, meaning that a rough configuration of the PLR, our scheme should cover this case.

5 Related Work

Forward error correction has been used earlier in soft handoffs in CDMA systems [7]. A mobile node in a soft handoff sends and receives two bit streams via two different base stations which are combined into a single stream at the receiver. The two streams are synchronized, and contain redundant information for handling transmission errors. The use of FEC in IP based communications handoffs has also been proposed earlier for seamless horizontal handoffs for multicast video traffic [8]. A more general solution for horizontal handoffs was proposed by Matsuoka et al. In their proposal which resembles the CDMA soft handoffs, a mobile node receives multiple IP packet streams encoded with a Reed Solomon code via different WLAN access points and combines them into a single stream with redundancy [9]. This work, which is so far the closest to our contribution, proposes to spread the redundancy over two different links. The main difference with our scheme is that the authors split the encoded streams while we propose to duplicate the FEC encoded stream with other linear packet combinations. This allows to increase the error capability correction compare to their scheme.

Finally, these approaches can not be directly applied to vertical handoff scenarios due to the different delay and bandwidth characteristics of the networks in a handoff. Further, they only address the time after the connectivity on the new network has been established. Our work focuses on vertical handoffs and uses a separate stream of selectively delivered redundancy packets to deal with the asymmetric delays and to address the potentially long handoff process itself.

6 Conclusion

In this paper, we presented an improved architecture for localized mobility management providing a mechanism for recovering packets lost during the handoff. The architecture applies forward error correction on packet level to reduce the buffer occupancy of the SafetyNet architecture while providing error recovery. We provided an initial evaluation on the efficiency of the mechanism which indicates that the architecture may provide recovery of lost packets at a smaller cost than the original SafetyNet proposal which already reduces the resending of packets significantly.

The analytical evaluation of the algorithm suggests that the use of forward error correction may be an interesting strategy for vertical handoffs. As a next step, we are planning to implement the algorithm as a part of the SafetyNet architecture and evaluate the performance empirically using our SafetyNet implementation.

References

1. Chakravorty, P., Vidales, K., Subramanian, I., Pratt, J., Crowcroft, J.: Performance Issues with vertical handovers-experiences from GPRS cellular and WLAN hot-spots integration. In: IEEE PerCom (2004)
2. Petander, H., Perera, E., Seneviratne, A.: Multicasting with selective delivery: A SafetyNet for vertical handoffs. Springer Journal on Personal Wireless Communication 43(3)
3. Lin, S., Costello, D.: Error Control Coding: Fundamentals and Applications. Prentice-Hall, Englewood Cliffs (1983)
4. Perkins, C.: IP Mobility Support for IPv4. Request For Comments 3344 (2002)
5. Johnson, D., Perkins, C., Arkko, J.: Mobility Support in IPv6. Request For Comments 3375 (2004)
6. Huitema, C.: Case for Packet level FEC. In: Proceedings of the TC6 WG6.1/6.4 Fifth International Workshop on Protocols for High Speed Networks (1996)
7. Bevan, D., et al.: Improved Soft Handoff Method For Uplink Wireless Communications. International patent no. WO/2004/004155 (2002)
8. Liu, H., et al.: A Staggered FEC System for Seamless Handoff in Wireless LANs: Implementation Experience and Experimental Study. In: IEEE ISM (2007)
9. Matsuoka, H., Yoshimura, T., Ohya, T.: End-to-end robust IP soft handover. In: IEEE ICC (2003)

A Mechanism for Vertical Handover Based on SAW Using IEEE 802.21*

Jorge Lima de Oliveira Filho and Edmundo Madeira

Institute of Computing – University of Campinas
Av. Albert Einstein 1271, 13083-970, Campinas, Brazil
{jlima,edmundo}@ic.unicamp.br

Abstract. Nowadays, there is a lot of devices which are able to access wireless networks through a wide range of access technologies. For a device to move among these heterogeneous networks and stay always connected, mechanisms of vertical handover are needed. This paper proposes a handover decision mechanism using the Simple Additive Weighting (SAW) in a heterogeneous wireless network environment using the IEEE 802.21. The proposed mechanism considers user preferences like cost as parameters of the candidate network to choose the best available network. We present some experiments that use a developed simulator to validate our mechanism. The results of these experiments show that the proposed solution distributes better the mobile nodes among the networks.

Keywords: Wireless Networks, Next Generator Networks, Handover, IEEE 802.21.

1 Introduction

Nowadays, there is a lot of devices and wireless network technologies. Each technology has advantages and drawbacks depending on the scenario, for instance, the use of a Wi-Fi network is suitable for wireless LANs. The WiMAX (Worldwide Interoperability for Microware Access) is suitable for WANs. The cellular networks will be based on the IP protocol and these new networks are known as Next-Generation Wireless Systems (NGWS) [1].

Currently, one of the main research challenges for next generation of IP wireless networks is the development of intelligent techniques for mobility management that take advantages of IP based technology to reach global roaming among wireless technologies [1]. In an environment of multiple access technologies, the concept of being always connected becomes always best connected (ABC) [2]. Many issues about network integration (composition) are a challenge to the research community.

The IEEE 802.21 standard [3,4] defines a common Media Independent Handover Function (MIHF) between layers 2 and 3, which enables mobility across

* The authors would like to thank FAPESP (2007/57336-0) and CNPq for supporting this work.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 96–108, 2011.

heterogeneous networks including both IEEE 802 and non-IEEE 802 networks. The main goal of this standard is to maintain connection during handover across different networks. Media Independent Handover Services define link-layer services to enable handovers among different radio air interfaces. Examples of 802.21 handovers are IEEE 802.11 networks to/from IEEE 802.16 networks, or IEEE 802 networks to/from cellular networks [5]. Despite offering mechanisms to the network, the 802.21 standard does not specify the handover decision algorithm.

In current wireless network technologies, such as IEEE 802.11 networks and cellular networks, handover decisions are usually based on the level of the received signal strength and IEEE 802.11 priority. The user's preferences are not taken into account and in the case of vertical handover, the signal can not be enough to decide the handover due to asymmetrical nature of radio technologies. In the next wireless network generation systems, the users can have a lot of networks alternatives and each one of them can have many features like cost, bandwidth, power consumption, etc, and the users must choose which is the best one to connect according to their needs.

In this paper we propose a mechanism for decision making of handovers that uses the SAW method. The proposed mechanism has three main parts: An algorithm for decision making handover, extensions of the IEEE 802.21 messages and the Information Server (IS) entity, proposed in this standard, and the creation of the SAW module inside the IS. These three main parts together are responsible for making the handover decision for the user.

The SAW (Simple Additive Weight) method is a type of Multiple Attribute Decision Making (MADM). The MADM refers to making decisions in the presence of multiple, usually conflicting criteria [6]. There are many methods to solve MADM problems, for instance, ELECTRE, TOPSIS, AHP (Analytic Hierarchy Process), etc. The SAW method has been chosen as a suitable option to solve this kind of problems, as discussed in [7]. The SAW method is used to rank the available networks to the mobile nodes. The users preferences like cost and bandwidth are used as handover parameters in order to make the decision. Besides these parameters, the network current load is considered to achieve load balancing, thus improving the network utilization. All computing processing for handover decisions is performed by a network entity, consequently, the mobile node (MN) minimizes the power consumption.

The rest of this paper is organized as follows. In Section 2, we briefly review related work. In Section 3 an overview about SAW method and the IEEE 802.21 standard is presented. We propose the IEEE 802.21 handover protocol extensions and the handover decision algorithm (HDA) in Section 4.2. The Section 5 shows simulation results, comparison, and analysis. Finally, we present conclusion in Section 6.

2 Related Work

In the future networks, there will be several alternatives of wireless access technologies, therefore IEEE 802.21 will be increasingly important to integrate these

networks in a seamless way for the users. The handover decision mechanisms and algorithms are a key part in this process. There are several handover decision algorithms in the literature.

Tawill et all. [8] propose a handover decision algorithm using the SAW method in a distributed manner. They use, as evaluation metrics to making handover decision, bandwidth, dropping probability, and cost.

Lee et all. [9] propose a vertical handover decision algorithm based on a utility function to satisfy the QoS requirements. This utility function considers signal to interference plus noise ratio (SINR), bandwidth, traffic load and user's mobility, its goal is to maximize the network throughput.

Kim and Jang [10] propose a vertical handover decision algorithm using IEEE 802.21. The mobile node periodically measures the Receive Signal Strength (RSS). The mobile node computes an RSS's mean to figure out which is the best time to execute the handover. The solution is based on same idea of the traditional handover that makes the decision based on the signal power.

The work closer to the proposed solution is presented by Tawill et all.[8]. Despite the authors use the SAW method to ranking the networks and computing the network score in a distributed way, the solution does not take into account the networks' current traffic load (bandwidth) at the time of decision. Using just SAW method to decide which is the better network, a network that fits the user's needs could remain more loaded than others, reducing their quality of service. Futhermore, the solution does not use the IEEE 802.21 services to obtain the networks' information, thus it is more complicated than the proposed solution. In the works developed by Yang et all. [11] and Lee et all. [9], the whole decision calculation is performed by the MN, decreasing their battery level quickly.

3 Background

This section presents some basic concepts about IEEE 802.21 standard and SAW method.

3.1 IEEE 802.21

Due to the heterogeneity of protocols and technologies involved, vertical handover is much more difficult to implement than horizontal handover. The IEEE 802.21 standard comes up to facilitate the vertical handover. This proposal provides a framework that enables the optimization of handover among heterogeneous networks. The purpose is to improve the user experience of an MN by faciliting handover among networks including both 802 and non 802 networks, and wired and wireless networks [3]. As show in Fig. 1. the IEEE 802.21 defines a service layer between the network and link networks. This layer is called Media Independent Handover Function (MIHF) and provides three types of services: the media-independent event service (MIES), the media-independent command service (MICS) and the media-independent information service (MIIS).

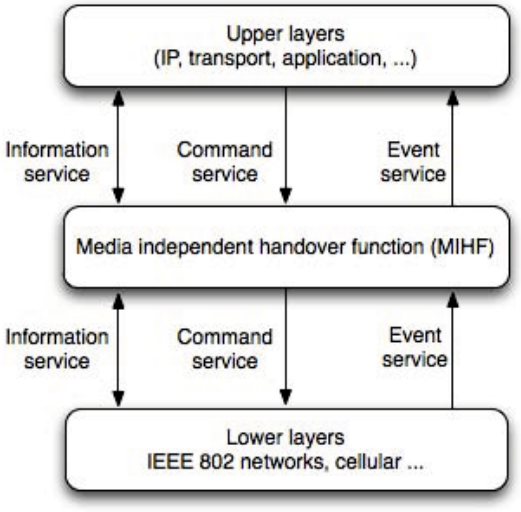


Fig. 1. MIH services

The main role of MIES is to detect events that occur in local or remote interfaces and report them to upper layers. Examples of events that may be reported to upper layers are degradation of the signal power, unavailability of link, among others.

The MICS is responsible for making available commands to uppers layers. These commands are responsible for controlling the lower layers, more specifically to control the handover procedures.

The service information (MIIS) provides information about network's service and available networks. This information is used by the MN to decide which network is a good choice to connect to. This information may include, for instance, GPS coordination, channels information, etc.

The standard proposes an information server (IS) to cope with all related network information. This information supports MN choosing the best available network, therefore facilitating network handovers. IEEE does not define where the IS is located. For instance, it can be located within the MN domain. Similarly, the database structure is not defined. The IS has a key role in the development of our proposed solution. The proposed algorithm uses information provided by the IS.

3.2 SAW - Simple Additive Weight

Decision makers often deal with problems that involve multiple, usually, conflicting criteria. In our problem, the user has several options to join a network and needs to decide what is the best alternative based on its preferences.

The SAW method, also called weighted sum method, is the simplest and still the widest used MADM method. In order to produce a value, a decision

Table 1. Decision Table

Network	Bandwidth (Mbps)	Weight	Cost	Weight
Wi-Fi	54	0.5	$0.70	0.5
WiMAX	70	0.4	$1.00	0.6
Cellular	42	0.1	$5.00	0.9

table must be mounted based on each alternative and criterion available. Each alternative is a line and each criterion is a column. First, all elements of the decision table need to be normalized, then, the SAW method can be used for any type and any number of criteria. The Table 1 shows an example decision table. In this table each alternative is a network with its respective criteria and weights.

A weight is given for each criterion and the sum of all weights must be 1 [6]. The importance of each criterion is determined by a weight. This combination of criterion plus weight composes the score. The score of the SAW method is given by:

$$Score_i = \sum_{j=1}^{M} W_j(m_{ij})normal \tag{1}$$

where $Score_i$ is the score of the alternative i; W_j is the weight of criterion j; M is the number of available criteria; m_{ij} is the criterion value of the alternative i for the criterion j; *normal* are the values normalized.

4 Handover Protocol and Proposed Decision Algorithm

Some extensions were made in the messages of IEEE 802.21 handover protocol in order to include metrics used by the proposed algorithm. In the next subsections the messages extensions of the IEEE 802.21 handover protocol are present as well as the proposed handover decision algorithm.

4.1 IEEE 802.21 Handover Protocol

The Fig. 2. presents the steps that the MN takes according to IEEE 802.21 to choose the best network to connect to. Proposed extensions in the standard are commented below.

First, the MN sends a message to the IS to figure out which networks are available in its coverage area (1). We have extended this original message including a handover field, thus the MN may send into the handover field "fromHandoff" or "newCon" string. If the MN is trying a handover, it includes the "fromHandoff" string into the field, otherwise, "newCon" string is included meaning that the MN wants to perform a new connection.

Second, the IS verifies which networks are available to the MN and sends a message to all of them (2). We've added an information in this step, the IS

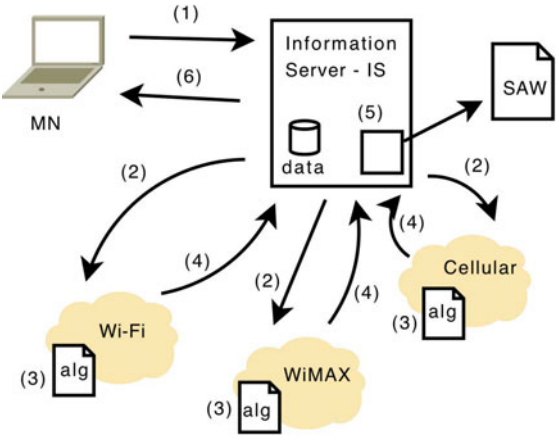


Fig. 2. Operation of the solution

informs to the networks if the MN has one network option in its coverage area through the boolean field isUnique. If there is just one option, the isUnique field is true.

Then, each network runs the Algorithm I (3) (see next section), and verifies if it can accept the MN or not. The answer message is sent to the IS (4). We have extended this message adding the actual_load_bw field, thus each network sends to IS its current traffic load inside the answer message.

Next, the IS mounts a decision table, as shown in Table 1 with all network's answers and sends this table to the SAW module (5). We added the SAW module in the IS to calculate the SAW score. The SAW module computes the final score of all networks. Each network score is divided by its current traffic load. After that, the SAW module sends the scores to the IS. Then, the IS sends a message to the MN (6) with all available networks with their respective scores. Finally, the MN chooses the network with the higher score to connect to.

The Fig. 3. shows a use case diagram of the IEEE 802.21 protocol with proposed extensions. All messages that are exchanged between the MN and the IS are showed. First, the user sets its preferences (bandwidth and cost) with its respective weights (1). Then, the application (UP layers) sends an MIH_Get_Information_request message to the MIH User (2). In this message the fields bw, weightB, cost, weightC were added to inform to the IS the bandwidth (bw), bandwidth's weight (weightB), cost (cost) and cost's weight (weightC). These fields represent the user's requirements or user's criteria used by the IS to choose the best network. In this message the handover field is filled out as explained before. After that, the MIH User uses the MIHF to send a message to IS (3). Then, the IS executes the procedures showed before (Fig. 2.) and sends an MIH_N2N_HO_Query_Resources_request message to all candidate networks (4). In this message we add the field isUnique. Through this field, the IS informs if the MN has one or more available network(s) to connect to. Then, an

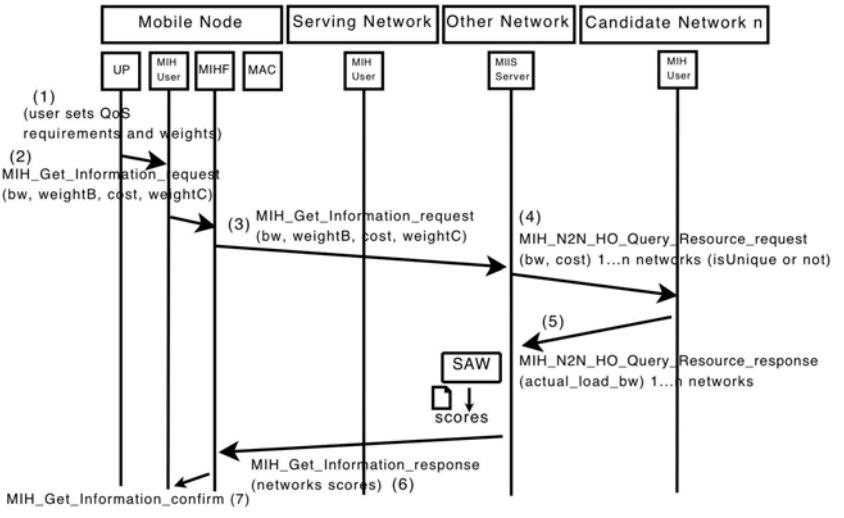


Fig. 3. Use case diagram of the IEEE 802.21 with proposed extensions

MIH_N2N_HO_Query_Resources_response message comes into the IS (5). As we explain before, we added the actual_load_bw field that informs the current load of each network. Now the IS executes the showed procedures (Fig 2. (5)). Finally the IS sends an MIH_Get_Information_response (6) to the MN.

4.2 Handover Decision Algorithm (HDA)

Algorithm I shows the proposed algorithm. This algorithm uses information provided by the IS and the MN to decide if the network may accept the MN or not. This algorithm is performed in each network as shown in Figure 2, step (3).

Initially, the algorithm verifies if the network can accept the MN by testing the user's criteria (bandwidth and cost) (line 1). This test compares the user's requirement with the network's resource, in other words, if the network may offer a bandwidth greater or equal than the user's request bandwidth and if the network's cost is lower or equal than the user's request cost. Next, the algorithm performs three tests (lines 2, 9 and 16) in order to distinguish the kind of connection of the MN, and then verifies two thresholds that will admit or deny the MN's connection. We defined two thresholds. The threshold1 (thr1) is used when the MN requests a new connection and there are more than one available network to connect to. The threshold2 (thr2) is used when the MN requests a new connection and there is just one available network to connect to as well as if the MN is coming from another network (handover=fromHandoff) and there are more than one available network to connect to. The use of two thresholds in the Algorithm I allows a carrier to assign priority to MNs. The algorithm can differentiate if the MN is trying to perform a handover (handover=fromHandoff) or

is starting a new connection (handover = newCon). These thresholds are useful to maintain the MN's connection in the case where the MN is already connected and it performs a handover.

Algorithm I - Handover Decision Algorithm

```
 1. if( (cost >= networkCost) or (bw <= networkBw) ) {
 2.     if (handover=newCon) and (not isUnique) {
 3.         if (((actual_load_bw + bw) <= thr1) and
 4.             ((actual_load_bw + bw) <=
                   max_load_bandwidth) ) {
 5.                 actual_load_bw = actual_load_bw +
                         bandwidth;
 6.                             connection accepted; }
 7.          }else
 8.              connection refused; }
 9.     else if((handover=newCon)and(isUnique)) or
            ((handover=fromHandoff)and(not isUnique)){
10.             if (((actual_load_bw + bandwidth) <= thr2) and
11.                ((actual_load_bw + bandwidth) <=
                                    max_load_bandwidth) ) {
12.                   actual_load_bw = actual_load_bw +
                            bandwidth;
13.                              connection accepted;
14.             }else
15.                    connection refused; }
16.     else if (handover=fromHandoff) and (isUnique) {
17.          if (((actual_load_bw + bandwidth) <=
                                   max_load_bandwidth) {
18.              actual_load_bw = actual_load_bw +
                       bandwidth;
19.                         connection accepted; }
10.          else
21.              connection refused; }
22.     else
23.         connection refused;
24. } else
25.     connection refused;
```

The first test (line 2) verifies if the MN's connection is not unique (not isUnique) and also if the connection is new (handover=newCon), in this case, it verifies (lines 3 and 4) if the network's current traffic load (actual_load_bw) plus the requested bandwidth (bw) is less than the thr1 and less than the max bandwidth supported by the network (max_load_bw). If these conditions are satisfied, the network admits the MN (line 6), otherwise the MN is refused by the network (line 8). The second test occurs when the MN is starting a new connection

(handover=newCon) and this connection is the only one available (isUnique), or the MN is trying to perform a handover (handover=fromHandoff) and has more than one available connection (line 9). In this case, the thr2 is the limit to admit (line 13) or deny (line 15) the MNs. In the third and last test, the algorithm tests if the MN is trying to perform a handover (handover=fromHandoff) and if the connection is the only available (isUnique) (line 16). The algorithm verifies if the bandwidth requested plus the current traffic load (actual_load_bw) is less than the max bandwidth supported by the network (max_load_bw).

In the cases where the network accepts the MN's request from the IS, the bandwidth is reserved (lines 5, 12 and 18) and the request is stored with a unique id, and a timeout is started. In the case the MN does not perform a request to connect to the network later, the reserved resource is released when the timeout expires. The steps after the handover decision are out of the scope of this paper.

5 Performance Evaluation

The purpose of this study was to compare how the proposed solution distributes mobile nodes through available networks. The resource utilization of networks was measured as well as the blocked ratio. The proposed solution was compared with an ordinary solution, where MNs choose the network, based on the network priority [12]. The MN first tries to connect to the Wi-Fi network. If it can not connect, it tries a WiMAX network and finally it tries the cellular network. In the next sections we call the ordinary decision, Wi-Fi algorithm, and our entire proposed mechanism, HDA. We show that our mechanism improves average network utilization. In the next paragraph, we describe our simulator as well as the developed experiments.

5.1 Simulator Implementation

A discrete event simulator was developed in Java. First of all, a traffic generator creates an event list sorted by time. Each event has a corresponding Mobile Node. This MN has an id, user's preferences (bandwidth, cost), MN's coordinates, and the kind of the MN's connection (fromhandoff or newCon). The id is a sequencial number, the other informations are generated using a uniform distribution up to a limit value. To each event created by the traffic generator, a departure event is created to define the time that the mobile node will leave the system. This time is generated using an exponential distribution using the MEANDEPARTURETIME variable as mean.

A control plane is a module responsible to control all functionalities of the simulator. There are three main modules: the stat, the network and the decision module. The stat module defines all statistics of the simulator, and it is customizable, in other words, it is able to be changed in order to be suitable to different situations. The network module deals with the information about the networks stored in the system. This module stores all information about the

networks, for instance, the kind of network (e.g Wi-Fi, WiMAX), max supported bandwidth, the propagation model, the transmit power, the thresholds and the access point/base station (AP/BS) coordinates. The communication radius in MNs is calculated by the network module using a combination of information such transmit power and depends on the propagation model defined in the simulator's configuration. The Tworayground propagation model was implemented in this simulator. This model is the same used by ns simulator [13]. The decision module is responsible to run the decision algorithm, and emulates messages produced by the IS.

5.2 Simulation Results

The topology of the simulated scenario consists of three networks. A Wi-Fi, a WiMAX and a Cellular network. In the Wi-Fi network, the transmission power of the AP is 0.281838db. The max supported bandwidth is 25.6 Mbps. In the WiMAX network, the transmission power of the BS is 0.481838db. The max supported bandwidth is 65 Mbps. In the Cellular network, the transmission power of the BS is 4.982838db. The max supported bandwidth is 41 Mbps. The AP/BS location is in the same place and the MNs are uniformly distributed around it, as shown in Fig. 4. We assumed no interference between the MNs and the AP/BS. For this scenario, it is assumed the presence of an admission control mechanism, so that the MN just receives the requested bandwidth. We ran each experiment 14 times, and we calculated a 95 percent of confidence interval.

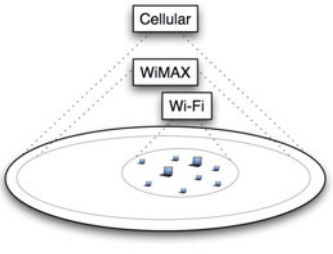


Fig. 4. Distribution of mobile nodes across the coverage area of the networks

The simulation results of the utilization ratio of resources are shown in Figs. 5-8. We can learn that the MNs were well distributed among the available networks by the HDA.

Figure. 5 shows the distribution of 26 MNs among the networks. Comparison between the HDA and Wi-Fi algorithm shows that our proposed mechanism distributes better the MNs among the networks. The network load of 26 MNs is low, smaller than the total capacity of the three networks. Note that almost all MNs connect in the Wi-Fi network using the Wi-Fi algorithm, meanwhile the other networks are empty. This occurs because Wi-Fi algorithm selects the Wi-Fi network first to connect to. The requested bandwidth of almost all MNs fits to Wi-Fi network capacity, leaving other networks empty.

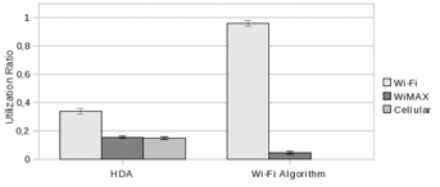


Fig. 5. Utilization Ratio of resources with 26 MNs

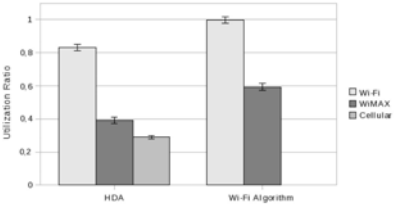


Fig. 6. Utilization Ratio of resources with 50 MNs

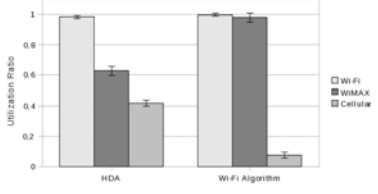


Fig. 7. Utilization Ratio of resources with 70 MNs

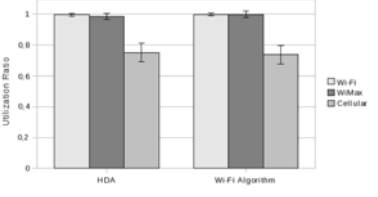


Fig. 8. Utilization Ratio of resources with 100 MNs

In Figures 6, 7, 8, the number of MN's request was increased to 50, 70 and 100, consequently, the network load was increased.

Up to 70 MNs, the HDA distributes the MNs among the networks better than the Wi-Fi algorithm, causing a load balancing among the networks. In the case where the number of MN is 100, both approaches distribute in the same way as shown in Fig. 8, but the blocked ratio of Wi-Fi algorithm is higher than HDA as shown in Table. 5, which causes a delay in the handover procedure.

The simulation results of the blocked ratio can be seen in Tables. 2-5. When the network load is low, the blocked ratio is low in both approaches. As the network load increases, the Wi-Fi algorithm blocks more MNs than the HDA, initially to enter into the Wi-Fi network, and after to enter into WiMAX and Cellular networks. When the MN's request is 100, the Wi-Fi algorithm blocks twice MNs than HDA for the Wi-Fi network, almost 40 percent more than HDA for the WiMAX network. There are no blocks in both approaches for the Cellular network.

Table 2. 26 MNs - Blocked Ratio

	Wi-Fi	WiMAX	Cellular
HDA	0	0	0
Wi-Fi	0.05	0	0

Table 3. 50 MNs - Blocked Ratio

	Wi-Fi	WiMAX	Cellular
HDA	0	0	0
Wi-Fi	0.43	0	0

Table 4. 70 MNs - Blocked Ratio

	Wi-Fi	WiMAX	Cellular
HDA	0.1	0	0
Wi-Fi	0.59	0.07	0

Table 5. 100 MNs - Blocked Ratio

	Wi-Fi	WiMAX	Cellular
HDA	0.36	0.06	0
Wi-Fi	0.71	0.43	0

6 Conclusion

This paper presents a handover decision mechanism using the IEEE 802.21 standard and the SAW method. The proposed mechanism leverages a load balancing among heterogeneous networks. Moreover, it considers the user's preferences as well as the traffic load of the networks, thus, the users can choose the networks according to their needs and the carriers have the benefit of a better resource utilization, supporting more users. For future work, we would extend our simulator to take into account moviment patterns of the mobile nodes.

References

1. Akyildiz, I., Jiang, X., Mohanty, S.: A survey of mobility management in next-generation all-ip-based wireless systems. IEEE Wireless Communications 11(4), 16–28 (2004)
2. Gustafsson, E., Jonsson, A., Res, E., Stockholm, S.: Always best connected. IEEE Wireless Communications 10(1), 49–55 (2003)
3. 802.21, I.S.: Ieee standard for local and metropolitan area networks - media independent handover (January 2009)
4. Griffith, D., Rouil, R., Golmie, N.: Performance Metrics for IEEE 802.21 Media Independent Handover (MIH) Signaling. Wireless Personal Communications 52(3), 537–567 (2010)
5. Eastwood, L., Migaldi, S., Xie, Q., Gupta, V.: Mobility using IEEE 802.21 in a heterogeneous IEEE 802.16/802.11-based, IMT-advanced (4g) network. IEEE Wireless Communications [IEEE Personal Communications] 15(2), 26–34 (2008)
6. Rao, R.: Decision making in the manufacturing environment: using graph theory and fuzzy multiple attribute decision making methods. Springer, Heidelberg (2007)
7. Zanakis, S., Solomon, A., Wishart, N., Dublish, S.: Multi-attribute decision making: A simulation comparison of select methods. European Journal of Operational Research 107(3), 507–529 (1998)
8. Tawil, R., Salazar, O., Pujolle, G.: Vertical Handoff Decision Scheme Using MADM for Wireless Networks. In: Proc. of IEEE WCNC 2008 (2008)
9. Lee, D., Han, Y., Hwang, J.: QoS-Based vertical handoff decision algorithm in heterogeneous systems. In: 2006 IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1–5 (2006)
10. Kim, J., Jang, J.: Low Latency Vertical Handover Using MIH L2-Trigger Algorithm in Mobile IP Networks. In: Stojmenovic, I., Thulasiram, R.K., Yang, L.T., Jia, W., Guo, M., de Mello, R.F. (eds.) ISPA 2007. LNCS, vol. 4742, pp. 707–718. Springer, Heidelberg (2007)

11. Yang, S., Wu, J., Huang, H.: A vertical Media-Independent Handover decision algorithm across Wi-Fi and WiMAX networks. In: 5th IFIP International Conference on Wireless and Optical Communications Networks, WOCN 2008, pp. 1–5 (2008)
12. Mola, G.: Interactions of vertical handoffs with 802.11 b wireless LANs: handoff policy. Masters theses, Department of Microelectronics and Information Technology, Royal Institute of Technology (KTH), Stockholm, Sweden (2004)
13. NS2: Ns2 network simulator, http://www.isi.edu/nsnam/ns/

Proactive Vertical Handover Optimizations in the 3GPP Evolved Packet Core

Marius Corici, Dragos Vingarzan, Thomas Magedanz, Cornel Pampu, and Qing Zhou

Fraunhofer FOKUS Institute and Huawei Technologies
Berlin, Germany
{marius-iulian.corici,dragos.vingarzan,
thomas.magedanz}@fokus.fraunhofer.de,
{cornel.pampu,zhou.qing}@huawei.com

Abstract. Mobility in a wireless heterogeneous scenario in which the mobile devices are able to connect to more than one access technology available in their vicinity requires a re-consideration of the access network reselection mechanisms as to ensure seamless handovers for real deployments. This paper describes and evaluates a new proactive vertical handover optimization which enables a fast reselection, independent and in addition to the classic proactive procedures. It uses as central concept the separation between the proactive context establishment and the actual handover triggered operations which may be at their turn active or proactive. This concept is exemplified on the 3GPP Evolved Packet Core and evaluated on a minimal prototype implementation.

Keywords: Access Network Discovery and Selection, Heterogeneous Networks, Fast Handovers, Evolved Packet Core.

1 Introduction

A multitude of wireless access networks are already available (e.g. WiFi, WiMAX, LTE, UMTS etc.) having different delays for attachment and offering various levels of resources to the mobile users. A growing number of deployments are already using more than one access technology in specific locations for an enhanced throughput of the wireless environment as to be able to offer extended services to the mobile users.

Due to the large number of wireless access networks available, the mobile devices are not able any more to select by themselves the most appropriate target access network. Because of this, multiple handover network architectures (e.g. 3GPP Evolved Packet Core, IEEE 802.21 Media Independent Handover etc.) introduce a network function which offers to the mobile device information on the momentary preference of the operator for the target access network and its parameters in order to be able to execute a fast handover which supports the service continuity.

These functions offer the discovery and selection information to the mobile devices either based on a request from the mobile device or based on some other internal network triggers. The information does not imply any connection to the actual execution of the handover which includes the operations of authentication and authorization, the link layer attachment and the network layer reachability establishment (i.e. IP address

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 109–122, 2011.

allocation and mobility association – Mobile IP procedures). Although, the information is already available in the network prior to the handover which is triggered by the mobile device due to loss or prediction of loss of connectivity to the source access network, it does not influence directly the handover procedures. One reason for this is that all the proactive procedures considered in the literature presume that a complete context is established in the target access network when they are executed which implies a high consumption of the resources of the target access network.

This paper considers a new type of context to be established on the target access network immediately when the network located selection functionality takes the decision to which access network the mobile device would handover to in case this is necessary in the near future. The context does not presume the reservation of the actual resources on the target access network, but a proactive preparation for the case the mobile device triggers a handover. Using this shallow context, the mobile device is able to connect faster to the target access network and by this reducing the delay of the handover procedures, which is especially benefic for handovers from source accesses in which the connectivity is lost fast (e.g. WiFi) to target accesses in which the context establishment has a large duration (e.g. UMTS, LTE etc.). The procedures of establishment, activation and release of the shallow context are further exemplified using the 3GPP Evolved Packet Core (EPC) and they can be applied for any other convergent network architecture. The Evolved Packet Core is standardized by 3GPP as an IP based multi-access core network which integrates both 3GPP e.g. GSM, UMTS, LTE etc. and non-3GPP e.g. CDMA, WiFi, WiMAX access technologies. A main concept of the EPC architecture is the mobility of the users based on coordination between link, network and application layer. The concept here presented integrates in the EPC as an extension of the access network selection and discovery functionality.

The remainder of this paper is organized as follows: Section 2 provides the background of the proposed method. Section 3 describes the general concept which is then exemplified on the 3GPP Evolved Packet Core architecture. Section 4 describes the testbed followed by the evaluation of the experimental results and in Section 5 conclusions are provided.

2 Background

The current state of the art considers separately the procedures of network based access network discovery and selection and the handover optimizations e.g. the 3GPP Evolved Packet Core ([1], [2]), IEEE Media Independent Handover ([6]) and IETF NetLMM ([8]) etc. Related to handover, for the network discovery and selection the following two mechanisms are defined, as depicted in Fig. 1:

1. Network discovery and selection decision independent of the triggering or of the handover operation – in the 3GPP EPC architecture
2. Network discovery and selection decision triggered by the trigger of the handover operation – in the IEEE MIH architecture.

In the first case, the network makes the network discovery and selection decision and transmits the information to the mobile device depending on other external parameters. When a handover is triggered by the mobile device, due to parameters

which can be measured only at the mobile device i.e. signal strength or internal policies, the handover operation is executed. This implies that no proactive procedure is executed on the target access network before the handover is triggered, either actively or proactively. Considering the active handover case in which the connection over the source access network is suddenly lost and that the procedures over the target access network have a large delay, the handover procedures have a large delay which impedes the seamless quality of the communication.

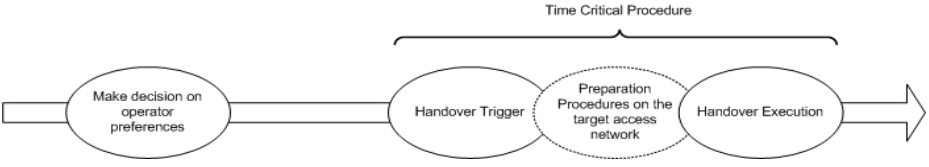


Fig. 1. State of the art in handover procedures

In certain scenarios, the network entity which makes the network selection decision is aware with a high probability of the access network to which the handover will be executed, but not of the moment of the handover trigger (active or proactive) which is mobile device dependent. In these cases, the network selection can be followed by an indication to the target access network in order to prepare for a future handover. Using this preparation, the time critical execution of the handover can be minimized.

In the second case, the network makes the network discovery and selection decision and transmits it to the mobile device only after the handover is triggered by the mobile device. In this case, all the procedures related to the handover are executed after the handover trigger, namely after the network receives the network discovery and selection request from the mobile device. Even though the handover trigger may be proactive (e.g. based on a rapid decrease of the signal strength), all the procedures related to the handover are executed after the event happens, which from the perspective of this article is similar to the previous case.

Thus, for both cases, only after the event of an imminent handover is available at the mobile device, the procedures of handover are executed which may contain a proactive or a preparation phase and an actual handover execution phase. The procedures for the handover have to be faster than the loss of connectivity to the source access network otherwise there is an interruption in the service due to the handover. Even though information is already available in the network, it is not used before the handover trigger due especially to the resources consumed by such a procedure in case the handover is not executed.

For example, in the IEEE 802.21 Media Independent Handover (MIH) architecture [7], [12] two types of operations are defined depending on the location of the Media Independent Handover Function (MIHF) which receives the handover event. For the first type, the handover is controlled by the MIHF of the mobile device while in the second it is controlled by the network MIHF. In the mobile device controlled scenario, its MIHF detects the event of handover and then requests to the network MIHF network discovery and selection information. All the handover related procedures are executed after the handover trigger which can include proactive procedures over the target access network.

In the network controlled handover, the network MIHF executes some proactive procedures and only afterwards transmits a handover command to the mobile device. In this case, the handover decision is not made by the mobile device and the handover trigger is received from the network. By this, the execution of the proactive procedures and the execution of the handover are tightly coupled. The handover has to be and is executed immediately after receiving the information from the network and it is not based on events which may be received only by the mobile device e.g. Link_Going_Down or Link_Down.

This second approach was extensively developed by the research community together and apart from the development of the 802.21 Media Independent Handover standard. For example, several approaches ([15], [17], [18]) consider that the required resources can be completely reserved over the target access network during the proactive phase of the handovers and prior to the actual handover execution. Other approaches ([16], [19]) consider also the integration in the proactive context of pre-authentication and pre-service signaling in the proactive context establishment.

However, due to the complete context which is created over the target access network, for the mobile device the resources are reserved twice: once in the source network and once in the target access network which are not strictly needed for the operation and have to be cleared in case the handover does not occur. For this reason, the proactive context establishment was bound to the proactive handover trigger as to have a higher degree of certainty of the handover. Also, because the context is established after the proactive handover trigger, due to the time constraint of the handover procedure, it may not be completely established for handovers between specific access technologies such as from technologies in which the signal strength is rapidly decreasing (e.g. WiFi) to access technologies in which the context establishment has a long duration (e.g. 3GPP access technologies). For this reason, the establishment of a shallower context prior to the handover triggers which may or may not be completed through a proactive procedure as the previous ones presented from the standards and literature is enabling a reduced delay of the overall procedure.

As an example architecture to validate the concepts here presented the 3GPP Evolved Packet Core (EPC) ([1], [2], [13]) was chosen due to its goal of providing network convergence between different types of accesses and because it may be considered the next-generation transport level for signaled services, by offering to the mobile terminals transparent support for mobility between different heterogeneous access networks.

As depicted in Fig. 2, EPC architecture includes a number of gateways which are transparently unifying the parameters of the different access technologies like LTE, UMTS, WiMAX, CDMA2000, WiFi etc. The Serving Gateway (S-GW) and the evolved Packet Data Gateway (ePDG) are responsible for the 3GPP and respectively the untrusted non-3GPP accesses while other technology specific gateways are used for trusted non-3GPP accesses e.g. for WiMAX or CDMA2000. The Packet Data Network Gateway (PDN GW) is the central anchor for all the data traffic. The gateways ensure the reservation of resources required by the mobile device, named User Endpoint (UE), through both the wireless environment and the EPC. Also they support the mobility protocol at network layer i.e. Mobile IP and variants.

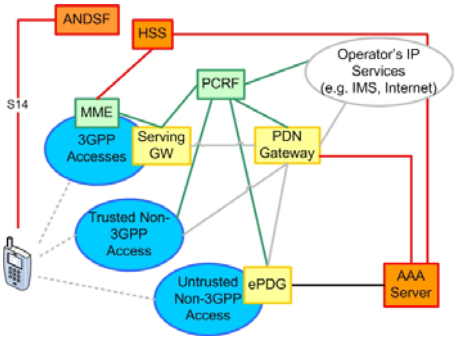


Fig. 2. 3GPP Evolved Packet Core

The Policy and Charging Rules Function (PCRF) ([2], [4], [5]) maintains the control over the resource reservations according to the user profile and to the communication with the service providers. It makes policy based decisions on the QoS levels that can be momentarily reserved for the UEs and enforces them on the gateways.

For the authentication and authorization of the UE the EPC uses a Home Subscriber Service (HSS) for the core network and an AAA Server for the Trusted and Un-trusted Non-3GPP accesses. For inter-3GPP technologies access control and mobility a Mobility Management Entity (MME) is used. It maintains the user context and offers services like reachability, resource management and handover signaling between the prior considered access networks.

The Access Network Discovery and Selection Function (ANDSF) ([3], [14]) transmits information available in the network to the UE related to the access networks to which it may connect to. In the current 3GPP standards only the interface to the mobile device (S14) is defined, thus being independent of the rest of the procedures of the EPC, therefore making the information transmitted to the UE independent of any handover execution procedures. On specific conditions, the mobile device alone initiates the handover procedures, without giving to the ANDSF the possibility of transmitting indications on immediate reselection.

Even though the ANDSF is aware that one or more of the access networks are highly probable to be selected by the mobile device in the near future and that a handover procedure is to be executed, no indication is transmitted to any other network entity in order to prepare for this future handover. When the handover event happens at the mobile device, it executes the handover procedures without having any proactive procedure executed in the network. Also in this case, all the re-selection procedures are executed after the handover trigger is received. If the attachment to the target access network takes a long period of time (e.g. in 3GPP accesses – GPRS, UMTS, LTE) and the connectivity to the source access network is fast lost (e.g. WiFi accesses), then a reduction of the duration of the procedures executed after the handover trigger has to be considered.

Therefore neither the 3GPP nor the MIH consider that any procedures are executed on the target access network prior to the active or proactive handover trigger, leaving that all the handover procedures are executed after it.

3 Concept

This article proposes a novel method for the execution of a proactive phase prior to the handover trigger, by executing a reduced context establishment on the target access network. The context is denominated from now as shallow context. It also considers that a new interface between the entity in the access network which makes the decision from the network side to which access network the mobile device may attach to in case a handover event occurs and a correspondent entity in the target access network which establishes the shallow context.

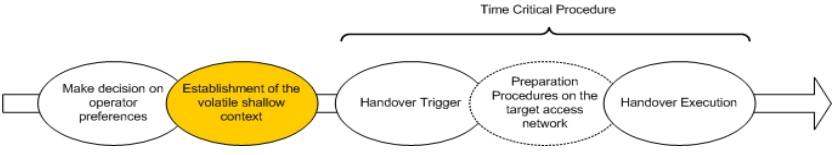


Fig. 3. Concept

The indication for the establishment of the shallow context in the target access network is transmitted when the information on which access network may be reselected is transmitted to the mobile device. The solution considers that the context is created due to the new information transmitted to the mobile device; independent of the moment when the handover event is received by the mobile device as depicted in Fig. 3.

The operations selected for the establishment of the shallow context are part of the steps which are executed during the handover procedure to the target access network. They don't have to be executed anymore after the handover event is received by the mobile device. The conceptual architecture is made of the following functional elements:

Network Access Selection Decision Function (N-ASDF) – it makes the decisions on the access network reselection from the network side. This function maps to the ANDSF in EPC.
Mobile Node Access Selection Decision Function (MN-ASDF) - it makes the decisions on the access network selection due to the triggers that are available only at the mobile device.
Network Access Selection Enforcement Function (N-ASEF) – it is attached to the target access network and executes the establishment if the context of the mobile device. It is a generic representation of the target access network entities.
Mobile Node Access Selection Enforcement Function (MN-ASEF) – it executes the selection procedures according to the decision taken by the MN-ASDF and it is equivalent to the attachment functionality of the UE in EPC.

Using this generic architecture, this article proposes the following access network reselection method:

Step 1: N-ASDF makes the decision on network selection policies that the MN-ASDF has to execute in case a handover event happens

Step 2: N-ASDF sends an indication with the decision to the MN-ASDF.
Step 3: N-ASDF makes the decision on which is the most probable access network that the N-ASDF would select in case a handover event happens
Step 4: N-ASDF sends an indication to the N-ASEF to establish a shallow context on the access network decided in Step 3.
Step 5: N-ASEF establishes the shallow context which will be further detailed for the EPC in the following sections.
Step 6: Independent of the previous steps, the MN-ASDF receives a handover event and makes the decision which access network to be selected.
Step 7: The decision is enforced to the MN-ASEF which selects the access network on which the N-ASEF has established the shallow context. The context is activated by this enforcement.

The shallow context should include information on the identity of the mobile node as to be able to identify during handover for which node the shallow context was created. The information is available in the N-ASDF because it is also the identification for the communication with the MN-ASDF.

It should also contain information necessary for the authentication of the mobile device in the target access network as to reduce the authentication procedures during handover – e.g. authentication vectors for the specific mobile node.

The context may include information on the active data flows of the mobile device which enables a faster resource reservation for the mobile device after the handover event. This information may be received from the MN-ASDF and/or may be augmented by information available already in the network.

It also should include information on the link layer, network layer and mobility protocol which otherwise would be available only after the handover event. For example, in the case Proxy Mobile IPv6 is deployed, it contains the home prefixes of the mobile device in the Mobility Access Gateway (MAG) and a secondary tunnel between the MAG and the Local Mobility Anchor (LMA).

In the EPC, the role of the N-ASDF is taken by the ANDSF as being the function which offers information to the mobile device on the accesses; while the role of the N-ASEF is taken by the gateways for the non-3GPP accesses i.e. ePDG and by the management entities for the 3GPP accesses i.e. S-GW or MME.

In order to establish the volatile context, the N-NREF can initiate: pre-authentication procedures based on the identity of the MN received from the N-ASDF (in the EPC the communication with the HSS), discovery of a path through the target access network (in EPC the discovery of the PDN GW) and its preparation (e.g. for Proxy Mobile IPv6 the creation of the tunnel between the gateway of the specific access and the PDN-GW anchor). It also may initiate policy and charging rules which bring the information on the QoS that has to be enforced in case the handover will occur (e.g. Gateway Control Session and PCEF Initiated IP-CAN Session Modification procedures in case of the EPC).

The context should not be initialized during the establishment procedure, but the information should be maintained on the network entities (e.g. for the EPC, the PCC and QoS rules and event triggers are available in the network entities in case a handover event happens, but they are not enforced).

The shallow context establishment does not presume the reservation of the actual resources on the data path. The resources which may be required in case of a

handover can be further considered as available resources in the specific access network. This may lead to a failure of resource reservation in case of a handover, due to the allocation of these resources to other mobile devices. For this reason, the shallow context here introduced should be integrated with the proactive handover procedures as extensively researched in the literature.

During the context establishment, the network part of the wireless context may be retrieved by the gateway of the specific access networks e.g. for UTRAN the PDP context information.

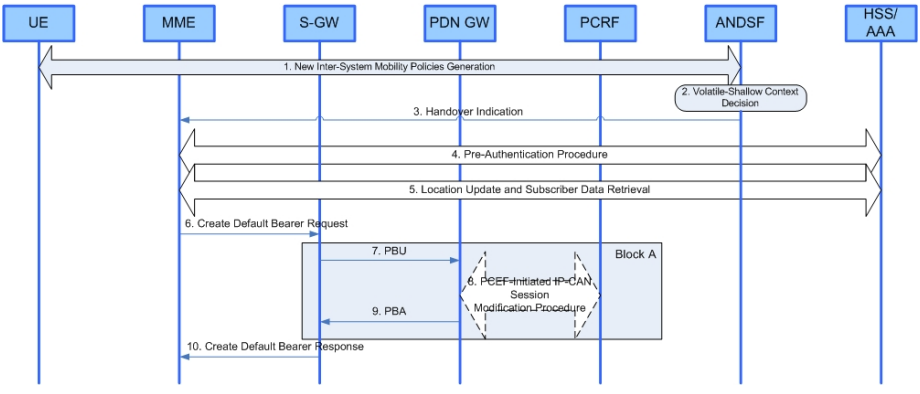


Fig. 4. E-UTRAN Shallow Context Establishment

Having the shallow context established when a handover occurs it is activated by a reduced number of procedures in case the resources required are still available. Otherwise, the handover procedures are similar to the state of the art procedures in which the resources required by the mobile device are not available in the target access network, scenario which was extensively presented in the literature. For the authentication, the communication is reduced to the one between the mobile node and the first network entity (e.g. the gateway or the manager of the access network). No inside network communication is necessary anymore. Also no communication for the resource reservation is necessary anymore, as the resources should be reserved by the activation of the shallow context which should be executed by each entity when the first upload packet is received or by the minimal authentication procedure considered in this paragraph.

For each type of access networks of EPC a different procedure is to be executed for the establishment and the activation of the shallow context. In the example of this article, E-UTRAN access network is chosen as the target access network and Proxy Mobile IPv6 as the mobility protocol as depicted in Fig 4.

Step 1: New inter-system mobility policies are generated for the UE by the ANDSF. The new policies are transmitted to the UE containing the accesses to which the UE may select for ensuring service continuity.
Step 2: ANDSF makes a decision for which access network from the information transmitted to the UE, a shallow context may be beneficial, based on the access network to which the UE is already connected to, the target access network type and

the probability of a handover. For the E-UTRAN, ANDSF selects the MME which would serve the UE in case of a handover.
Step 3: The ANDSF transmits a handover indication to the MME. It is considered that by knowing the access network to which the UE may handover to, the ANDSF is aware also of the MME which will serve the UE in that access network.
Step 4: The MME contacts the HSS and authenticates the UE. The authentication procedure is based on executed only on the network side and includes the retrieval of the authentication vectors from the HSS based on the identifier received from the ANDSF in the previous step as presented in 3GPP TS 33.401 section 6.2.1.
Step 5: The MME retrieves the subscriber data from the HSS, including the PDN GW identity and the information on the PDNs the UE is connected to over the source access network.
Step 6: The MME selects a Serving GW for the case no APN was provided by the UE as described in [1]. The MME sends a Create Default Bearer Request to the selected S-GW which includes the information received from the ANDSF.
Step 7: The S-GW sends a modified Proxy Binding Update (PBU) to the PDN GW which contains for the Handover Indication (HI) a new value which indicates that the tunnel has only to be created and that the data traffic should not be routed through it unless uplink data traffic is received or the source tunnel is not available anymore.
Step 8: Since the PDN GW is aware that a shallow context is created, it executes the mobility and resource reservation procedures as follows. When the PBU with the Handover Indication is received, a new tunnel is created, without routing the data traffic to it. An indication to the PCRF is transmitted which contains the flag that a shallow context is established similar to the resource modification of the standard procedures. The PCRF correlates the request with the identity of the UE and maintains in the subscriber structure that a shallow context is established. The PCRF makes the authorization and the policy decision considering that a handover may occur in the near future. A response is sent to the PDN GW indicating which resources have to be reserved and which event triggers are to be activated case the connectivity over the source access network is terminated or that the connectivity over the target access network is established. The resource reservation rules and event triggers received from the PCRF are maintained by the PDN GW until the shallow context is activated. No operations are executed.
Step 9: The PDN GW responds with a Proxy Binding Acknowledgement to the S-GW which contains the IP address or prefixes to be assigned to the UE in case of a handover.
Step 10: The S-GW responds with a Create Default Bearer Response to the MME. This message allows the MME to determine that the shallow context was established.

When the UE attaches to the E-UTRAN, the following operations are executed:

Step 1-2: UE selects E-UTRAN as a result of a handover event and of the policies received from the ANDSF.
Step 3: The authentication procedure is executed. As the MME is aware of the identity of the UE, the procedure is executed only between the UE and the MME and may be entirely skipped.
Step 4-9: The wireless link is configured. The UE may receive its IP address at this step or through the router solicitation/router advertisement mechanism.

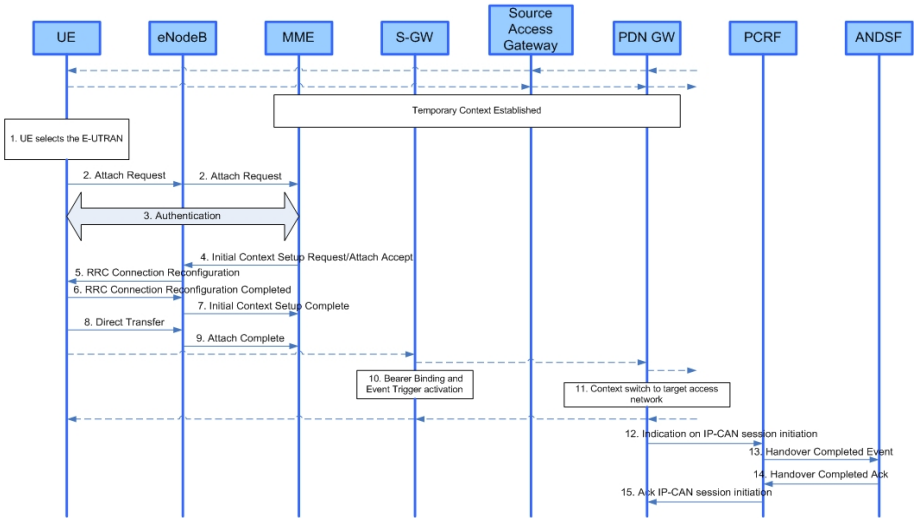


Fig. 5. E-UTRAN Shallow Context Activation

Step 10: The first upload packet triggers the activation of the shallow context in the S-GW. The activation includes the enforcement of the QoS rules and the bearer activation. It also includes the activation of the PMIPv6 tunnel from the S-GW side. No procedure for PMIPv6 or for required resources is executed as in the state of the art.
Step 11: The first upload packet triggers the activation also in the PDN GW for the resources, the PMIPv6 tunnel and the event triggers. No procedure is to be executed between the network entities as the information is already available from the shallow context establishment. The data traffic is exchanged bi-directionally on the target access network
Step 12-15: An indication is sent by the PDN GW to the PCRF and from the PCRF to the ANDSF in order to acknowledge the activation of the context.

The attachment procedure and the shallow context establishment may happen synchronously. As the control of the E-UTRAN is given to the MME, it may act as a synchronization point. For example, in the authentication procedures, if the pre-authentication procedure was already initiated when the request is received from the UE, the MME waits for the response from the HSS and responds to the UE. The delay is smaller than the case in which the MME issues a new request. Similar inferences are valid also for the other operations of the handover.

4 Testbed, Results and Evaluation

To test the applicability of the proposed procedures in a real environment an IPv6 based testbed was built as depicted in Fig. 6. A UE able to connect to a WiFi access network and to a public operator UMTS network was chosen. As the public operator

offers services only using IPv4, all packets being sent by the UE to that network are encapsulated in IPv4 packets and are sent through the network architecture of the public operator which is out of control of our testing system.

Two gateways were deployed: a S-GW for the UMTS which acts as the MME and S-GW together from the previously presented procedure and an ePDG for the WiFi. The gateways have PMIPv6 Mobility Access Gateway (MAG) functionality and are connected to DHCPv6 serves which offer the IPv6 address to the mobile device.

A PDN+PCRF server takes the role of the PMIPv6 Local Mobility Anchor (LMA).

An ANDSF was also deployed able to make handover decisions which are then transmitted to the mobile device using a minimal OMA DM PUSH mode mechanism and to the S-GW in the form of requests for establishing a shallow context.

The public 3G network presents a variability of the delay which affects the measurements in the mobile device, in the same way the need to tunnel IPv6 packets also adds a small delay in both the UE and the S-GW, but in such a realistic scenario the delay of the handover procedures is made more visible as this procedure typically does not only involve the internal network procedures.

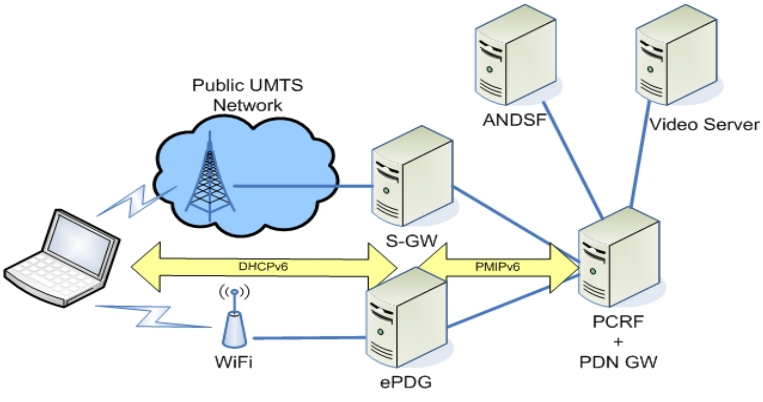


Fig. 6. Testbed Architecture

The implementation of the communication between the network entities is based on the FOKUS Diameter Stack for Linux while the minimal PMIPv6 implementation is based on Ubuntu Linux-based Operating System primitives. For the signaling part user-space raw sockets were used in both the PMIPv6 LMA and MAG. For the data tunneling, the standard IPv6 over IPv6 tunnels offered by the Linux kernel were considered.

The scenario measured and evaluated by this article considers that the UE is already connected to the WiFi network. An indication is transmitted by the ANDSF to the UE that in case an application is started to use the 3G network instead. At the same time a similar indication is transmitted to the S-GW to establish a shallow context.

The testbed considers that a shallow context contains the information on the IP address of the mobile node and the establishment of an inactive PMIPv6 tunnel between the S-GW and the PDN GW; limiting the efficiency of the shallow context to the network layer procedures (no resource reservation or authentication was considered).

When the user initiates the application, the UE executes a handover to the 3G network and only afterwards executes the service establishment.

This scenario based on an application initiation trigger is similar to the loss of signal in WiFi scenario, but it enables the operator of the testbed to easily run multiple consecutive tests.

Based on this minimal scenario, at the following entities, different delays have been measured:

UE attachment delay – The duration of the attachment with and without a shallow context including the DHCPv6 and the PMIPv6 procedures.

S-GW delay – The duration of the attachment with or without shallow context measured at the S-GW which translates into the duration between the moment the attachment request is received from the UE and the moment the attachment completion response is transmitted, in the case of the testbed, the duration between the DHCPv6 solicit and response.

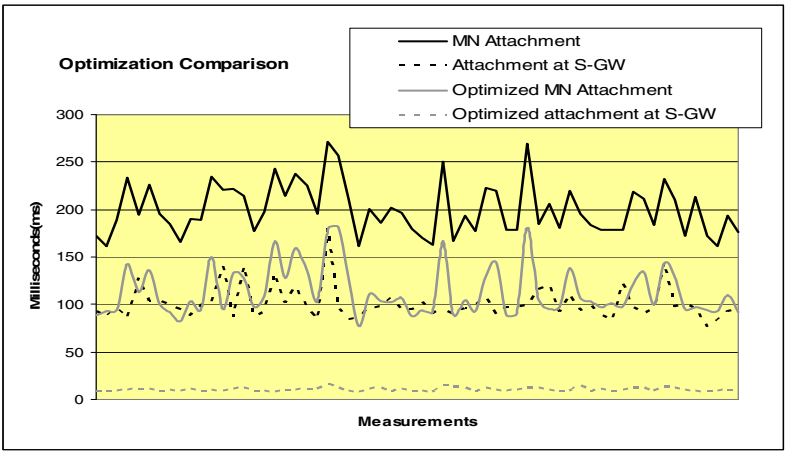


Fig. 7. Measured Values

As depicted in Fig. 7, the values obtained translated in percents bring an optimization of a median of 90.22% of the procedures executed in the S-GW for the connection of the MN and 47.62% for the overall connection procedures of the MN which also include the delivery over the wireless access system, with instable delay and not controlled by the testbed.

Although the number of measurements was limited and the procedures were executed for one MN and only for the network layer operations necessary for an access network reselection, the measured values prove the optimizations proposed by this article. A complete use of the shallow context would include also the pre-authentication operations and the subscriber data and the required resources retrieval in the gateway of the specific accesses or in the MME. It is expected that the same optimization is obtainable for these operations too, around 50% of the overall procedure time, which

makes the shallow context usage beneficial for the handover between access networks in which the signal is suddenly lost and the ones in which the attachment procedures have a long duration.

Thus, the opportunity of using a shallow context in the real deployments depends first on the physical characteristics of the access networks and on the determined level of probability that a handover to the specific access network will be executed.

5 Conclusions

This article presents a new concept of optimization for vertical handover procedures in which a set of proactive procedures are executed immediately after the network located access network selection makes the decision which access network has a high probability to be selected by the mobile device in the near future.

These procedures enable the creation of a shallow context on the possible target access networks. Due to the high degree of uncertainty that a handover will occur to the target access network, the shallow context does not presume any resource reservation on the data path as considered by previous proactive handover solutions. Instead, only the information on the resources which may be required in case of a handover is transmitted to the target access network entities. For this reason, the here presented procedures are introduced as an addition to the existing proactive procedures.

From another perspective, this concept enables the access networks to prepare for a future possible handover which means that the network is dynamically reactive not only to the attachment of the mobile device to the different access networks, but also to the logical decisions on which access network to select. Thus, the solution considers a two stage type of preparation: one in which the information on the requirements of the mobile devices is transmitted to the entities on the target access network and a reservation one when the handover trigger (active or proactive) is received.

The concept was exemplified in the 3GPP Evolved Packet Core and evaluated and validated on a minimal prototype implementation. From the delay values obtained we can conclude that the solution is feasible to be directly implemented in the real-life deployments as the delay decrease is beneficial compared to the functionality introduced in the network. However, the solution has to be further evaluated as impact on the network of the establishment of the shallow context in case the handover to the specific accesses does not occur and also how it integrates as an addition to the current proactive handover procedures.

References

1. 3GPP TS 23.401, General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Access Network (E-UTRAN), `http://www.3gpp.org` (access)
2. 3GPP TS 23.402, Architecture enhancements for non-3GPP accesses, version 8.4.1 (January 2009)
3. 3GPP TS 23.203, Policy and Charging Control Architecture, version 8.4.0, (December 2008)

4. 3GPP TS 24.302, Access to the Evolved Packet Core (EPC) via non-3GPP access networks, version 8.0.0 (December 2008)
5. 3GPP TS 29.212, Policy and Charging Control over the Gx reference point, version 8.2.0, (December 2008)
6. 3GPP TS 29.214, Policy and Charging Control over the Rx reference point, version 8.3.0 (December 2008)
7. IEEE P802.21/D05.00, Draft IEEE Standard for Local and Metropolitan Area Networks: Media Independent Handover Services (April 2007)
8. IETF Network-based Localized Mobility Management Workgroup (NETLMM), http://tools.ietf.org/wg/netlmm/
9. IETF RFC 5213, Proxy Mobile IPv6 (August 2008), http://www.ietf.org
10. Linux Advanced Routing & Traffic Control (March 2004), http://lartc.org
11. Open Mobile Alliance, Device Management Group, http://www.openmobilealliance.org
12. Piri, E., Pentikousis, K.: IEEE 802.21: Media Independent Handover Services. The Internet Protocol Journal 12(2) (June 2009)
13. Corici, M., Magedanz, T., Vingarzan, D.: The 3GPP Evolved Packet Core – the Mass Wireless Broadband All-IP Architecture. In: World Telecommunication Congress 2010, Vienna, Austria (September 2010)
14. Corici, M., Diez, A., Vingarzan, D., Magedanz, T., Pampu, C., Zhou, Q.: Enhanced Access Network Discovery and Selection in 3GPP Evolved Packet Core. In: IEEE LCN 3rd Workshop on User Mobility and Vehicular Networks (ON-MOVE), Zurich, Switzerland (October 2009)
15. Prior, R., Sargento, S., Gomes, D., Aguiar, L.R.: Heterogeneous Signaling Framework for End-to-End QoS Support in Next Generation Networks. In: 38th Annual Hawaii International Conference on System Sciences, Hawaii, USA (2005)
16. Dutta, A., Das, S., Famolari, D., Ohba, Y., Taniuchi, K., Fajardo, V., Lopez, R.M., Kodama, T., Schulzrinne, H.: Seamless proactive handover across heterogeneous access Networks. Wireless Personal Communication (2007)
17. Papadopoulou, E., McBurney, S., Taylor, N., Williams, M.H.: A Dynamic Approach to Dealing with User Preferences in a Pervasive System. In: International Symposium on Parallel and Distributed Processing with Applications, ISPA 2008, Sidney, Australia (2008)
18. Kim, I., Jung, Y.-C., Kim, Y.-T.: Low Latency Proactive Handover Scheme for Proxy MIPv6 with MIH. In: Ma, Y., Choi, D., Ata, S. (eds.) APNOMS 2008. LNCS, vol. 5297, pp. 344–353. Springer, Heidelberg (2008)
19. Boysen, E.S., Kjuus, H.E., Maseng, T.: Proactive Handover in Heterogeneous Networks Using SIPs. In: Seventh International Conference on Networking (icn 2008), pp. 719–724 (2008)

Key Distribution Mechanisms for IEEE 802.21-Assisted Wireless Heterogeneous Networks

F. Bernal-Hidalgo, R. Marin-Lopez, and A. F. Gómez-Skarmeta

Faculty of Computer Science, Dept. Information and Communications Engineering, University of Murcia
{fbernal,rafa,skarmeta}@um.es

Abstract. In recent years there has been a significant growth in the deployment of heterogeneous wireless technologies. Due to its diversity, new multi-interface terminals have appeared and pose new challenges to mobility management and security in wireless networks. In order to achieve a solution to these new challenges several standardisation groups are working to provide solutions that enable a seamless handoff in heterogeneous wireless networks by reducing the latency to obtain network access. In particular, the standardisation task group IEEE 802.21a is studying new media-independent services that allow a secure handoff process as well as mechanisms to reduce the latency during network access control after a mobile handoff. In this article, we analyse, three well-known key distribution mechanisms, in the context of IEEE 802.21a, for secure handover and how these mechanisms can help to reduce the network access time after a handoff in IEEE 802.21-assisted networks.

1 Introduction

In the last years the evolution of data networks and wireless devices have risen dramatically. Moreover, the proliferation of wireless access technologies implies that network subscribers can connect anywhere at any time using real time (e.g. voice calls or video streaming) applications that usually require high performance networks. For that reason, nowadays, several devices support different wireless technologies such as WiFi [1], third generation wireless connectivity (3G), or WiMAX. Due to this increasing diversity, operators must facilitate access to multiple wireless technologies through a single device.

Supporting handoff by avoiding loss of connectivity is the key enabling operation for seamless roaming and high-quality content delivery. Moreover, inter-technology handoff must be supported due to the growing network heterogeneity. An important factor that notably affects the provision of a seamless handoff between heterogeneous wireless technologies is the network access authentication and authorisation processes, by which operators control their subscribers, when they try to access the network service.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 123–134, 2011.

Different standardisation bodies are providing solutions to handle these kinds of problems in heterogeneous wireless networks. One general approach is the so-called SRHO. In this mechanism, a multi-interface terminal only transmits, at any given time, through a single radio interface during the handoff process. By means of SRHO, most of the processes (e.g. association, authentication, etc...) required to get network access are performed by using the single radio interface with the network where the mobile is intended to associate in the near future. For example, WiMAX forum [2] is working on a SRHO solution that handles both WiMAX-WiFi and WiFi-WiMAX inter-technology handoffs [3]. Also, 3GPP [4] is working on a Single Radio solution called SRVCC [5] which provides a service continuation between 3GPP and 3GPP2 [6]. It is also expected that IP-based 3GPP services will be provided through various access technologies, including existing broadband radio access standards like WiMAX.

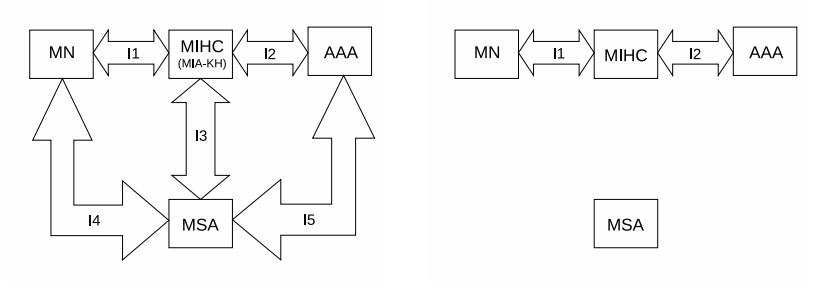
Additionally, IEEE 802.21 [7] standard defines media-access independent mechanisms to facilitate and enable optimisations to improve handoffs between heterogeneous wireless networks. So that, the main aim of this specification is to achieve seamless handoff between heterogeneous technologies. The standard defines all necessary elements required to exchange information, events and commands to facilitate handoff initiation (network discovery and selection process) and handoff preparation (network establishment before the movement). Specifically, there are several tasks groups which are defining new extensions to IEEE 802.21. For example, part of these extensions are being discussed in the IEEE 802.21c task group [8], where some MIH messages are being defined to transport link-layer frames to assist the SRHO mechanisms in WiMAX-WiFi and WiFi-WiMAX handoffs. Moreover, IEEE 802.21a task group [9] is defining mechanisms to further reduce the latency introduced by the authentication and authorisation processes usually required to get network access during the handoff process (working item #1); and security extensions to protect 802.21 messages (working item #2).

In particular, our contribution analyses and describes in detail the integration of three general and well-known key distribution mechanisms in the context of IEEE 802.21-assisted wireless networks in order to provide a secure handover and reduce the latency to get network access after the mobile handoff process. That is, the analysis which we describe in this article is focused in IEEE 802.21a working item #1.

This article is organised as follows. The section 2 describes the general architecture and specific details about the integration of the three general key distribution mechanisms in IEEE 802.21-assisted wireless networks. In particular, we will describe how to carry out a fast and secure heterogeneous handoff by using these key distribution mechanisms. Moreover, a deployment analysis is provided in section 3. Section 4 provides a performance analysis and shows how the key distribution mechanisms can further reduce the latency in solutions based on SRHO. Finally, section 5 provides some important conclusions and shows some future directions.

Table 1. Summary of used acronyms

Acronym	Definition	Acronym	Definition
SRHO	Single Radio HandOver	SRVCC	single Radio Voice Call Continuity
MIA-KH	Media Independent Authenticator and Key Holder	MSA-KH	Media Specific Authenticator and Key Holder
MN	Mobile Node	MIHC	Media Independent Handover Controller
MS-PMK	Media Specific Pairwise Master Key	MI-PMK	Media Independent Pairwise Master Key
TN-PMK	Tunnel Pairwise Master Key	PSK	Pre-Shared Key



(a) General Architecture

(b) Media Independent Authentication Architecture

Fig. 1. General Media Independent Architecture

2 Key Distribution for Media Independent Handoff

2.1 General Architecture

In the context of IEEE 802.21a, the proposal specified in [12] defines an architecture based on a MIA-KH entity (see Fig.1(a)). In this architecture, this entity controls and interacts with a set of MSA-KHs and facilitates the MN to perform a (proactive) authentication *before* moving to a *target* MSA-KH. Nevertheless, in order to complete the network access process, a key distribution mechanism is required to distribute key material to the target MSA-KH in order to establish a security association between the MN and the MSA-KH. If this key distribution process is carried out before the handoff, it will cause a reduction of (and in some cases, even eliminate) the authentication process. In general, it is initially assumed that a target MSA-KH will require an authentication based on the EAP [10] (unless some optimised key distribution is deployed), which has been recognised as a very flexible authentication protocol and used in multiple wireless technologies (e.g. WiFi or WiMAX) but is not as appropriate for authentication in mobile networks [11].

However, how key distribution is performed has not yet been deeply discussed in the context of the proposal [12]. Thus, in our contribution, we analyse and describe the integration of three well-known key distribution mechanisms but take into account the MIA-KH/MSA-KH architecture described above. We have considered these key distribution mechanisms as MIH services provided by the MIA-KH to the MN. According to this approach, we consider that the MN should

be authenticated and authorised to use these services. As such, we believe that MIA-KH can be such an entity since it allows a (proactive) media-independent authentication. However, we consider that in order to embrace the concept of MIH service, a more generic name is required, as not only media-independent authentication, but also key distribution services are provided by that entity. Therefore, we have renamed MIA-KH to MIHC and its functionality has been updated to provide both secure MIH signalling and help to reduce the network access time by providing different key distribution services: push, reactive pull and proactive pull key distribution.

Thus, the use of the MIHC entity has two main goals: one is to authenticate and authorise the use of the MIH services and the second is to assist the proper execution of them.

Table 2. Summary of External Interfaces for Media Independent Authentication and Key Distribution

Interface	Functionality
I1	It is used for performing the Media Independent Authentication, Push and Proactive Pull Key Dist. mechanisms. It is in charge of transporting MIH signalling, all required information for the key dist. method and the authentication protocol for media-independent authentication
I2	It is used to transport the authentication protocol to the MN's home domain in order to perform the authentication (e.g. AAA protocol)
I3	It enables, in Push Key Distribution, the MIHC to install a MS-PMK in the target MSA-KH. Moreover, in Proactive Pull Key Distribution, it transports the target technology level two frames to the MSA-KH
I4	It is used to communicate the MN with the MSA-KH
I5	It is used by the target MSA-KH to communicate with the AAA server

In order to achieve these goals, entities needs to communicate through several interfaces. Table 2 shows a summary of the interfaces used by each entity depending on the key distribution method used and the interfaces required for the media-independent authentication.

2.2 Media Independent Authentication Process

Before providing any MIH key distribution service, a media-independent authentication (Fig. 1(b)) by using some extensions to the MIH protocol (I1) is required between the MN and MIHC. The media independent authentication is composed by four phases:

1. *Negotiation phase.* Both the MN and the MIHC exchange unprotected MIH messages in order to agree on the type of key distribution mechanism to be used in that session and other related parameters.
2. *Media-Independent Authentication phase.* The MN and MIHC authenticate each other by using MIH signaling (I1) in order to get access to the key

distribution services. Moreover, MIHC may contact a backed authentication server (e.g. AAA server) to verify MN's credentials by using I2. In general, this media-independent authentication will be performed with the MIHC before the MN moves to a target MSA-KH under the control of the MIHC, in a so-called *proactive media-independent authentication*. In this case, we refer to the MIHC as *Candidate MIHC*. After performing the (proactive) authentication, key material will be shared between the MN and the MIHC, so that the rest of the MIH communication (I1) can be protected using this key material. This shared key, which is used as a root key for further key derivation, is a so-called *Media Independent Pairwise Master Key* (MI-PMK). At the end, the negotiated parameters in the negotiation phase are confirmed and an authentication session is established. For the purpose of this article, we assume that EAP is used as the authentication protocol since it provides a flexible way [10] to perform such authentication process.

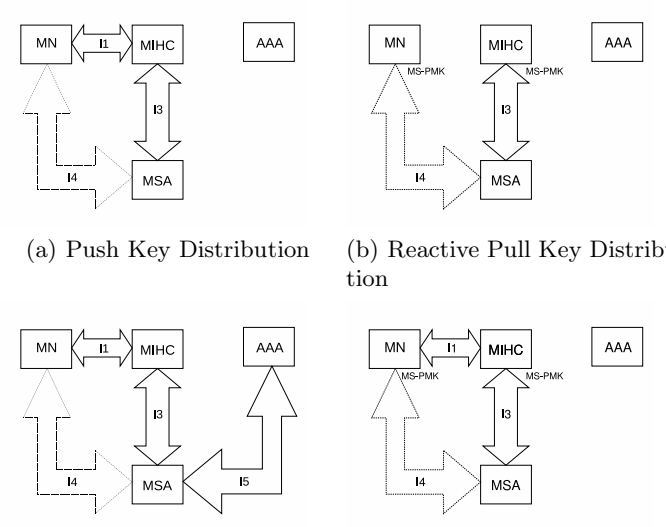
3. *Authenticated/Authorised phase*. In this phase, the MN is already authenticated and authorised to use the key distribution services provided by the MIHC.
4. *Finalisation phase*. When either the MN or the MIHC desire to finish the authentication session, they send protected MIH messages in order to release the MN's state related to the provided MIH services.

Once the MN is authenticated, we propose the use of three types of key distribution mechanisms (Push, Reactive Pull and Proactive Pull key distribution) which bring some interesting advantages (but also some associated disadvantages) to reduce network access latency as we describe in the following sections.

2.3 Push Key Distribution

In the *Push Key Distribution* mechanism (Fig. 2(a)), the MIHC pushes a key into the target MSA-KH on the a MN's request (*mobile-initiated process*) or some decision made by the MIHC itself (*network-initiated process*). The distribution process must be signaled by using protected MIH messages (I1) before the handoff to the target MSA-KH. Otherwise, an attacker could initiate the process. Then, the MN and MIHC will derive from the MI-PMK a specific MS-PMK for the target MSA-KH. To complete the process, the MIHC will push the MS-PMK into the target MSA-KH (under the control of the MIHC) by using interface I3.

Once the key has been installed, the MN can perform the handoff to the target MSA-KH, and establish the link-layer association and the security association to protect link-layer frames using I4. In general, the MN can pro-actively request pushing a new key in another MSA-KH under the control of the MIHC. Thus, using the key hierarchy derived from the MI-PMK, it is possible to access different MSA-KH without performing an EAP authentication each time the MN handoffs to a new MSA-KH under the same MIHC. So, the network access time after handoff can be reduced considerably.



(a) Push Key Distribution

(b) Reactive Pull Key Distribution

(c) Regular Proactive Pull Key Distribution

(d) Optimized Proactive Pull Key Distribution

Fig. 2. Key Distribution Mechanisms

2.4 Reactive Pull Key Distribution

The *Reactive Pull Key Distribution* mechanism (Fig. 2(b)) operates with the assumption that the MN and MIHC shares a symmetric key MS-PMK derived from the MI-PMK. In this sense, the MS-PMK should be considered as a *mid-term* credential shared between MN and MIHC for this type of key distribution. The MS-PMK will be used in a media-specific EAP re-authentication process based on an EAP method or mechanism (e.g. ERP [13]) that uses symmetric keys. This EAP re-authentication will happen *after* the MN moves to the target MSA-KH. In this EAP authentication, MIHC will act as EAP/AAA server and the target MSA-KH as an EAP authenticator.

Then, as a consequence of the media-specific EAP re-authentication which is performed by means of interface I4 and I3, an MSK is derived for the MSA-KH by the MIHC acting as AAA server. This MSK will be used to establish the security association between the MN and the target MSA-KH.

In order to achieve improvement with this key distribution mechanism, we propose the use of a temporal (NAI) [14] which must be provided for the MN. The temporal NAI can be provided during the media-independent authentication with the MIHC by means of interface I1 (Fig.1(b)). In general, this temporal NAI will have the format $user@mihc-realm$ where $mihc-realm$ can be the Fully Qualified Named (FQDN) of the MIHC. With this information, the target MSA-KH and/or the AAA infrastructure behind, can route the authentication and authorization (AAA) information to the MIHC during the media-specific

EAP re-authentication. The key aspect to reducing time with this type of key distribution is that the MIHC is assumed to be very near to the target MSA-KH so the latency between this entity and the MIHC is low.

2.5 Proactive Pull Key Distribution

The *Proactive Pull Key Distribution* mechanism allows media-specific EAP authentication without the need of the MN being directly connected to the wireless link of the target MSA-KH. To carry out this mechanism, it is necessary to transport link-layer authentication frames for the corresponding target MSA-KH wireless technology over a media-independent tunnel between the MN and MIHC (I1); and to convey those link-layer frames between the MIHC to the target MSA-KH by means of another tunnel (I3). In this way, after successful completion of the proactive media-specific authentication, a MSK will be pulled by the target MSA-KH as happens in a typical EAP authentication.

We have considered two main cases with this mechanism: The *regular Proactive Pull Key Distribution* (Fig.2(c)) and the *optimized Proactive Pull Key Distribution* (Fig 2(d)). In the former, the MN uses its NAI (e.g. $user@home-domain$) with the home domain (where the MN is subscribed to) in the proactive media-specific EAP authentication with the target MSA-KH. In this case, the authentication and authorisation process will be routed to the AAA server in the MN's home domain. In the latter, it is assumed that a MS-PMK is shared between the MN and MIHC as happens in Reactive Pull Key Distribution. Thus, it is also necessary to use a temporary NAI for EAP re-authentication purposes with the same format described in section 2.4, in order to forward the information to the corresponding MIHC (acting as a local AAA server).

Finally, it is worth noting that in order to implement the media-independent tunnel between MN and MIHC, we have considered two options: 1) carrying link-layer authentication frames over protected MIH signaling (this option is being considered on IEEE 802.21c as well); or 2) by the establishment of a dynamic secure IP tunnel (e.g. IKEv2) between the mobile node and the MIHC. This first option uses the interface I1 to transport link-layer authentication frames and the second option uses I1 to request the establishment of this secure IP tunnel. In the second option, the MN and MIHC will derive a pre-shared key (TN-PMK) obtained from key hierarchy rooted in the MI-PMK. With that key the secure IP tunnel will be established (e.g. IKEv2 with PSK authentication where the TN-PMK will be used as PSK).

From our point of view, we consider the second option as a better option because it separates MIH signaling from the tasks purely assigned to tunneling purposes.

3 Deployment Implications

In this section we describe the most important implications (advantages and disadvantages) that must be taken into account in the deployment of the three key distribution mechanisms that we have described in the previous sections.

For example, to use the *Push Key Distribution* mechanism the target MSA-KH must provide an implementation of interface (I3) in order to allow MIHC to push a key into the MSA-KH. Thus, to the best of our knowledge, no wireless technologies have standardised any interface that allows an external entity to push a key. To install a key, SNMP [15] could be used, but several changes must be carried out to do this. Conversely, the operation of the *Reactive Pull Key Distribution* does not need any change in existing wireless standards and, therefore, in the existing deployed target MSA-KHs. However, in the *Reactive Pull Key Distribution*, one possible deployment issue is that the MIHC also needs to act as AAA server in this process and there is no current entity nowadays which acts as an AAA client (for media independent authentication) and as a AAA server at the same time.

Additionally, the use of the (regular or optimized) *Proactive Pull Key Distribution* mechanism requires that the target MSA-KHs accepts wireless link-layer authentication frames over a wired link (the same issue is raised in single-radio handover case). Furthermore, a protocol to transport these frames from the MIHC to the target MSA-KH is required. Moreover, for the specific case of the optimized Proactive Pull Key Distribution, the MIHC must act again as AAA server in the proactive media-specific EAP re-authentication.

Also, obviously, MIHC functionality must be deployed on the existing networks. There are several alternatives. The first one is that the MIHC entity could be co-located with other existing entities (e.g. local AAA server). In this case, the software for these entities must be modified in order to support the new functionalities provided by the MIHC. In the second alternative, the MIHC could be a separate entity. In this case, the new entity must be deployed and connected with the rest of the network entities. Taking into account that other standardisation work considers including new entities and functionalities (e.g. WiMAX and WiFi Signal Forwarding Functions [3]) the first alternative seems the most promising option.

Finally, on the MN side, mobile terminals must be updated to support MIH protocol, manage the new key hierarchy and implement the different interfaces needed to support these key distribution methods.

4 Remarks on Performance

Nowadays, there is no existing implementation of these key distribution mechanisms in the context of IEEE 802.21-assisted wireless networks. So, this makes it difficult to obtain real experimental values in order to evaluate the key distribution mechanisms. For that reason, we have used real measurements taken from [16] [17] [18], where simulations and real scenarios have been used, to compute an approximate authentication delay, and to provide a rough analysis on how these key distribution mechanisms provide benefits during handoffs. Furthermore, using these mechanisms, other proposals (e.g. SRHO, IEEE 802.21c) can also improve their performance.

The following notation is used. T_{assoc} and T_{re_assoc} represents the time of performing a complete association and re-association *after* the attachment to the target MSA-KH, respectively. Note that, in general $T_{assoc} >\approx T_{re\_assoc}$ because performing a re-association process usually involves less messages. For example, if we consider IEEE 802.11, we assume that T_{assoc} includes 3 roundtrips [1]; whereas T_{re_assoc} only includes one roundtrip, as happens in IEEE 802.11r [19]. In [18] T_{assoc} implies a time of $\approx 20ms$ so, taking into account [18] and [19] we will roughly say a time of $6-7ms$ for T_{re_assoc}.

$T_{ms-auth}$ refers to the time consumed in carrying out a media-specific full EAP authentication between the MN and the MSA-KH that involves the MN's home domain. Based on [16] we can assume that $T_{ms-auth}$ is $\approx 600ms$ (in roaming case). $T_{ms-fast\_reauth}$ refers to the media-specific EAP re-authentication time without the need to contact with the MN's home domain but with the contact with the MIHC acting as AAA server. $T_{ms-fast\_reauth}$ could take $\approx 100ms$ according to [17]. Therefore, as [16] and [17] show, the assumption that $T_{ms-auth} >> T_{ms-fast\_reauth}$ is feasible. T_{sap} denotes the time of performing a secure association protocol (e.g. 4-way handshake in WiFi networks). Based on the measurements in [16] T_{sap} may take $\approx 10ms$. T_{push_key} refers to the time involved in contact with the target MSA-KH to push the corresponding key. Based on [16] we can assume a time of $\approx 10ms$. $T_{MIH_{push}}$ will refer to the time involved in the signaling required to indicate to the MIHC to install a key. Although there is no experimental data, we expect that this time will involve one exchange, as T_{push_key}. So, we could roughly assume that $T_{MIH_{push}}$ may also take $\approx 10ms$.

Taking into account these assumptions, when a MN moves to a target MSA-KH and no improvement is provided to reduce the network access delay, the latency to get network access through that MSA-KH can be computed in a very general way as:

$$T_{networkaccess_{total}} = T_{assoc} + T_{ms-auth} + T_{sap}$$

The table 3 shows the times only applicable to the different key distribution mechanisms described above. It shows two different time components: the time spent before the MN moves to the target MSA-KH (*Handoff Preparation Time*) and the time spent after the MN moves to the target MSA-KH (*Handoff Execution Time*). To obtain a rough estimation about the benefit of the performance of the three key distribution mechanisms adapted to the IEEE 802.21 context, we also assume that the MN has already performed the media-independent authentication with MIHC.

For this general analysis, we assume that $T_{MIH_{push}} + T_{push\_key} <\approx T_{ms-fast\_reauth}$. The reason for this assumption is that, $T_{MIH_{push}}$ involves one roundtrip between the MN and MIHC and T_{push_key} one roundtrip between the MIHC and the target MSA-KH. In $T_{ms-fast\_reauth}$ a similar number of roundtrips are required. For example, that happens when using ERP [13], since it is the most optimised and fast re-authentication solution defined in the IETF,

Table 3. Computed times for key distribution mechanism

Mechanism	Handoff Prep. Time	Handoff Exec. Time
Push	$T_{MIH_{push}} + T_{push\_key}$	$T_{assoc} + T_{sap}$
Reactive Pull		$T_{assoc} + T_{ms-fast\_reauth} + T_{sap}$
Proactive Pull	$T_{ms-auth}$	$T_{assoc} + T_{sap}$
Proactive Pull (optimized)	$T_{ms-fast\_reauth}$	$T_{assoc} + T_{sap}$

Table 4. Times when applying key distribution mechanisms to SRHO

Mechanism	Handover Prep. Time	Handover Exec. Time
SRHO	$T_{assoc} + T_{ms-auth} + T_{sap}$	T_{re_assoc}
SRHO+Push Key Dist.	$((T_{MIH_{push}} + T_{push\_key}) << T_{ms-auth}) + T_{assoc} + T_{sap}$	T_{re_assoc}
SRHO+ Proactive Pull (opt.)	$T_{assoc} + ((T_{ms-fast\_reauth}) << T_{ms-auth}) + T_{sap}$	T_{re_assoc}

where, in the best case, only one roundtrip is required. However ERP (or other authentication protocols) may require, in some specific cases, some additional message to complete a fast re-authentication process.

Taking into account these assumptions, we may observe that the Push Key Distribution will (potentially) reduce the network access time to only $T_{assoc} + T_{sap}$ when the MN attaches the target MSA-KH in each inter-MSA-KH handoff under the same MIHC. Although this equals the value of others key distribution mechanisms (proactive and optimized proactive pull key distribution), it has low latency at handoff preparation which is an advantage when the MN moves quickly to a new target MSA-KH under the same MIHC[1]. However, *Push Key Distribution* has some important deployment issues as discussed in section 3.

As we may also observe from table 3, the *Reactive Pull Key Distribution* will contribute with additional latency to network access control process after the MN attaches the target MSA-KH and will exhibit worse performance due to the MN fast re-authentication ($T_{ms-fast\_reauth}$) being carried out *after* MN moves to the target MSA-KH; this implies an increment in the latency with respect to other key distribution mechanisms but it actually represents a trade-off between easier deployment (see section 3) and fast network access.

It is also very important to note that some of the key distribution mechanisms could be used to minimise the handoff preparation time in SRHO process discussed in IEEE 802.21c. This is important, as commented before, if the MN moves quickly between MSA-KHs. In this manner, SRHO combined with either Push or Optimised Proactive Pull Key Distribution mechanisms can considerably reduce the SRHO handover preparation time and improve its benefits. Specifically, table 4 shows the combination of both methods and the time reduction achieved in SRHO. The first row in the table shows the general times involved in a traditional SRHO process. The rest of rows show how the use of the key distribution mechanisms can help the SRHO process. Basically, following the same assumptions as before, SRHO + Push Key Distribution can obtain better

[1] If a MN moves quickly the handoff preparation could not finish and not improvement would be achieved.

performance for the same reason as we already described. That is, *Push Key Distribution* takes a reduced time including preparation and execution handoff. So, this combination minimizes the problem of a MN moving to a target MSA-KH without completing the handoff preparation process.

5 Conclusions

In this work, we have analysed and described in detail how to integrate three general and well-known key distribution mechanisms in the context of IEEE 802.21-assisted heterogeneous wireless networks. The general architecture is based on an entity that we have called MIHC, which is an extension of the architecture described in [12]. The interfaces and the entities involved in the MIHC architecture have been defined and described, as well as, each key distribution mechanism. We have also discussed several deployment issues that must be taken into account in the real deployment of each key distribution mechanism. Finally, a general performance discussion about how the general and well-known key distribution mechanisms integrated in IEEE 802.21, can help to reduce the network access latency during handoff.

As future goals, we are now focused not only on the definition of the specific MIH messages for supporting the key distribution mechanisms introduced but also the implementation, based on ODTONE [20], of the mechanisms described in this manuscript to obtain experimental results. In continuation, with the specific values obtained during our experiments, we will perform simulations to observe the advantages and disadvantages of the described key distribution mechanisms, considering different types of use cases, wireless technologies and number of mobile nodes.

Acknowledgements

This work has been supported by the Funding Program for Research Groups of Excellence with code 04552/GERM/06 granted by the Seneca Foundation. Thanks also to the project CICYT TIN2008-06441-C02-02 which has also supported this work.

References

1. IEEE 802.11, `http://www.ieee802.org/11/`
2. WiMAX-3GPP Interworking WMF-T37-002-R010v3 (January 2008)
3. WiFi-WiMAX Interworking DRAFT, `https://mentor.ieee.org/802.21/dcn/10/21-10-0014-00-0000-wifi-wimax-iwk-spec.pdf`
4. 3GPP 3rd Generation Partnership Project (3GPP), `http://www.3gpp.org/`
5. 3GPP Single Radio Voice Call Continuity (SRVCC) TS 23.216
6. 3GPP 3rd Generation Partnership Project 2(3GPP2), `http://www.3gpp2.org/`
7. IEEE 802.21 Media Independent Handover Working Group, `http://ieee802.org/21/`

8. IEEE 802.21 Optimized Single Radio Handovers PAR and 5C 21-09-0146-05-0000-single-radio-handovers-par-and-5c.doc (November 2009)
9. IEEE Security SG Technical Report 21-08-0172-02-0sec-21-08-0012-02-0sec-mih-security-technical-report.doc (December 2008)
10. Aboba, B., Blunk, L., Vollbrecht, J., Carlson, J., Levkowetz, H.: Extensible Authentication Protocol (EAP) RFC 3748 (June 2004)
11. Clancy, T., Nakhjiri, M., Narayanan, V., Dondeti, L.: Handover Key Management and Re-Authentication Problem RFC 5169 (March 2008)
12. IEEE 802.21a Proactive Authentication and MIH Security 21-09-0102-02-0sec-proactive-authentication-and-mih-security.doc (September 2009)
13. Narayanan, V., Dondeti, L.: EAP Extensions for EAP Re-authentication Protocol (ERP). RFC 5296 (August 2008)
14. Aboba, B., Beadles, M., Arkko, J., Eronen, P.: The Network Access Identifier. RFC 4282 (December 2005)
15. Case, J., Fedor, M., Schoffstall, M., Davin, J.: A Simple Network Management Protocol (SNMP) RFC 1157 (May 1990)
16. Lopez, R.M., Dutta, A., Ohba, Y., Schulzrinne, H., Gomez, A.F.: Skarmeta Network-Layer Assisted Mechanism to Optimize Authentication Delay During Handoff in 802.11 Networks mobiquitous, MobiQuitous, pp.1–8 (2007)
17. Marin-Lopez, R., Pereiguez-Garcia, F., Ohba, Y., Bernal-Hidalgo, F., Gomez, A.F.: A Kerberized Architecture for Fast Re-authentication in Heterogeneous Wireless Networks. Mobile Networks & Applications, 1–21 (2010)
18. Machan, P., Wozniak, J.: Simultaneous handover scheme for IEEE 802.11 WLANs with IEEE 802.21 triggers. Springer Science+Business Media (2009)
19. IEEE 802.11r-2008, `http://www.ieee802.org/11/`
20. ODTONE, `http://hng.av.it.pt/projects/odtone`

Optimum Selection of Access Networks within Heterogeneous Wireless Environments Based on Linear Programming Techniques

Johnny Choque, Ramón Agüero, Eva-María Hortigüela, and Luis Muñoz

University of Cantabria, Santander, Spain
jchoque@tlmat.unican.es

Abstract. In this work we analyze the possibilities which are brought about by the use of linear programming techniques in the framework of access selection procedures within heterogeneous wireless network environments. We present a tool which has been designed and implemented (based on the *GLPK* package) to tackle this problem. This tool, starting from a particular network model, can be used to retrieve the optimum assignment of access elements. To fulfil this goal, we introduce a flexible cost (*utility*) function, which allows modulating the relevance given to the different aspects which could be taken into consideration while deciding the access alternative to be used: connection with a preferred operator, minimizing the number of handovers, or link quality, amongst others. Afterwards, the tools is used to study a set of canonical access selection strategies, so as to establish the combination of parameters which might lead to better performances.

1 Introduction

The recent proliferation of different wireless communication technologies has brought about the need of developing proper mechanisms to handle heterogeneity at the devices worn by regular users. The consequence is that near-future communication scenarios (in some cases they are referred to as 4G) will be characterized by offering a broad range of access alternatives, not only considering the involved technologies, but also from the point of view of the entities which manage the available networks.

In the aforementioned framework, it becomes necessary to revisit the new challenges which pop up in the access selection procedures, which are nowadays mostly based on non-automated processes, requiring in several cases, the direct intervention of the end-user. In contrast with this legacy approach, novel automated algorithms to ensure the most appropriate access alternative to the end user are deemed necessary; these should take into account various aspects, such as the user preferences, the particular network condition, as well as the punctual requirements of the current services.

In order to reach such an optimum performance, the involved network entities (base stations, access points, users, operators, etc.) must cooperate between

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 135–149, 2011.

them, by means of novel signalling mechanisms to transport the required control information. This information would, obviously, have a local *scope*, limited to the end-user, since he might not be aware of the potential consequences his decision can have over other network nodes. Additionally, in order to establish the suitability of the employed access mechanisms, it is fundamental to be aware of which is actually the best performance to be reached.

In this work we aim at analyzing such optimum behavior, using linear programming techniques. In this sense, we model a heterogeneous wireless networking scenario (embracing various technologies and operators) as an optimization problem, using an objective, *utility*, function, able to modulate the weight given to the various merit parameters. The *GLPK* library is used to solve the corresponding problem. The achieved results might help to establish, amongst others, the combination of the various weights which should be used by the access selection algorithms so as to guarantee an optimum behavior.

In order to fulfil the previously established goal, this paper has been structured as follows: Section 2 discusses some of the related work, highlighting the main differences with our approach. Section 3 presents the optimization problem to be solved with the framework depicted in Section 4. Section 5 describes the networking scenario which will be used to analyze a number of access selection strategies, whose performance will be discussed in Section 6. Finally, Section 7 concludes the paper, advocating some items which are left for future work.

2 Related Work

As it has been previously mentioned, the broad range of currently available radio access technologies (RAT), together with the increase of the number of devices which actually integrates various of them are some of the elements which promote the path towards future wireless access systems, mainly characterized by the large degree of heterogeneity, more notably from a technological perspective. In this type of scenarios, the procedures which are being currently used to perform access selection (which are mostly rather static, requiring the direct participation from the end-user) are not longer valid; as an alternative, the scientific community has been recently analyzing a number of proposals to reach, as their main goal, the *Always Best Connected* paradigm, considering user preferences, network situation and service requirements. As an illustrative example of the interest that this type of schemes has recently generated, it is worth mentioning the fact that the not only the scientific community, but also the relevant standardization bodies, have paid attention to them; in fact they are working towards the specification of procedures to promote the interoperability between heterogeneous networks. In this sense, the role taken by the IEEE 802.21 group clearly outstands amongst them; it has defined [1,15] a framework to facilitate handovers between heterogenous network technologies (*Media Independent Handover Framework, MIHF*).

The main advantage of being able to use various radio technologies is that it brings about the possibility of using, any time, the most appropriate characteristic of each of them in particular moments of time; on the other hand, this also imposes a series of drawbacks, derived from the complexity which is introduced into the system, since it becomes necessary managing a larger number of parameters. In this sense, there exist a number of works which propose solutions to tackle these challenges from various perspectives. For instance, the authors of [13] present the *Common Radio Resource Management* (CRRM), which is an evolution of the *Join Radio Resource Management* (JRRM) [6], proposing a joint management of the available radio resources, assuming a complete overlap of the coverage areas of the base stations (as it is also the case in [2]), and giving the full responsibility of the resource management to the network, thus following a clearly centralized approach. Another architecture which shares many of the characteristics of the CRRM is the so-called *Multi-Radio Resource Management* (MRRM) [14], in which a feature to be distinguished is how it deals with the cooperation strategies between different operators [9]. In the same line, different algorithms have been proposed to improve the assignment of resources, considering specific parameters of the involved wireless technologies [6,16], or additionally based on micro/macro economic profiles from either the end user [7] or the network [5].

The aim of this work is to go beyond the aforementioned works by considering, not only network preferences, but also (and as the most relevant aspect) end-user ones, while managing the available radio resources. By using linear programming techniques, the goal which we pursue is twofold: on the one hand we seek the establishment of a number of upper bounds on the performance which could be achieved with this type of frameworks, thus broadening the capacity to carry out sensible comparisons between different access selection algorithms; furthermore, by defining flexible utility functions (using various parameters), we could analyze the particular configurations which might lead to better behaviors. Although the use of linear programming techniques for this type of problems might seem to be a sensible choice, there are not, to our best knowledge, many works which have explored their possibilities, as it is discussed in [4]. Some of the existing ones, e.g. [12,3,11], are based on a reduced set of parameters which are combined in the corresponding objective function, resulting on a more rigid approach. Besides, it is worth highlighting another set of existing works, which benefit from the application of *Multi-Attribute Decision Making* (MADM) techniques, see e.g. [8,17,10]. The most relevant difference comes from the fact that they can be employed with the biased information which could be acquired by individual network elements (that is, as potential access selection algorithms), since one of their advantages is their computational efficiency; however they could not probably be used so as to analyze the overall optimum performance which could be achieved by the whole networks, which is precisely the main objective of this work; in fact, based on the conclusions which could be extracted by using the framework presented in this work, we could e.g. establish the criteria to be applied while carrying out access selection procedures based on MADM.

3 Optimization Problem Modeling

We consider a particular network scenario, in which M access elements (either base station or access points) are deployed. We assume that each of them would have a particular RAT and an associated capacity ϕ_j, as the maximum number of users they could serve. On the other hand, N end users aim at establishing a connection with one of the M access elements, since they wear devices able to use any of the involved RATs. We further assume that there are a number of different operators in the scenario, so that any of the access elements (j) would belong to a particular operator (ζ_j) and each user (i) would also have a preferred operator (η_i).

We define the set of basic variables, x_{ij}, defined as:

$$x_{ij} = \begin{cases} 1 & \text{if user } i \text{ is connected with access } j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Starting from this basic variables, we establish a number of parameters which will be favored when carrying out the optimization process. These aspects mainly represent the preferences which an end-user might have when taking a decision about the access alternative to connect to. In particular, the parameters which are defined below will be used.

- *Connectivity.* In this case we simply model the desire of any end-user to be connected to the network, so this can be somehow seen as the basic aspect to be optimized. We will use the parameter σ_{ij}, defined as:

$$\sigma_{ij} = \begin{cases} 1 & \text{if user } i \in \text{ coverage of access } j \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

- *Preferred operator.* This parameter aims at reflecting the fact that any end-user would like to connect to an access element from his preferred operator (agreement, better prices, etc). ψ_{ij} will be used to model this parameter, defined as:

$$\psi_{ij} = \begin{cases} 1 & \text{if } \eta_i = \zeta_j \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

- *Handovers.* Once an end-user is connected with an access element, he would like to maintain the connection as long as possible, so that he does not need to incur in the overhead (and degradation) associated to a handover procedure. In this sense, if the previous access element is known, we define the parameter λ_{ij} as:

$$\lambda_{ij} = \begin{cases} 1 & \text{if user } i \text{ was connected to access } j \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

- *Link quality.* When taking a decision on the access alternative to connect to, one of the parameters which has been traditionally used is the quality of the corresponding wireless link. Obviously, this is as aspect which depends

heavily on the particular wireless technology and the selected propagation model. For the sake of generality, it can be said that it can be modeled as a decreasing function with the distance to the access element[1]; in this case, a simple triangle function has been used, which takes the maximum value (1) at the very same position of the access element and the minimum (0), on the edge of the corresponding coverage area, and thus θ_{ij} is defined as:

$$\theta_{ij} = \begin{cases} 1 - \dfrac{d_{ij}}{\omega_j} & \text{if } d_{ij} < \omega_j \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where ω_j is the coverage of the access element j and d_{ij} the distance to the user i.

By combining these parameters, we define a *utility* function (u_{ij}), bringing about the possibility to establish a classification of the available access elements. As can be seen, if there is not physical connection between the end-user i and the access element j ($\sigma_{ij} = 0$), the utility function is 0.0.

$$u_{ij} = [\alpha + \beta \cdot \psi_{ij} + \gamma \cdot \lambda_{ij} + \delta \cdot \theta_{ij}]\, \sigma_{ij} \tag{6}$$

In order to make such function as flexible as possible, each of the aforementioned aspects is modulated with a weight; as such, α promotes the connectivity with any element; β emphasizes the connection with a preferred operator; γ is used to minimize the need to perform handovers; finally, δ empowers the quality of the wireless links. By examining all the parameters introduced above, it can be seen that all of them are delimited in $[0, 1]$, so if we fix that the sum of the four weights equals 1.0 ($\alpha + \beta + \gamma + \delta = 1$), we can also bound the utility of the connection between end-user i and access element j in the same interval. Furthermore, this can be also used to easily tweak the relevance associated to each of the parameters, bringing about the possibility to define differen access selection strategies.

Taking all of the above into consideration, the goal is to solve the following optimization problem:

$$\begin{aligned} \textbf{Max.} \quad & \sum_{i=0,j=0}^{N-1,M-1} u_{ij} \cdot x_{ij} \\ \textbf{s.t.} \quad & \sum_{j=0}^{M-1} x_{ij} \leq 1 \qquad i = 0 \ldots N-1 \\ & \sum_{i=0}^{N-1} x_{ij} \leq \phi_j \qquad j = 0 \ldots M-1 \end{aligned} \tag{7}$$

[1] It is worth highlighting that the goal of this work is not to accurately model the propagation model, but it focuses on the optimum selection of an access alternative. However, the implementation is flexible enough, so as the integration of more complex empirical propagation models would not impose great difficulties.

The first set of constraints fixes that an end-user can only establish a connection with just one access elements, while the second block limits the number of users connected to a particular access elements to its capacity (ϕ_j). It is important to mention that, for a particular network deployment, all the elements of the previously defined u_{ij} function exclusively depend on such scenario and thus, the only variables which need to be considered in the optimization problem are x_{ij}.

4 Design and Implementation of a Tool to Solve the Optimization Problem

As it has been said before, we will use the *GLPK* library to solve the problem which was presented in Section 3. This library offers an API which can be used from other application, and this eases the integration of the solver module. On the other hand the dimensions of the problem to be solved is within the limits which can be handled by this library, since the number of variables stays below 10^4 and, in addition, many of the corresponding coefficients (of both the objective function and the constraints) are 0.

Hence, a proprietary tool will be designed and implemented, to perform the following functionalities:

1. Read network parameters from a configuration file.
2. Deploy access elements from an external file.
3. Deploy the end-users from another external file.
4. Deploy and process the network scenario, obtaining all the parameters of the u_{ij} functions.
5. Solve the optimization problem with the *GLPK* library.
6. Process the solution.

As it was said before, one of the aspects that are considering to establish the value of u_{ij} is the previous connection of the end-user (so as to reduce, if possible, the number of required handovers). In this case, it is assumed that the position of the users varies (according to some specific mobility model) and, thus, the main programme must iterate a number of input files (*snapshots*), which reflect such mobility patterns. Each iteration must also use the solution of the previous one, so as to be able to establish the value of λ_{ij}.

Figure 1 depicts the high level flow diagram of the whole process.

Although the framework has been designed so that the deployment of both the access elements and the end-users could be done randomly, taking their positions from input files offers more possibilities, since it facilitates the integration with other platforms, once the format of the corresponding files has been fixed. This brings about the possibility of using the very same network deployments and movement traces, so as to foster more accurate comparisons with complementary approaches.

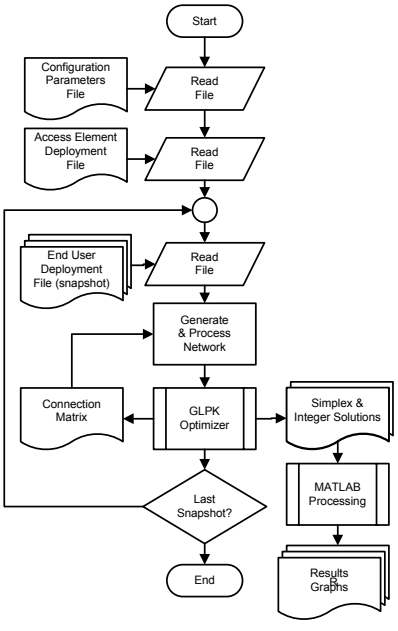


Fig. 1. Block diagram of the implemented optimization tool

5 Scenario Description

As it has been previously said, the main idea is to analyze different strategies to ensure the optimum access selection considering a heterogenous network deployment. The heterogeneity does not only extend to the involved technologies, but it also expands to the presence of various operators. In order to cover these requirements, three different access elements will be used, as can be seen in Table 1. The first one emulates a technology which could be closer to the characteristics of traditional cellular communications (e.g. GSM), since it has a notably larger coverage as well as a higher capacity. The two other access elements are clearly closer to WLAN access points, with a coverage and capacity much lower. It is important to mention that the capacity is modeled, generically, as the the maximum number of users which can be associated to the access element, e.g. without considering the subjacent traffic.

Table 1. Involved technologies

ID	Coverage (m)	Capacity	# Elements
ρ_0	600	20	4
ρ_1	80	5	16
ρ_2	60	5	20

We also assume that there are two different operators. The first one (A) is the incumbent operator, and such it manages the access elements of the legacy cellular RAT (ρ_0), while the second one (B) would emulate novel operator, which offers a less conventional access, by means of the access elements with technologies ρ_1 and ρ_2.

We consider a $1000 \times 1000\ m^2$ area, over which we randomly deploy (i.e. without any previous planning) the access element; note that we respect a minimum distance between access elements, provided that they belong to the same operator and use the same technology. Figure 2 shows the particular deployment which will be used; as can be seen, the 4 elements of RAT ρ_0 cover the whole scenario, resulting in highly overlapped areas. The surface covered by operator B is remarkably lower.

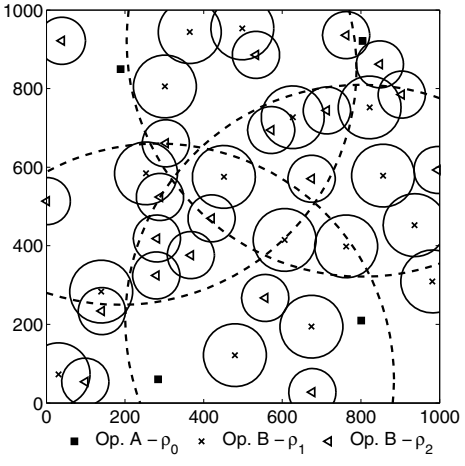


Fig. 2. Network deployment used during the analysis

Over the same scenario, 200 users are deployed, assuming that 60% of them belong to operator A, while the others have an agreement with the less conventional operator (B). It is further assumed that all users are able to use any of the involved technologies. All of them are randomly deployed at the beginning of the simulation, but afterwards they move according to the *Random Waypoint* model, with the characteristics which are enumerated in Table 2.

Once the scenario which will be used to carry out the analysis has been introduced, Table 3 depicts the different access selection strategies which will be studied. As can be seen, we modify the value given to the different weights, thus strengthening and prioritizing some of the different aspects which were previously discussed. In this sense strategy **A** has the only goal to maximize the number of connected users; **B** slightly favors the connections to the preferred operator and link quality, while **C** also prioritizes the reduction of the number of handovers. Strategies **D**, **E** and **F** focus (each of them) on a single parameter

Table 2. Characteristics of the *Random Waypoint* model used during the analysis

Characteristic	Value
Speed	$\mathbb{U}[2, 3]$ m/s
Movement time	$\mathbb{U}[100, 120]$ m/s
Pause time	$\mathbb{U}[5, 10]$ s
Edge policy	Reflection

Table 3. Analysis access selection strategies

Parameter	A	B	C	D	E	F	G
α	1.0	0.6	0.4	0.1	0.1	0.1	0.1
β	0.0	0.2	0.2	0.9	0.0	0.0	0.0
γ	0.0	0.0	0.2	0.0	0.9	0.0	0.7
δ	0.0	0.2	0.2	0.0	0.0	0.9	0.2

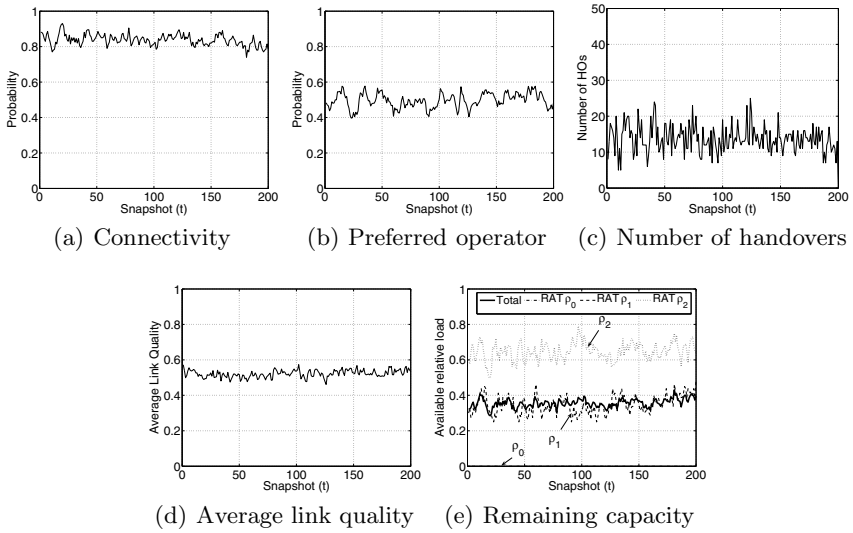
(preferred operator, number of handovers and link quality, respectively), providing a residual relevance to the connection probability. Finally, strategy **G** mainly aims at minimizing the number of handovers, but also considers link quality.

6 Discussion of Results

In the following we discuss the results which were obtained by using the 7 access selection strategies which have been previously presented. Each of he simulations lasts 2000 seconds, taking *snapshots* of the end-user position every 10 s, so the complete scenario involves the resolution of 200 optimization problems.

In order to evaluate the performance of the different alternatives, a number of metrics will be studied:

- *Connection probability.* It is the basic parameter, since it just considers whether the end-user was able to establish a connection or not.
- *Connection with the preferred operator.* It considers whether the connection was with an access element belonging to the preferred operator of the end-user.
- *Number of handovers.* It can be used to analyze the number of access element changes which must be performed.
- *Average link quality.* It averages the qualities of all the links which are established, using the function which was defined in Section 3.
- *Available load.* It is used so as to analyze the remaining capacity (for each of the involved technologies); this parameter can be used so as to determine whether more users could be accepted by the network.
- *Operator traffic.* In this case, the load accepted by each of the two operators will be analyzed, identifying the percentage of connections of users belonging to other operators, which could be used to estimate the incomes due to *roaming* circumstances.



(a) Connectivity (b) Preferred operator (c) Number of handovers

(d) Average link quality (e) Remaining capacity

Fig. 3. Instantaneous variation of the different merit parameters for the access selection strategy **G**

As a first step, to ascertain the validity of the framework, we will study the instantaneous evolution of the performance of a particular strategy; this would also shed some light on the stability of the results for a single scenario (set of *snapshots* with a single movement trace). Afterwards, 10 independent runs will be executed for each of the cases, averaging the results, so as to ensure tight confidence intervals, and to be able to compare the behaviors of the 7 access selection strategies.

Figure 3 presents the connectivity probability, the percentage of connections with the preferred operator, the number of handovers, the average link quality as well as the remaining network capacity (for the various RATs) for strategy **G**. As can be seen, none of the cases presents a remarkable variation around the average behavior. Connection probability reaches a value close to 85%, while only half of the connections are with the preferred operator; it is worth recalling that in this specific access selection algorithm, the weight given to that parameter is 0, and the distribution of the end-users between the two operators (*Market Share*) tells that 60% of them are *clients* of the incumbent operator. We also observe that the parameter in which a higher variability is observed is the number of required handovers, while it stays (for this particular configuration) below 10%. Regarding the average link quality, we can see that it maintains a value rather close to 50%, with little variability around it.

What it is even more interesting is the analysis of the remaining capacity for the different access element types. It can be seen that the available capacity of the whole networks is slightly below 40%. Although it might seem weird (by looking at the connectivity results), we must take into account the fact that the

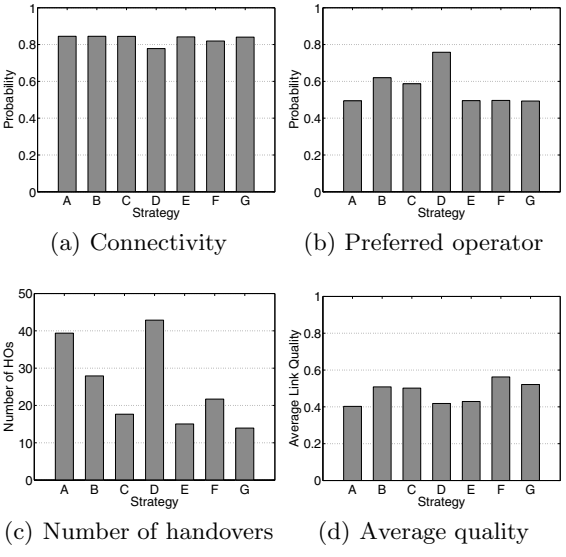


Fig. 4. Performance of the access selection strategies

access elements with RATs ρ_1 and ρ_2 have, overall, a relevant capacity (80+100), being even higher than the one offered by ρ_0, but they cover a much lower area. The consequence is that there are not enough end-users within the area covered by such RATs. In fact, the figure also allows us inferring the higher coverage of ρ_1, since the available capacity is lower than the one observed for ρ_2, whose utilization factor is rather low. Finally, we can also see that the access elements having the RAT ρ_0 are fully loaded, since as was previously seen completely cover the whole area and, thus, it is quite likely that there exist some users whose only alternative is, precisely, connecting to such access elements.

Although the instantaneous results permit acquiring a preliminary idea on the behavior of the various access selection strategies, it becomes more interesting being able to compare them, so as to study their advantages and drawbacks. Eventually, if we swept all the potential combinations for the different parameters, it would be possible to establish the combination which leads to the optimum performance. In this case, we will compare the 7 alternatives which were presented in 3, carrying out, for each of them, 10 independent measurements, averaging the corresponding results.

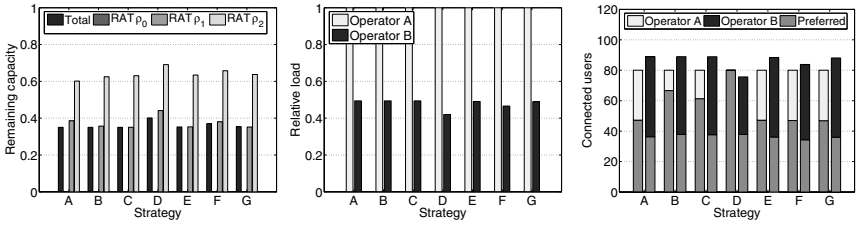
First, Figure 4(a) shows the connectivity which was observed for the 7 strategies; we can conclude that the different combinations do not have a strong influence over the achieved results, since it can be seen that in all cases, the connectivity probability is very close to 0.8. The graph yields a slight decrease in both **D** and **F** strategies (higher in the first one); by using **D** connections with the preferred operator are favored, and thus, the access elements using ρ_0, which belong to the incumbent operator, easily fill their available capacity; this, together with the fact that establishing a connection with either ρ_1 or ρ_2 is less

likely (since they cover a smaller area) justifies the observed decrease, which stays below 10%. The reduction is even less relevant for strategy **F** due to the fact that in this case, the connection with closer (higher quality) access elements is strengthened, thus increasing the *utility* give to ρ_0 access elements.

On the other hand, we can clearly see, in 4(b), that the differences regarding the connections with non-preferred operators are much more relevant, depending on the particular access selection strategy which is being used. In those combinations in which the weight given to the β parameter in the utility function is 0, the probability of establishing a connection with an access element of the preferred operator does not vary (staying around 50%); we must take into account that 60% of the users are *clients* of operator which manages the access elements with a wider coverage (ρ_0). On the other hand, by comparing the results of strategies **B**, **C** and **D** it can be inferred that the additional gain which is achieved when increasing β from 0.2 (in the two first ones) to 0.9 (in the last one) is, approximately, 15%; last, the connection to the preferred operator is slightly lower in strategy **C** than in **B**, since in the first one, the utility function also favors the quality of the wireless link, which slightly decreases the number of connections with the preferred operator.

Figures 4(c) and 4(d) shows the results for both the number of handovers and the average link quality for the 7 analyzed strategies. First, we can see that there exists a clear relationship between both parameters. In this sense, if we compare the results of strategies **A** and **B** (in which the γ parameter, weight which favors the reduction of handover events, takes the same value), we shall see that the number of handovers in the first strategy is notably higher ($\gtrsim 25\%$); the reason behind this is that favoring the links with a higher quality ($\delta = 0.2$ in strategy **B**) causes that more connections are established with closer access elements, and thus, it is less likely that a handover will be required in the following iterations. This aspect is reflected somehow if comparing strategies **E** and **G**; the first only just favors the minimization of handovers ($\gamma = 0.9$), while in the second one, γ is reduced 20%, so as to correspondingly augment the weight given to the link quality to $\delta = 0.2$; as can be seen, the consequence is that the average quality of the established connections is notably improved (around 10%), without negatively affecting the number of handovers; in fact there is a slight reduction (although the value of γ is lower).

Besides, Figure 5 can be used so as to analyze the influence of the different weight combinations over the available network capacity, both from the perspective of the type of RAT and operator. As it was observed before, access elements of RAT ρ_0 (which match the overall capacity of operator A) use up all their resources, for all the strategies which have been studied. Obviously, the connectivity results (Figure 4(a)) correspond to the remaining capacity and it is strategy **D** the only one in which we observe a higher available capacity, since the access elements from operator B slightly reduce their occupancy. With a clear relationship with these latest results, Figure 5(c) shows the number of connected users per operator, highlighting those which are connected to their preferred operator; these results reflect the relevance of the β parameter, since



(a) Remaining capacity (b) Load per operator (c) Connected users

Fig. 5. Performance of the access selection strategies - capacity and load

in those strategies in which it takes a value higher than 0 (**B**, **C** and **D**), there are more users connected to their preferred operator. We shall also see that this only affects the incumbent operator, since the number of users that, being clients of B, can connect to their preferred operator stays unchanged (around 40 users) for all cases. The consequence is that, in strategy **D**, all users connected to operator A (80, which is all its overall capacity) belong to its clients, and therefore the load carried by operator B is reduced. Thanks to these results we could probably aim at establishing different cooperation mechanism to increase their profit, by applying different prices for the traffic coming from own clients or *roaming*.

7 Conclusions

This works has exhibited the possibilities that the application of linear programming techniques open in the analysis of access selection strategies, over heterogeneous network scenarios (considering both the involved technologies and the operators). We have defined an optimization problem, in which the objective function (to be maximized) can be adapted using a number of parameters which establish the different priorities that an end-user might have when deciding on access alternative amongst the others.

In order to solve the problem, we have designed and implemented an application, which is based on the *GLPK* library. Such framework has been used to study a set of access selection strategies, in which the weight given to the various parameters was modified, so as to analyze which is their influence. Using a scenario with two operators and a remarkable heterogeneity (involved radio technologies) it has been shown that there are certain combinations which offer better performances than others (at least, in terms of the number of handovers and the quality of the established wireless links). The analysis has show, e.g. that prioritizing the choice of higher quality radio links brings about a decrease of the number of handovers. We have also analyzed the influence of the various parameters in how the load is distributed between the different types of access elements (and operators).

From the work presented in this paper, a relevant number of future research lines could be opened; first, the definition of the utility function is very flexible

so that the inclusion of new metrics within the access selection procedure or tweaking their weights could be done quite easily. This flexibility could be used e.g. to carry out a more refine variation of their values, so as to determine which is the optimum combination. Furthermore, the results obtained from this study, if we interpret them as the best that can be achieved for a particular network scenario, could be compared with complementary analysis, considering only the information (with a more local scope) that is confined to the end-user environment. Finally, we will incorporate more realistic traffic models, considering also the possibility to modify the capacity required when starting an application. Finally, we could also modify the parameters of the scenario, in terms of the market share, number of users, access element deployment, etc, so as to study their influence over the overall system performance.

Acknowledgements

The authors would like to express their gratitude to the Spanish government for its funding in the following two projects: Mobilia - CELTIC Program (Avanza I+D TSI-020400-2008-82) and "Cognitive, Cooperative Communications and autonomous SErvice Management", C3SEM (TEC2009-14598-C02-01).

References

1. IEEE standard for local and metropolitan area networks- part 21: Media independent handover. IEEE Std 802.21-2008 (2009)
2. Chen, B., Chan, M.: Resource management in heterogenous wireless networks with overlapping coverage. In: First International Conference on Communication System Software and Middleware, Comsware 2006 (2006)
3. Falowo, O., Anthony Chan, H.: Optimal joint radio resource management to improve connection-level qos in next generation wireless networks. In: Radio and Wireless Symposium. IEEE, Los Alamitos (2008)
4. Falowo, O.E., Chan, H.A.: Joint call admission control algorithms: Requirements, approaches, and design considerations. Comput. Commun. 31(6), 1200–1217 (2008)
5. Giupponi, L., Agusti, R., Perez-Romero, J., Salient, O.: Improved revenue and radio resource usage through inter-operator joint radio resource management. In: IEEE International Conference on Communications, ICC 2007 (June 2007)
6. Giupponi, L., Agusti, R., Perez-Romero, J., Sallent, O.: Joint radio resource management algorithm for multi-RAT networks. In: Global Telecommunications Conference, GLOBECOM 2005, vol. 6. IEEE, Los Alamitos (2005)
7. Giupponi, L., Agusti, R., Perez-Romero, J., Sallent, O.: WLC05-2: An economic-driven joint radio resource management with user profile differentiation in a beyond 3G cognitive network. In: Global Telecommunications Conference, GLOBECOM 2006. IEEE, Los Alamitos (2006)
8. Gu, C., Zhang, Y., Ma, W., Liu, N., Man, Y.: Universal modeling and optimization for multi-radio access selection. In: 5th International Conference on Wireless Communications, Networking and Mobile Computing, WiCom 2009 (2009)
9. Ho, L., Markendahl, J., Berg, M.: Business aspects of advertising and discovery concepts in Ambient Networks. In: IEEE 17th International Symposium on Personal, Indoor and Mobile Radio Communications (2006)

10. Hu, H., Zhou, W., Zhang, S., Song, J.: A novel network selection algorithm in next generation heterogeneous network for modern service industry. In: Asia-Pacific Services Computing Conference, APSCC 2008. IEEE, Los Alamitos (2008)
11. Karthikeyan Krishnasamy, N., Narayanasamy Palanisamy, P.: Bandwidth allocation scheme for multimedia mobile networks using optimization techniques. In: 2006 IEEE Conference on Cybernetics and Intelligent Systems (June 2006)
12. Lucas-Estan, M., Gozalvez, J., Sanchez-Soriano, J.: Common radio resource management policy for multimedia traffic in beyond 3G heterogeneous wireless systems. In: IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2008 (September 2008)
13. Perez-Romero, J., Sallent, O., Agusti, R., Karlsson, P., Barbaresi, A., Wang, L., Casadevall, F., Dohler, M., Gonzalez, H., Cabral-Pinto, F.: Common radio resource management: functional models and implementation requirements. In: IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2005 (2005)
14. Sachs, J., Aguero, R., Daoud, K., Gebert, J., Koudouridis, G., Meago, F., Prytz, M., Rinta-aho, T., Tang, H.: Generic abstraction of access performance and resources for multi-radio access management. In: 16th IST of Mobile and Wireless Communications Summit (July 2007)
15. Taniuchi, K., Ohba, Y., Fajardo, V., Das, S., Tauil, M., Cheng, Y.H., Dutta, A., Baker, D., Yajnik, M., Famolari, D.: IEEE 802.21: Media independent handover: Features, applicability, and realization. IEEE Communications Magazine 47(1), 112–120 (2009)
16. Tolli, A., Hakalin, P., Holma, H.: Performance evaluation of common radio resource management (CRRM). In: IEEE International Conference on Communications, ICC 2002, vol. 5 (2002)
17. Xing, B., Venkatasubramanian, N.: Multi-constraint dynamic access selection in always best connected networks. In: The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services, MobiQuitous 2005 (July 2005)

On the Empirical Analysis of Handover Latency Reduction by Means of Multi-RAT Devices: A Prototypical Approach

David Gómez[1], Ramón Agüero[1], Jesús Herrero[2], Bruno Cendón[2], and Luis Muñoz[1]

[1] University of Cantabria, Santander, Spain
[2] TST Sistemas, Santander, Spain
{dgomez,ramon,luis}@tlmat.unican.es,
{jherrero,bcendon}@tst-sistemas.es

Abstract. In this paper we present an fully empirical assessment of the possibilities which are brought about by an architecture able to handle multiple wireless access technologies. Heterogeneity is believed to play a key role in forthcoming communication scenarios and therefore some entities able to appropriately tackle the new challenges are deemed necessary. Although reaching the *Always Best Connected* paradigm has gathered the interest of the scientific community, many of the existing works are descriptive (architectural papers) or based on simulation and/or emulation. In this paper we go a step beyond this and starting from the architecture proposed in the Mobilia Celtic project, we deploy a real platform to showcase two illustrative handover examples, triggered either by the end-user and the network. Additionally, we use the same platform to quantitatively analyze the enhancement which the use of multi-RAT devices may provide, in terms of the handover latency reduction.

1 Introduction and Objectives

The advent of new wireless technologies has paved the way of a largely heterogeneous wireless communication environments, where new challenges emerge. One of the most remarkable ones is the need to foster the *Always Best Connected* paradigm, where the end user is always served with the most appropriate access alternative, depending on a set of factors. These embrace user preferences and policies, requirements from the current services, as well as the particular conditions of the available networks and access alternatives. On the other hand the existence of agreements (cooperative) between different entities must also be considered. Mechanisms which are available at the time of writing clearly do not fulfil with this requirements and usually require the direct involvement of the end-user.

Opposed to that we present an architecture able to deal with the large heterogeneity which will characterize forthcoming wireless communication scenarios, since it offers the end-user with the most appropriate access. The proposal is the cornerstone of the *Mobility concepts for IMT-Advanced (Mobilia)* project,

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 150–162, 2011.

belonging to the CELTIC programme. In particular, in this work we present a real platform which has been used so as to assess the feasibility of the this proposal. We describe two use cases, illustrating different ways to trigger access selection procedures (initiated either by the network or by the end-user). In addition, the paper also illustratively shows the gains which could be expected if taking full advantage of the possibility of having multi-RAT devices, in terms of the latency reduction during a handover event.

In order to tackle the aforementioned objectives, the paper has been structured as follows: Section 2 discusses related work, paying special attention to other works which have pursued an experimental approach; Section 3 presents the architecture which has been designed in the framework of the Mobilia project. Section 4 describes how this architecture was translated into a real platform, depicting the use cases which have been used to challenge its feasibility, while Section 5 presents a set of results achieved with such platform, to highlight the enhancements that the smart use of multi-RAT devices (conveniently fostered) in terms of latency reduction during handover events. Finally, Section 6 concludes the paper, advocating some items for future work.

2 Related Work

During the latest years, the presence of heterogeneous wireless networks has gathered the attention from the scientific community. As a result, there are a number of proposals which have tackled the problems that arise in such scenarios. Some of them share many of the characteristics of the Mobilia architecture, and therefore they are worth mentioning herewith.

Two of the most complete ones are the *Common Radio Resource Management (CRRM)* and *Joint Radio Resource Management (JRRM)* (see e.g. [8,3] and the references therein) and the *Multi-Radio Resource Management (MRRM)* ([9] and references therein). Both of them assume the presence of an entity which harmonizes the information from the lower layers, so that it can be compared on a technology-agnostic way. Afterwards, a smart entity, using such information, together with other requirements takes a decision on the most appropriate available access to serve the current demand.

In parallel, the growing heterogeneity has been also addressed by the relevant standardization bodies. In this sense, the IEEE 802.21 has defined the *Media Independent Handover Framework (MIHF)*, whose main goal is to provide some means to transport relevant signalling information which should be taken into consideration during handover procedures [10,1]. One of the distinguishing characteristics of the Mobilia architecture is that we have considered the use of IEEE 802.21 as a focal aspect; therefore, we do not consider it *as a signalling transport mechanism*, but we extend its original scope and we provide APIs for the various entities which are part of the proposed architecture. The idea is that we do not establish any bounds on where and how to use such facilities. The Mobilia architecture will also consider the possibility to support cognitive radio capability, since this feature is believed to play a key role in future wireless communication networks.

The focus of this paper is to present the real platform which was deployed in order to assess the feasibility of the proposed architecture and, using such framework, provide illustrative figures of the gain which could be expected if taking full advantage of wearing multi-RAT devices. In particular, the paper will provide some performance figures of the latency which might be expected during a handover process. It is important to highlight that all the results are completely based on real components and their performances and no emulation has been employed.

There are some works which also seek the real implementation of this type of architectures. For instance, in [2] the authors present a real platform which reflects the Multi Radio Architecture which was proposed in the framework of the Ambient Networks project. However, no results are presented, and the different entities were not included in the network elements (access points and base stations), thus limiting the possibility to initiate handovers from the network. Other works from the same project, e.g. [6,7] analyze the overhead during a vertical handover process, based on the *Host Identity Protocol (HIP)* mobility solution; they use real heterogeneous technologies (namely 3G and IEEE 802.11), but they do not incorporate any functionality within the network elements and therefore the use cases which are analyzed are always triggered at the end user device. The solution is based on a triggering entity which delivers events to interested entities.

Another completely different approach is the use of large testbeds which emulate the behavior of heterogeneous wireless networks; one of the most relevant ones is the AROMA platform [4], which has been used to analyze the performance of the aforementioned CRRM. In these approaches, the goal is to perform an exhaustive performance analysis, usually mimicking the role which traditionally has been assigned to simulators, but offering a greater degree of flexibility and accuracy; however, they need to incorporate different models for the movement partners, traffic sources, etc. Opposed to that we have a more concrete scenario, which we use to characterize specific procedures, studying the enhancements which might be brought about when using different access discovery and selection mechanisms over real platforms.

3 Mobilia Architecture

Mobilia is trying to face the new challenges appearing with the forthcoming wireless communication scenarios, where heterogeneity is one of the most common factors. One important goal in this project is supporting an efficient and transparent cooperation between heterogeneous access networks, following the concept of Always Best Connected. For that reason, a great effort has been made working in the definition of an architecture which enables to deal with the selection of access networks in highly heterogeneous environments. The global framework proposal is based on various principles, such as the abstraction of the subjacent technologies, the consideration of a number of metrics in the selection procedures and the use of the signaling framework, proposed by Media

Independent Handover (MIH), as the transport medium for the exchange of information between different entities in the network [1]. This signaling framework is being standardized by the IEEE 802.21 working group and means the distinctive feature to be explored in Mobilia. The IEEE 802.21 signaling is based on a Type-Length-Value (TLV) codification, which offers different advantages from the implementation point of view. The next figure shows the global architecture to be implemented in the mobile user involved in the platform implementation. In addition, it would be necessary to include some of these functionalities in another entities appearing in the prototype, such as the Access Points and the Network Server.

Several entities have been presented in Figure 1 in order to ensure a suitable management of vertical handover situations. The next paragraphs make a short description of the functionalities associated with each of them.

Link Layer. Based on the MIH_LINK_USER SAP defined on the IEEE 802.21 specification, it facilitates to the different Radio Access Technologies (RATs) the data exchange with the Abstraction Layer.

Media Independent Handover Function (MIHF). It is the core element of the architecture; its main function is to enable, either locally or remotely, the exchange of information and commands between the different devices involved in making and executing handovers decisions. The MIHF enables a fair and technology-agnostic comparison of the characteristics of the subjacent radio

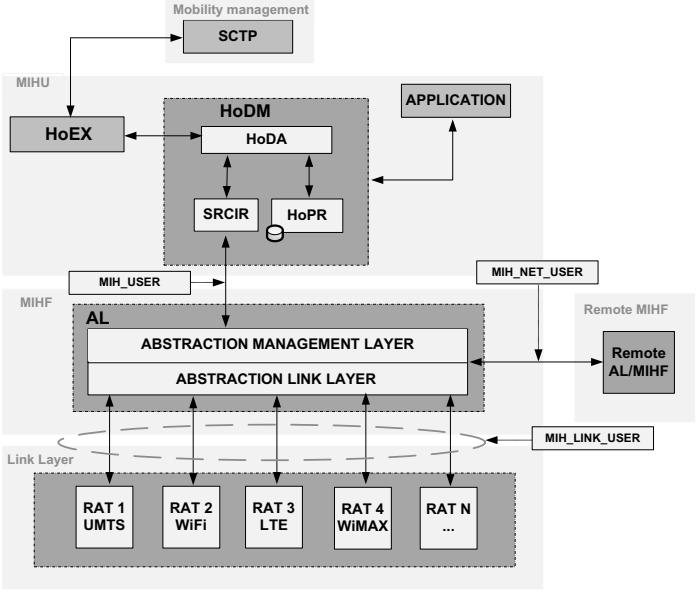


Fig. 1. A global vision of the architecture proposed by Mobilia to handle vertical handover situations

technologies and it has been structured into two different components: the Abstraction Management Layer (AML) which is in charge of managing the User MIH interface (MIH_SAP) and the remote interface (MIH_NET_SAP) and the Abstraction Link Layer (ALL) which manages the interfaces with the Link Layer (MIH_LINK_SAP).

Media Independent Handover User (MIHU). The MIHU is performed by two elements: The Handover Decision Manager (HoDM) and the Handover Execution Manager (HoEM). The HoDM is the most important part of this entity, being the element with the final responsibility during the execution of the handover. All the data meaning and decision criteria are located at this level of the architecture. It is composed by the following elements:

- Handover Decision Algorithm Module (HoDA) which includes the necessary intelligence to decide when a handover shall be done.
- Handover Policies Repository (HoPR), which is used by the HoDA to set the basis of the conditions to order a handover. This information is combined with the one provided dynamically by the Service Requirements Collector Information Repository (SRCIR).
- Service Requirements Collector Information Repository (SRCIR), which implements the MIH_USER interface to the AL/MIHF and act as a dynamic repository for the information collected from the MN and the Network. This information is used by the HoDA.

The HoEM interfaces directly with the MIHF through the SRCIR to manage the primitive exchange in order to execute the handover. This element owns the necessary intelligence to manage the event reporting and error handling that the handover procedure might create.

Remote MIHF. It provides the framework with the capabilities to share relevant information with remote elements.

Stream Control Transmission Protocol (SCTP). It is a reliable, general-purpose transport layer protocol for use on IP networks that addresses mobility needs [2] and allows high availability, increases reliability, and improves security for socket initiation.

4 Demonstration Setup

As it was previously discussed, one of the main goals of this work is to carry out an empirical assessment of the architecture presented in the previous section, thus filling the gap which exists regarding this type of approaches. However, this poses several constraints that need to be taken into account: first, the availability of the wireless technologies is limited; although there are various *off-the-shelf* wireless technologies available, most of them do not offer the flexibility which we deem necessary to challenge the Mobilia architecture. Basically, this embraces the availability of appropriate drivers and the possibility to introduce modifications at the network side (i.e. at the access element). Considering these intrinsic limitations, the platform is based on the IEEE 802.11 technology; we emulate

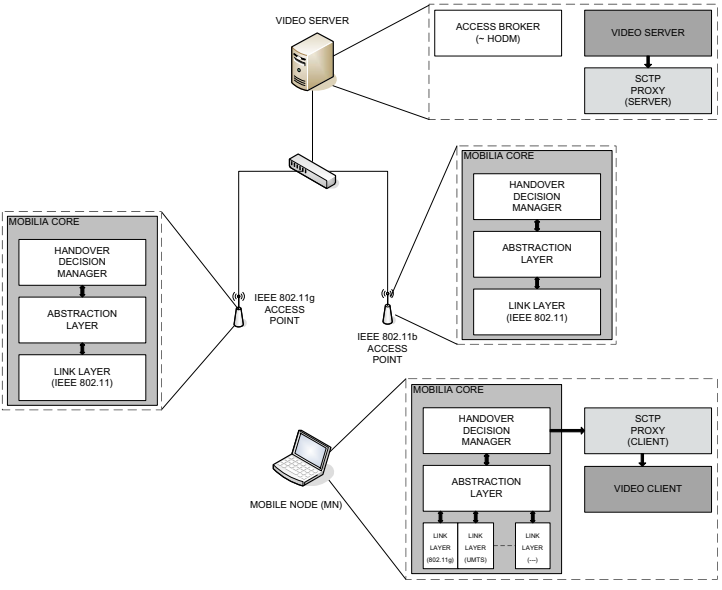


Fig. 2. Demonstrator platform

the presence of heterogeneous networks by deploying two access points, configured in two orthogonal channels, which take the role of access elements. This approach, which undoubtedly present some limitations, also has clear advantages: the support of the corresponding drivers is quite important, and available interfaces allow the user obtaining several pieces of information (which can be used as elements to be considered during the access selection process); in addition, by means of the `madwifi` project [5], it is possible to use a regular laptop to deploy an access point, therefore bringing about the possibility to introduce deployments at the network side.

With the aforementioned limitations in mind, the proposed (minimum) platform comprises four laptops, as can be seen in Figure 2. Their role is briefly introduced below.

- **Mobile Terminal (MT).** It takes the role of the device which an end-user would carry; its main characteristic is that it might include more than one wireless interface.
- **Access Elements (AE).** These two boxes emulate the role of any network element providing access (Access Point, Base Station, etc.); in order to emulate heterogeneous technologies, the two access elements are configured to work in orthogonal (non-overlapping) wireless channels.
- **Access Broker (AB).** In order to promote an efficient management of the available resources, the role of an access broker (either centralized or distributed) might be fundamental; in the deployed framework, this box

is connected with the two access elements by means of an infrastructure network (Ethernet); in addition, it includes some of the applications which will be used to generate traffic during the various experiments.

Both the **MT** and the **AE** incorporate the whole Mobilia architecture, while the **AB** only includes an instantiation of the *HoDM* module, since it does not directly manage any wireless resources, and thus it does not need to incorporate the *AL* nor any *LL*. In addition, in order to follow the demonstration on a friendly way, each of the elements incorporate a *GUI* which graphically shows the most relevant events and can also be used so as to configure some of the parameters.

As it was briefly introduced before, the platform adopts SCTP as the mobility solution. More precisely, it includes a SCTP proxy in both sides of the communication (both the end-user and the video server); this proxy basically forwards the traffic through a *tunnel*, while the real end-point can be dynamically adjusted by means of proprietary commands (those were implemented through `ioctl` after the commands sent by the corresponding Mobilia entities).

Two different use cases are challenged over this platform: the first one mimics a traditional handover triggered by the end-user when he detects a decrease of the link quality with the current access element; in the second case, the network decides to initiate a handover, due to a high congestion situation.

The detailed message interchange flow chart is shown in the following section.

4.1 Use Case 1: End-User Initiated Handover

In this case, the device perceives a decrease on the link quality with the current access network; as a consequence, it triggers a handover request, which might eventually lead to a change on the serving network.

There are two specific aspects which are worth highlighting: first, all the process is handled by the different elements of the Mobilia architecture and, therefore, the end-user should not perceive any quality of service degradation; furthermore, we benefit from the *GUI* and we emulate the link quality indicator (e.g. *RSSI*)[1].

Once the end-user is connected, and a video streaming session is initiated from the server, the *HODM* configures the *AL* so that to receive notification when a predefined threshold is crossed (this configuration can be done from the *GUI*, which uses a proprietary message to the *HODM*). By using the link quality emulation facility of the *GUI*, the link quality is decreased, and the corresponding event (*LINK DOWN*) is generated from the *LL* and travels up to the *HODM*. In this case, there are various alternatives, depending on the particular capacities of the device: if it has more than one interface, the end user would be able to benefit from this, minimizing the latency during a handover event; another possibility would be to use some sort of control channel to retrieve

[1] This is done as a means to facilitate the whole demonstration process since the **LL** is able to get the estimation provided by the subjacent wireless cards; it has to be considered that nowadays, wireless activity is usually rather heavy, and this might lead to undesirable fluctuations.

information from the network, which might provide location-aware data to ease the access selection procedure. In the particular case of this work, since there are two interfaces, the change of access takes place without any service disruption (carrying out a *make-before-break* handover); before actually changing the video flow, the device connects to the destination network, and once this is ready, the flow can be changed. In the demo, the SCTP proxy is in charge of finishing the handover.

The detailed interchange of primitives between the involved entities is depicted in the following section.

4.2 Use Case 2: Network Initiated Handover

In this case, the handover is triggered by the network; in this sense, an overload situation is emulated by means of the *GUI* of the current serving access element[2]. Again, once a predefined threshold is crossed, the *LL* notifies this situation to the *AL*, which passes it to the *HODM*. In this case, we assume that there is not any established agreement between the two access elements and therefore, the current serving one asks the access broker about the possibility to change the flow. Since the access broker is aware of the whole network topology (it would know e.g. whether there are enough resources in the networks which are accessible from the end-user position) will instruct the end-user to initiate the handover (this message goes from the *HODM* at the access broker to the *HODM* at the end-user device, through the current access element (which still gives access to the end-user). From this moment, the procedure looks like the previously described handover. Again, we benefit from the advantages brought about by the fact of having two interfaces so as to reduce the latency during the handover procedure.

5 Results

This section starts from the demo setup detailed in Section 4, evaluating the delay suffered by a mobile node when executing a user-transparent vertical handover from an access element to another one.

Specifically, we will challenge four different situations, which are briefly described below. It is worth highlighting that the main goal would be to assess the benefits that the use of multi-RAT devices may bring about.

A. The first one consists of a device with a single IEEE 802.11 interface (thus illustrating the case of most terminals at the time of writing), therefore it must carry out the whole handover process and all the involved tasks, such as scanning for networks and connecting to a new cell. This case can be described as the least desirable one, since it breaks the connection twice (see Figure 3), but it is included for comparison purposes. In this case, we assume that the IP address on the new access is statically configured and hence, there is not any additional delay due to IP assignment.

[2] Again this is just a simplification for the sake of demo purposes, since the *LL* of the access element is able to measure the load by means of the number of associated clients or using the size of the transmit buffer.

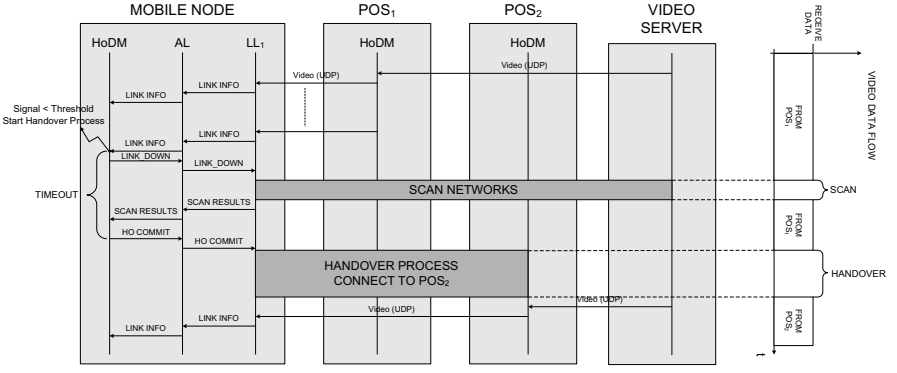


Fig. 3. Flow chart for the single-interface case

B. In this case a new IEEE 802.11 interface is added to the device. One of the available RATS is used for connecting to a new network, whereas the new one would only act when a scanning task is required. Hence, the data flow only stops at the handover process, saving the time waisted for looking for available networks. This situation also reflects what would happen if the network could provide location-aware information through a common transport channel (provided it is available), since this would allow the mobile to change the access element without scanning for other networks.

C. For this case, we take advantage of the fact that the device has more than one interface, so we can make a soft-handover, where the second interface is used for connecting to a new network when a LINK_GOING_DOWN message is received from the HODM, while the data flow continues through the other interface. Once the whole handover process is completed, the mobile node instructs the video server to modify the destination IP address[3], so the packets change the route to the new one, thus saving the additional time introduced by the 'standard handover', as can be inferred by looking at Figure 4. This technique is also known as *make-before-break*.

D. As we have already mentioned, in the previous three cases it is assumed that the assignment of IP address does not consume any additional time; this might not be realistic, and therefore, in this last case we would like to analyze the overhead caused by a legacy DHCP IP assignment. We take the first case as the starting point, and we configure the interface so as to request for an IP address once it is connected to the new network; needless to say, this increases the handover latency, as can be seen on Figure 5. It is worth mentioning that this delay would not affect any of the other cases, since the DHCP procedure might be running in background, and only when it is fully accomplished, the handover takes place.

[3] The default video server is not able to change an IP address by itself (while the application is running, so (as already mentioned before) we have introduced a proxy SCTP component in order to sort this out.

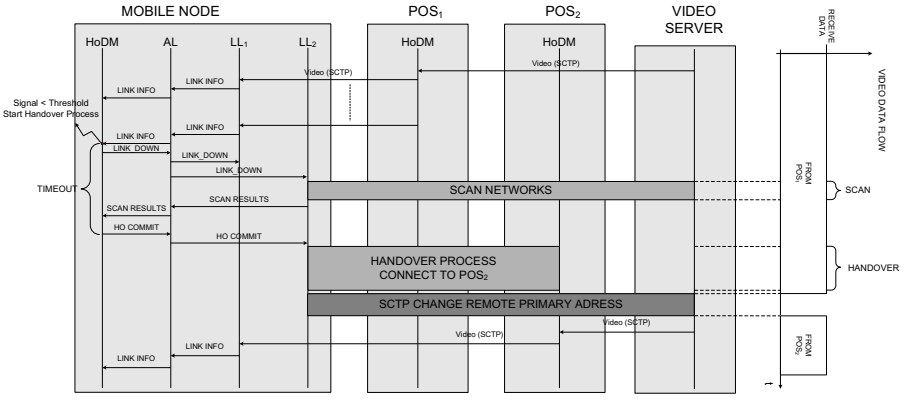


Fig. 4. Flow chart for the dual-interface case with SCTP

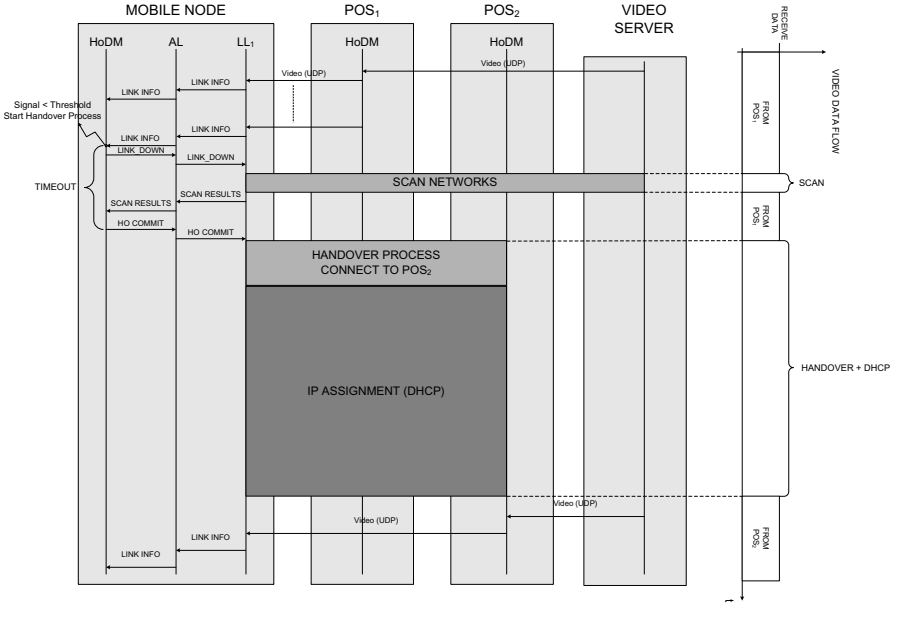


Fig. 5. Flow chart when IP address is automatically assigned with DHCP

Figure 6 shows the average handover latency and the 95% confidence interval (obtained by means of the *t-student distribution*) of a series of 10 independent measurements[4]. It is worth highlighting that, although 10 measurements might seem not to be enough, all the values were really close to each other, so we could conclude that the average performance figures are accurate.

[4] Without considering the time lost in carrying out the network scanning (since, it only affects case A).

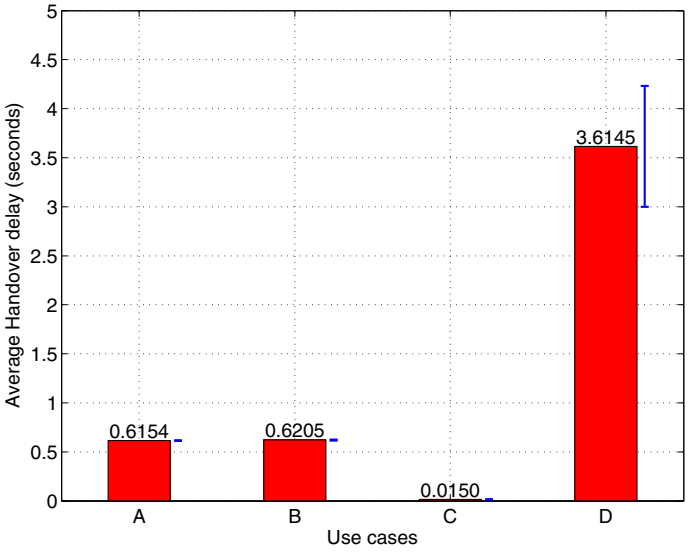


Fig. 6. Average handover latencies

These experiments have been measured at the Mobile Node by using the `Wireshark` [11] packet capture tool. The handover latency has been calculated as the interval between the reception of the last packet via the old link and the first packet received though the new one.

As can be seen, case A and B follow the same handover process, so the results are almost equal (the difference lies in which interface is in charge of scanning for the neighbor networks, breaking the connection for a while if it happens through the same interface), stopping the traffic for roughly half a second. Nevertheless, case C provides the system a smart handover process, avoiding the loss of connectivity due to the fact that the active connection is broken before changing to a new access element. Finally, case D shows the delay resulting between a link layer handover and the link layer IP address negotiation, which introduces a significant increase on the overall latency.

6 Conclusions

This paper has presented a real platform, based on the architecture proposed by the Mobilia Celtic project to showcase access selection procedures over heterogeneous wireless networks. In particular, we have presented two illustrative handover procedures: the first one is initiated by the end-user device (after a decrease on the link quality with the current serving network), while the second one mimics a situation where the handover is triggered by the network, after the access element detects an overload.

The implementation is afterwards used to analyze the improvements which might be brought about by the smart usage of multi-RAT devices; the handover latency of four different cases is empirically studied; the measurements show that taking advantage from having more than one interface might lead to remarkable performance enhancements.

In the future we plan to extend the platform in various ways: first, the signalling between all the different entities will be based on the IEEE 802.21 standard; furthermore, since the Mobilia architecture is rather orthogonal to the mobility solution, it would be interesting using other alternative solutions, like Mobile IP or HIP.

Acknowledgements

The authors would like to express their gratitude to the Spanish government for its funding in the following two projects: Mobilia - CELTIC Program (Avanza I+D TSI-020400-2008-82) and "Cognitive, Cooperative Communications and autonomous SErvice Management", C3SEM (TEC2009-14598-C02-01).

References

1. IEEE standard for local and metropolitan area networks- part 21: Media independent handover. IEEE Std 802.21-2008 (2009)
2. Aguero, R., Gebert, J., Choque, J., Eckhardt, H.: Towards a multi-access prototype in ambient networks. In: 16th IST of Mobile and Wireless Communications Summit (2007)
3. Giupponi, L., Agusti, R., Perez-Romero, J., Salient, O.: Improved revenue and radio resource usage through inter-operator joint radio resource management. In: IEEE International Conference on Communications, ICC 2007 (June 2007)
4. López-Benítez, M., Bernardo, F., Vucevic, N., Umbert, A.: Real-time emulation of heterogeneous wireless networks with end-to-edge quality of service guarantees: The AROMA testbed. EURASIP Journal on Wireless Communications and Networking (2010)
5. Madwifi Project: Madwifi – Multibrand Atheros Driver for Wireless Fidelity, http://madwifi.org/
6. Paakkonen, P., Salmela, P., Aguero, R., Choque, J.: An integrated ambient networks prototype. In: 15th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2007 (2007)
7. Paakkonen, P., Salmela, P., Aguero, R., Choque, J.: Performance analysis of HIP-based mobility and triggering. In: WOWMOM 2008: Proceedings of the 2008 International Symposium on a World of Wireless, Mobile and Multimedia Networks (2008)
8. Perez-Romero, J., Sallent, O., Agusti, R., Karlsson, P., Barbaresi, A., Wang, L., Casadevall, F., Dohler, M., Gonzalez, H., Cabral-Pinto, F.: Common radio resource management: functional models and implementation requirements. In: IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2005 (2005)

9. Sachs, J., Aguero, R., Daoud, K., Gebert, J., Koudouridis, G., Meago, F., Prytz, M., Rinta-aho, T., Tang, H.: Generic abstraction of access performance and resources for multi-radio access management. In: 16th IST of Mobile and Wireless Communications Summit 2007 (July 2007)
10. Taniuchi, K., Ohba, Y., Fajardo, V., Das, S., Tauil, M., Cheng, Y.H., Dutta, A., Baker, D., Yajnik, M., Famolari, D.: IEEE 802.21: Media independent handover: Features, applicability, and realization. IEEE Communications Magazine 47(1), 112–120 (2009)
11. Wireshark – The Network Protocol Analyzer, http://www.wireshark.org/

On the Performance of Static Inter-cell Interference Coordination in Realistic Cellular Layouts*

David González G, Mario García-Lozano, Silvia Ruiz, and Joan Olmos

Universitat Politecnica de Catalunya (UPC),
C/ Esteve Terradas, 7 - 08860 Castelldefels, Spain
{david.gonzalez.gonzalez,mario.garcia-lozano}@upc.edu

Abstract. Effective interference management has been recognized by the industry and standardization bodies as a key enabler for 4G systems. This work is about static Inter-Cell Interference Coordination for OFDMA based cellular networks such as LTE. The majority of previous ICIC studies, both theoretical and simulation-based, have been conducted considering synthetic and/or small cellular layouts. In this work, the performance of static ICIC strategies in non-regular cellular layout is studied introducing some related RRM functions in the methodology. The overall performance assessment gives special attention to the efficiency vs. fairness tradeoff and the elements associated to it. Results show that the design of suitable and effective ICIC schemes for realistic cellular networks can not be done by simply extending classical approaches.

Keywords: Long Term Evolution, Inter-cell Interference Coordination, Radio Resource Management, Soft and Fractional frequency reuse.

1 Introduction

The International Telecommunications Union - Radiocommunication Sector (ITU-R) has defined a set of features that must be fullfilled by the so called International Mobile Telecommunications-Advanced (IMT-A) systems. Broadly speaking, these systems must be able to support high-quality mobile multimedia applications and fullfill the evolving users' needs [1]. Consequently, mobile operators are optimizing and upgrading their networks according to the evolution of the most popular technologies: the Long Term Evolution (LTE) [2] and WiMAX [3]. In particular, LTE has been described as a 3.9G (beyond 3G but pre-4G) technology since its first release LTE does not meet IMT-advanced requirements for 4G. However, the 3rd Generation Partnership Project (3GPP) is currently developing LTE Advanced which is a preliminary mobile communication standard, formally submitted as a 4G system candidate to ITU as a major enhancement of the LTE standard [4,5,6]. The target of 3GPP LTE Advanced is to reach and surpass ITU requirements.

* This work has been funded through the project TEC2008-06817-C02-02 (Spanish Industry Ministry).

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 163–176, 2011.

LTE (and WiMAX) employs Orthogonal Frequency Division Multiple Access (OFDMA) as access technology for the downlink [7] mainly due to its flexibility for resource allocation and because OFDMA provides intrinsic orthogonality to the users within the cell, which translates into a almost null level of intra-cell interference. Therefore, inter-cell interference is the limiting factor when high reuse levels (to achieve higher spectral efficiency) are intended. On the other hand, the new requirements for IMT-A include delivering higher peak rates to support advanced services (up to 1 Gbps for low mobility) and enhanced and uniform levels of quality of service within the cell area [8]. Nevertheless, from the network perspective, this is not a trivial task since users far away from their serving access point, typically perceive a significant amount of inter cell interference. Then, as a direct consequence of this situation, the fairness among the Quality of Service of users is jeopardized.

Given this, initiatives and proposals have been formulated within the 3GPP to cope with inter-cell interference. In particular, three main strategies [9,10] are proposed : Inter-cell Interference (a) Coordination (ICIC) (b) randomization and (c) cancellation. Although, inter-cell interference randomization [11], and cancellation [12] have received some attention, ICIC has been the field in which more contributions are being done, and it has been idenfied as a key element by the industry [13,14], and the research community [15,16]. The so called *soft* and *fractional* frequency reuse schemes (SFR and FFR respectively), have been widely studied. Reference works are: [17,18,19,20,21]. Although, the scope and manner in which the authors address the subject is varied, there are some similarities. Most of the existing literature addresses the ICIC issue by means of synthetic (and very often small) cellular layouts. The extention of these results to realistics scenarios is questionable mainly due to two reasons: the inter-cell interference is not uniform, this means that not all the cells receive the same amount of interference, as it does happen in a perfectly geometric layout. Second, when the effects of inter-cell interference are studied, the scenario must be large enough to asure that a at least 3 or more interferer tiers are considered. The latter is especially important in OFDMA networks taking into account that the wireless channel is frequency selective.

To the best of the authors knowledge, only very recent works [22] have addressed ICIC by considering large scale/realistic scenarios. In this work, the authors claim that for real networks with an irregular coverage pattern, no simple reuse scheme can be applied in a straighforward way. In this excelent contribution, the authors estimate the average throughput map over the entire service area by considering only large scale fadings effects. This approach allows a cell edge performance assessment without significant complexity involved.

In this paper, we do consider a large scale/realistic network as in [22] but with some important differences. First, from the system model perspective, we have evaluated the ICIC gain considering the constrains associated to the frequency domain scheduling (proportional fair discipline is considered) and the effect of the Adaptive Modulation and Coding (AMC). We strongly agree with the authors in [23] in that *the impact of ICIC on the overall system throughput* must

be analyzed through a model in which the interactions with additional Radio Resource Management (RRM) functions, such as scheduling, adaptive modulation and coding and power control are captured. With this, the analysis is more precise in the sense that also takes the frequency selectiveness of the channel into account. Second, from the performance evaluation point of view, not only system oriented metrics are considered, but also user oriented ones. This is important since improving fairness while keeping spectrum efficiency as high as possible is the main target of ICIC strategies.

In this manner, the contribution of this paper is a comparative analysis of several static ICIC strategies in a very realistic test bench in which the cross effects between ICIC and the additional RRM functions are weighed up. Results not only quantify ICIC gains by means of a comprehensive set of performance metrics but also provide hints about where to go in future studies.

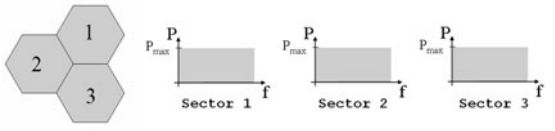
The paper is organized as follows, Section 2 introduces the ICIC schemes considered. Next, Section 3 includes the description of the simulation scenarios, the system model, the performance metrics and the particular configuration we have used for each scheme. The analysis of results, taken into account existing/related results from previous contributions, is presented in Section 4. Finally, conclusions and additional remarks close the paper in Section 5.

2 Description of Static ICIC Strategies

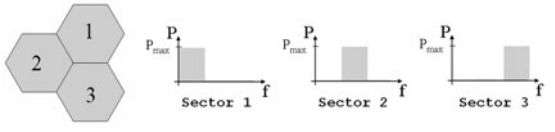
In this section, we present a detailed description of the static ICIC strategies considered for this study. In particular, 4 different schemes with different values of frequency reuse (FR) have been selected. The two first schemes, the so called full frequency reuse (FR=1) and fixed reuse 3 (FR=3), were selected as they represent the benchmark in terms of spectral efficiency and fairness respectively. Thus, a more coherent performance assessment of the two properly said ICIC strategies, *Soft Frequency Reuse* (SFR) and *Fractional Frequency Reuse* (FFR) can be done. These two strategies have been shown [24,25] to be intermediate points between FR=1 and FR=3. In addition, these strategies have been widely studied as ICIC schemes in the vast majority of contributions based on synthetic/theoretical scenarios. Therefore, the selection of these strategies is the natural choice according to our target. A generic representation of the selected static ICIC schemes is shown in Fig. 1.

In static ICIC schemes, the resources (bandwidth and power) allocated to each cell do not change over time. Then, each cell uses those resources autonomously and according to the rest of RRM functions. Nevertheless, depending on such *resource-to-cell allocation*, more or less freedom is left to finally pair the resources to the users.

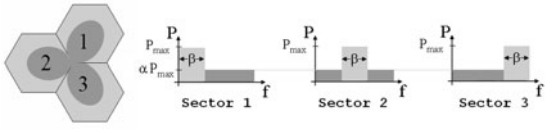
ICIC S1: Static ICIC scheme S1 is also known as full frequency reuse, since reuse factor 1 is aplied to the whole network. In this case, there is not constraint at all on the usage of resources within each cell. This scheme is attractive since it has been shown in [15] that it provides the best overall spectral efficiency.



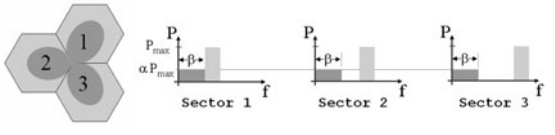
(a) ICIC S1.



(b) ICIC S2.



(c) ICIC S3.



(d) ICIC S4.

Fig. 1. Generic power profiles in static ICIC schemes

Nevertheless, users close to the cell edge experience a significant amount of inter-cell interference.

ICIC S2: Employing reuse factor 3 is the traditional choise followed by network operators in tri-sectorial deployments. In this case, the levels of inter-cell interference experienced by cell edge users are significantly reduced at expense of the overall network efficiency. Similarly to the previous case, cells employ their resources without restrictions.

ICIC S3: This strategy corresponds to soft frequency reuse. This approach can be considered as an intermediate point between the two previous strategies in the sense that reuse factor 3 is applied to the cell edge users while central users do more aggressive usage of the spectrum. Soft frequency reuse implies the need to classify users within each cell. The criterion is often based on the average channel quality [26], [27]. Two possible approaches can be taken into account:

1. *Class Proportionality*: SINR thresholds are selected so that each class has the same average number of users.
2. *Bandwidth Proportionality*: The threshold guarantees that the number of users is proportional to its allocated bandwidth.

Table 1. Simulation scenarios details

Scenario	Cells	Area [Km^2]	Density [$Cells/Km^2$]	$SINR_{TH}$ [dB]
A	171	33.2	5.15	1.85
B	42	32.9	1.28	2.50

In this study, only class proportionality is considered. The reason is twofold. First, study the impact of the classification thresholds on ICIC performance is out of the scope of this work, hence keep fixed this degree of fredom is a necessary condition to get valid results. Second, the particular choice of class proportionality is due to the fact that it is an approach commonly employed in previous contributions [28,26] dealing with static ICIC, thereby facilitating comparisson and analysis.

It is worth to note that while central users receive inter-cell interference of type *inter-class* (coming from users of different class) and *intra-class* (coming for users of the same class), cell edge ones only receive inter-class interference. Finally, the amount of interference received by the cell edge users and the their bandwidth size are controlled by the parameters α and β respectively.

ICIC S4: As in ICIC S3, two different classes are considered, nevertheless the main difference in this case is that the inter-class interference is completely removed, i.e. each class has exclusive use of its bandwidth. This is important because the performance in terms of throughput and fairness becomes independent of α since the SINR does not depend on the transmitted power (equal for all cells) as long as the inter-cell interference level is significantly higher than the noise floor. The parameter β controls the width of the band allocated to central users, hence it also determines the bandwidth available for outer ones.

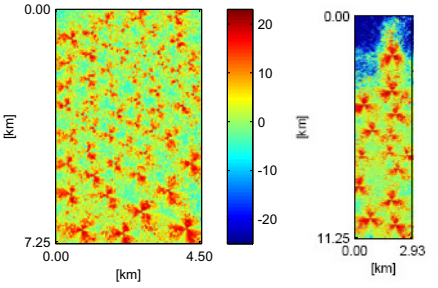
3 Experiments Description

In this section, a complete description of the experiments is provided. The different sub-sections explain the overall methodology.

3.1 Simulation Scenario

The simulated scenario is a realistic one covering the city of Vienna and its surroundings. The digital elevation model, system layout and propagation data has been obtained from the MORANS initiative [29]. This activity was framed within the European COST 273 Action and aimed at providing common system simulation environments so that different researchers can compare results.

In particular, two sub-areas having different cell densities were selected for comparisson purposes. Figure 2 depics the resulting average SINR map for both zones. Additional details are shown in Table 1.



(a) Scenario A (b) Scenario B

Fig. 2. Realistic simulation scenarios

3.2 System Model and Methodology

Additionally to the cellular layouts described previously, the cells' downlink configuration follow the setting established by the LTE standard [7,30,31]. Both the simulation scenario and the LTE's OFDMA setting complete the system model in which users are randomly allocated. For each scenario, an average density of 50 users per cell was considered. The traffic model considered for all users is full buffer. The choice of this model, from the ICIC performance assessment perspective, can be considered as a worst case since it assures that each cell will transmit power over its whole available bandwidth leading then to the worst situation in terms of inter-cell interference. The implementation was done by means of an OFDMA system level simulation platform that has been developed in C++. The link-to-system level interface largely follows the guidelines given by [32].

Specifically, the system has 100 physical resource blocks (PRB) available for the users (18 MHz, 1200 sub-carriers of 15 kHz). Note that a Physical Resource Block (PRB) is the minimum bandwidth the scheduler can assign to one single user. Transmission time intervals of 1 ms containing 10 OFDMA symbols are considered. The total available power at each cell is 43 dBm and sum power condition is always kept. ITU Extended Typical Urban (ETU) have been considered as channel model. 8 dB log-normal shadowing is applied following the model proposed in [33] with a correlation coefficient between cells equal to 0.5. It is important to stress that achievable rates were computed taking into account the instantaneous channel conditions (including the frequency selectiveness of the channel) and according to the adaptive modulation and coding used in LTE, as specified in [34]. This mapping has been done using the link abstraction model based in mutual information at modulation symbol level [35], which outperforms the classic Effective Exponential SINR model because it is able to predict the BLER with higher accuracy, particularly for higher order modulations, such as 64-QAM. Additionaly, Proportional Fair Scheduling is autonomously executed at each cell to make the final pairing of resources to users according to the following expression:

$$m^*(t,n) = \underset{m}{\operatorname{argmax}} \frac{T_{m,n}(t)}{\sum_{k=0}^{t-1} \sum_n T_{m,n}(k)} \qquad \forall n \in N_L \tag{1}$$

where $T_{m,n}(t)$ is the achievable throughput if the PRB n is assigned to the user m at time t. N_L is the set of PRBs available at cell L.

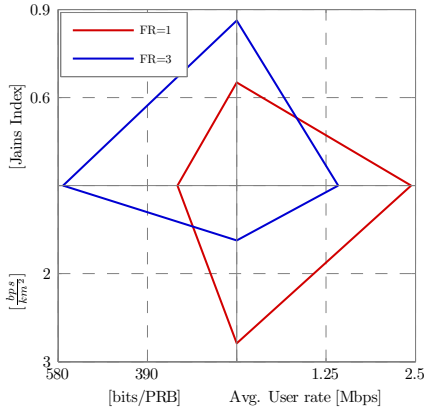
To conclude this section, the description of the experiments is provided. As stated earlier, two different scenarios (see Table 1) and 4 different static ICIC strategies (see Fig. 1) were considered. Specifically for ICIC schemes S3 and S4, the impact of the parameters α and β on the overall performance was also studied. In particular, for SFR, a linear variation of α (0.2 0.4 0.6 0.8) was considered while two values of β (0.16 0.33) were also taken into account. Note that for the tri-sectorial arrangement, β must be smaller than 0.33 when SFR (as in 1c) is applied. For FFR, the values of α were kept as in SFR but an additional value of β (0.67) was considered. The choice of these particular values is to keep consistency with previous contributions in which similar variations of these parameters was considered to assess the performance of static ICIC strategies in synthetic scenarios and so, obtain reasonable conclusions based on our results.

3.3 Performance Metrics

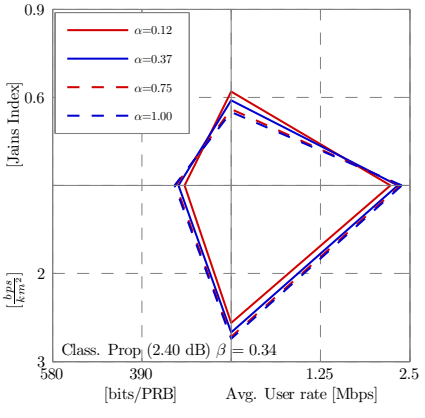
As it was commented previously, the analysis is mainly focused on the efficiency vs. fairness tradeoff. Neverthelss, additional metrics were also considered to better understand the whole network behaviour. In particular, the set of metrics includes: a fairness measure, the Jain's index [36], the system spectral efficiency per area unit ($\frac{bps}{Hz \cdot km^2}$), the average user rate (Mbps) and the average number of bits per PRB ($\frac{bits}{PRB}$) to take into account the effectiveness in the resources usage. All these metrics were computed based on Monte Carlo simulations.

4 Analysis of Numerical Results

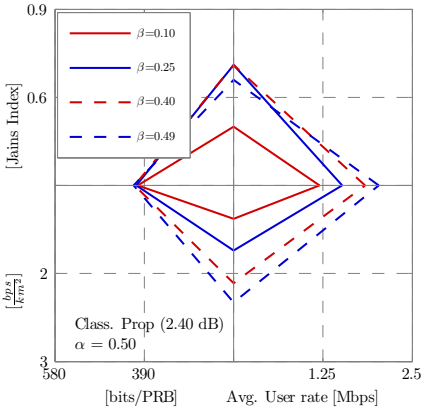
In this section, the results obtained for the experiments previously described are presented. It is important to recall that the main target of this study is to assess the performance of static ICIC strategies in realistic scenarios and compare them with the ones already reported for synthetic cellular layouts. To do this, we will take as reference results and conclusions from some previous contributions considering synthetic scenarios [28,37,17]. In particular, results taken from [28] are illustrated in Fig 3 as main reference. By comparing Fig 3a with Figs 3b and 3c, it is clear that schemes S1 and S2 are the benchmarks in terms of spectral efficiency and fairness respectively. This is a well known result in the context of static ICIC in synthetic cellular layouts. On the other hand, the results corresponding to this work (for realistic cellular layouts) are presented in Figs 4, 5 and 6. These figures correspond to the cases of FR=1/FR=3, SFR and FFR (for both scenario A and B) respectively.



(a) ICIC S1/S2: FR=1 FR=3



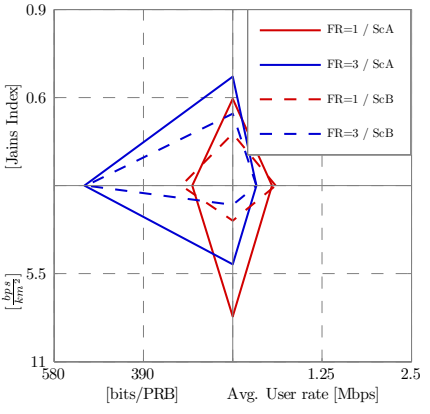
(b) ICIC S3/S4: SFR



(c) ICIC S3/S4: FFR

Fig. 3. Reference results from synthetic scenarios

Looking at Fig 4, the well known rule of thumb [27] that states that S1 offers the highest spectral efficiency with poor fairness, while S2 boosts the opposite also applies in realistic networks. The main difference appears when quantifying this losses/gains. Thus, for example the throughput gain in scenario B is clearly smaller than in A or synthetic scenarios. On the other hand, the fairness gain is proportionally smaller in the scenario A than in scneario B. Comparing the results of S1 and S2 in synthetic and realistic scenarios (beyond the absolute magnitudes), the impact of such schemes on the performance metrics is clearly different for the different scenarios, especially for the case of fairness and average users rate although the overall behaviour still holds in realistic networks in the sense that S1 favorates the efficiency while S3 favorates the fairness.



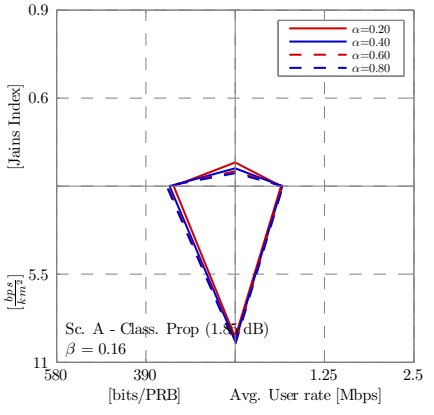
(a) ICIC S1/S2: FR=1 FR=3

Fig. 4. S1/S2: FR=1 FR=3 Sc. A/B

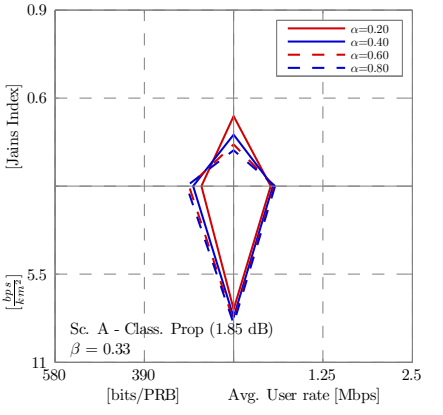
Figure 5 depicts the results corresonding to SFR in which different values of α were evaluated for two different values of β for both scenario A and B. Again, what we can see, by considering the case of β=0.34 is that the impact of the parameter α on the different metrics is different in realistic scenarios compared with the result shown in Fig 3b for synthetic cellular layouts. Nevertheless, the overall behaviour for SFR is, in general terms, similar to the one observed in synthetic scenarios: (a) the greater the value of α the greater the spectral efficiency and (b) the lower the value of α the greater the value of fairness. However, one common aspect to both realistic scenarios is that a greater value of β favorates the sensitivity of the metrics to the values of α and so an easier tunning can be achieved. This is expected as β controls the portion of bandwidth allocated to the exterior users.

In Fig 6, the perfoemance of S4 is shown for both scenario A and B. Note that in both cases only one value of α is considered (0.20). The reason is that, similarly to synthetic scenarios, in fractional frequency reuse schemes, the overall performance becomes independent of α. Thus, results shows an interesting situation. In both scenario A and B the value of β is proportional to the spectral efficiency, but the fairness shows a maximum point for a intermediate point of β (different for each scenario). Also, we can note that the sensitivity of the performance metrics to the value of β is quite different in realistic networks due to the irregular geometry. All this behaviour is expected since β has to do with the bandwidth sharing between classes, and the definition of such classes implies setting up the SINR thresholds which in turn depends strongly on the layout under consideration. So, it follows that the choice of an optimal value for β depends on the network under consideration.

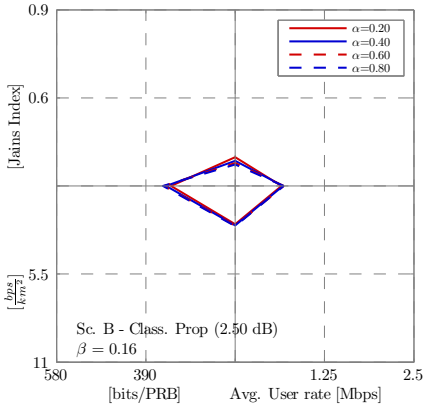
From the previous observations, it is clear that the optimal setting in terms of ICIC is particular to each geometry and that the best performance can not be



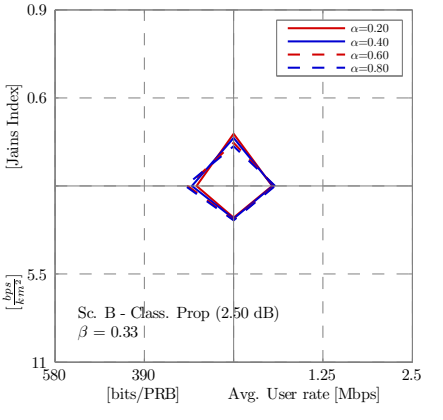
(a) ICIC S3: SFR in Sc. A β=0.16



(b) ICIC S3: SFR in Sc. A β=0.33



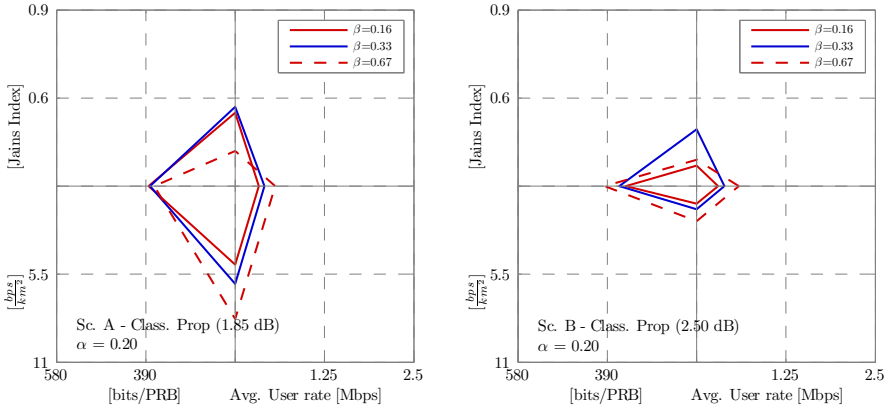
(c) ICIC S3: SFR in Sc. B β=0.16



(d) ICIC S3: SFR in Sc. B β=0.33

Fig. 5. S3: SFR in scenarios A and B

obtained by simply applying the traditional/homogeneous ICIC schemes (based on identical power mask for each cell). Even in situations in which the overall system performance seems to be good, an static/homogeneous ICIC pattern could penalize cells receiving more interference due to the irregular layout. Thus, although there are some important similarities between the results observed in synthetic scenarios with respect to realistic ones, there are also important differences. The first one, from the whole system performance perspective is that the optimum tunning is clearly network dependent and the second one, from the individual cells perspective, is that irregular cellular layouts leads inevitably to non-regular bandwidth allocations.



(a) ICIC S3: FFR in Sc. A α=0.20 (b) ICIC S3: FFR in Sc. B α=0.20

Fig. 6. S3: FFR in scenarios A and B

5 Conclusions and Future Work

A fair comparisson among different static ICIC strategies has been presented in this work. The results were obtained by running simulations over realistic cellular layouts and by considering additional RRM functions. The overall study was focused on the assessment of the performance of static ICIC strategies.

The main conclussions are summarized as follows:

- Classical ICIC strategies cannot be applied in a straighforward manner to non-regular cellular layouts claiming at their optimality at the same time. The performance of such strategies is different from one network to another. Tuning the network in realistic deployments must take into account several factors such as changing network load, traffic and mobility patterns, the local geometry and the associated/available network functions.
- The study of ICIC on non-regular cellular layouts open a promizing research line as it raises interesting open problems. A sufficiently generic framework that can be extended to realistic deployments has not been formulated. The relationship between ICIC and other RRM functions in the context of non-regular layouts has not been modeled neither. Because of this, the study of ICIC should not be addressed without the knowledge of the whole network picture.
- At this point, logical extensions to this work have been identified: (a) How to further exploit the connections between ICIC and additional RRM functions and network-dependent features appear as a natural direction to follow and (b) time varying network conditions could serve as a basis for designing of more flexible (but simpler) ICIC schemes. These schemes must be supported

by the limited amount of ICIC-oriented mechanisms available in the standards. In this sense, results clearly point towards the design of semi-static and dynamic ICIC schemes.

References

1. Radiocommunication Sector: Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000. In: International Telecommunication Union (ITU), M.1645 (2008)
2. Group Radio Access Network: Overall description, Stage 2. 3rd Generation Partnership Project (3GPP). (Mar 2010) TS 36.300, V9.3.0, (Release 9).
3. WiMAX Forum: Mobile WiMAX – Part I: A Technical Overview and Performance Evaluation (February 2006)
4. Parkvall, S., Dahlman, E., Furuskar, A., Jading, Y., Olsson, M., Wanstedt, S., Zangi, K.: LTE-Advanced - Evolving LTE towards IMT-Advanced. In: IEEE 68th of Vehicular Technology Conference, VTC 2008-Fall, pp. 1–5 (October 2008)
5. Ericsson: LTE Performance and IMT-Advanced Requirements. 3rd Generation Partnership Project (3GPP) (May 2009); R1-092022, TSG RAN WG1 Meeting #57: San Francisco, USA
6. Group Radio Access Network: Requirements for Further Advancements for E-UTRA (LTE-Advanced). 3rd Generation Partnership Project (3GPP) (June 2008); TR 36.913, V8.0.0 (Release 8)
7. Group Radio Access Network: LTE Physical Layer: General Description. 3rd Generation Partnership Project (3GPP) (December 2008); TS 36.201, V8.2.0, (Release 8)
8. Radiocommunication Sector: Guidelines for evaluation of radio interface technologies for IMT-Advanced. In: International Telecommunication Union (ITU), M.2135 (2008)
9. Instruments, T.: Performance of Inter-Cell Interference Mitigation with Semi-Static Frequency Planning. 3rd Generation Partnership Project (3GPP) (January 2006); R1-060067, TSG RAN WG1 Meeting #43: Helsinki, Finland
10. Nokia: A proposal for LTE TDD Uplink Multi-TTI Scheduling. 3rd Generation Partnership Project (3GPP) (March 2008); R1-081450, TSG RAN WG1 Meeting #52bis: Shenzhen, China
11. Hu, W., Willkomm, D., Abusubaih, M., Gross, J., Vlantis, G., Gerla, M., Wolisz, A.: Dynamic Frequency Hopping Communities for Efficient IEEE 802.22 Operation. IEEE Communications Magazine 45(5), 2393–2409 (2007)
12. Ping, L., Liu, L., Leung, W.: A simple approach to near-optimal multiuser detection: interleave-division multiple-access. In: Wireless Communications and Networking, WCNC 2003, vol. 1, pp. 391–396. IEEE, Los Alamitos (2003)
13. Alcatel: Interference Coordination in new OFDM DL air interface. 3rd Generation Partnership Project (3GPP) (May 2005); R1-050407, TSG RAN WG1 Meeting #41: Athens, Greece
14. Electronics, L.: Further aspects of interference coordination. 3rd Generation Partnership Project (3GPP) (January 2006); R1-060053, TSG RAN WG1 Meeting #43: Helsinki, Finland
15. Simonsson, A.: Frequency Reuse and Intercell Interference Coordination In E-UTRA. In: IEEE 65th of Vehicular Technology Conference, VTC 2007-Spring, pp. 3091–3095 (April 2007)

16. Necker, M.: Local Interference Coordination in Cellular OFDMA Networks. In: 2007 IEEE 66th of Vehicular Technology Conference, VTC 2007 Fall, pp. 1741–1746 (September 2007)
17. Huawei: Soft Frequency Reuse Scheme for UTRAN LTE. 3rd Generation Partnership Project (3GPP) (May 2005); R1-050507, TSG RAN WG1 Meeting #41: Athens, Greece
18. Huawei: Further Analysis of Soft Frequency Reuse Scheme. 3rd Generation Partnership Project (3GPP) (September 2005); R1-050841, TSG RAN WG1 Meeting #42: London, UK
19. Necker, M.: A Graph-Based Scheme for Distributed Interference Coordination in Cellular OFDMA Networks. In: Vehicular Technology Conference, VTC Spring 2008, pp. 713–718. IEEE, Los Alamitos (2008)
20. Dong, K., et al.: A Distributed Inter-Cell Interference Coordination Scheme in Downlink Multicell OFDMA Systems. In: 2010 7th IEEE of Consumer Communications and Networking Conference (CCNC), pp. 1–5 (2010)
21. Ali, S., Leung, V.: Dynamic Frequency Allocation in Fractional Frequency Reused OFDMA Networks. In: GLOBECOM Workshops, pp. 824–829. IEEE, Los Alamitos (2008)
22. Chen, L., Yuan, D.: Soft frequency reuse in large networks with irregular cell pattern: How much gain to expect? In: 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1467–1471 (2009)
23. Racz, A., Reider, N., Fodor, G.: On the Impact of Inter-Cell Interference in LTE. In: Global Telecommunications Conference, IEEE GLOBECOM 2008, pp. 1–6. IEEE, Los Alamitos (2008)
24. Huawei: R1-050507: Soft Frequency Reuse Scheme for UTRAN LTE. 3GPP (May 2005); TSG RAN WG1 Meeting #41: Athens, Greece
25. Samsung: R1-051341: Flexible Fractional Frequency Reuse Approach. 3GPP (November 2005); TSG RAN WG1 Meeting #43: Seoul, Korea
26. Koutsimanis, C.: Intercell Interference Coordination Techniques for Multicell OFDMA Networks Supporting Narrow Band and Elastic Services. Master's thesis, Royal Institute of Technology (KTH) (May 2007)
27. Hernández, A., Guío, I., Valdovinos, A.: Radio resource allocation for interference management in mobile broadband OFDMA based networks. Wireless Communications and Mobile Computing 9999(9999), 1530–8669 (2009)
28. Gonzalez, D., Garcia-Lozano, M., Ruiz Boqué, S., Olmos, J.: Static Inter-Cell Interference Coordination Techniques for LTE Networks: A Fair Performance Assessment. In: Vinel, A., Bellalta, B., Sacchi, C., Lyakhov, A., Telek, M., Oliver, M. (eds.) MACOM 2010. LNCS, vol. 6235, pp. 211–222. Springer, Heidelberg (2010)
29. Verdone, R., Buehler, H., Cardona, N., Munna, A., Patelli, R., Ruiz, S., Grazioso, P., Zanella, A., Eisenblätter, A., Geerdes, H.: MORANS White Paper - Update. Technical Report available as TD(04)062, COST 273, Athens, Greece, January 26-28 (2004)
30. Group Radio Access Network: Physical Channels and Modulation. 3rd Generation Partnership Project (3GPP) (December 2008); TS 36.211 v8.5.0 (Release 8)
31. Group Radio Access Network: Multiplexing and Channel Coding. 3rd Generation Partnership Project (3GPP) (December 2008); TS 36.212 v8.5.1 (Release 8)
32. Brueninghaus, K., Astely, D., Salzer, T., Visuri, S., Alexiou, A., Karger, S., Seraji, G.A.: Link performance models for system level simulations of broadband radio access systems. In: IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC 2005, vol. 4, pp. 2306–2311 (November 2005)

33. Fraile, R., Lázaro, O., Cardona, N.: Two Dimensional Shadowing Model. Technical Report available as TD(03)171, COST 273, Prague, Czech Rep., September 24-26 (2003)
34. Olmos, J., Serra, A., Ruiz, S., García-Lozano, M., González, D.: Exponential Effective SIR Metric for LTE Downlink. In: Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Comm. (PIMRC 2009), Tokyo, Japan, September 13-16 (2009)
35. Zheng, H., Wu, M., Choi, Y., Himayat, N., Zhang, J., Zhang, S.: Link Performance Abstraction for ML Receivers Based on RBIR Metrics. Technical Report C802.16m-08, IEEE (2008)
36. Jain, R.: The Art of Computer Systems Performance Analysis, 1st edn. John Wiley & Sons, New York (1991)
37. Gonzalez, D., Ruiz, S., Garcia-Lozano, M., Olmos, J., Serra, A.: System level evaluation of LTE networks with semidistributed intercell interference coordination. In: 2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications, pp. 1497–1501 (2009)

Location-Based Ubiquitous Context Exchange in Mobile Environments

Stefan Forsström, Victor Kardeby, Jamie Walters, and Theo Kanter

Mid Sweden University, Sundsvall 851 70, Sweden
{stefan.forsstrom,victor.kardeby,jamie.walters,theo.kanter}@miun.se

Abstract. Context-aware applications and services require ubiquitous access to context information of users. The limited scalability of centralized servers used in the provisioning of context information mandates the search for scalable peer-to-peer protocols. Furthermore, unnecessary signaling must be avoided in large-scale context networks, when location-based services only require nodes in a certain area with which to communicate context. To this end, we propose a lightweight model for composing and maintaining unstructured location-scoped networks of peer-to-peer nodes, which gossips in order to ensure quality of service for each user. The model is implemented in a prototype application running in a mobile environment, which is evaluated with respect to real-time properties. This model can also be extended to include more context dimensions, other than location.

Keywords: Ubiquity, location-based, context exchange, mobile system.

1 Introduction

This paper investigates whether a scalable user centric self organizing context delivery system provides a better solution for location-scoped acquisition of context information in real-time. This paper proposes as such an necessary improvement to previous research [1] regarding structured dissemination of context information to context-aware applications. Scalable personal networks can provide a solid model for ad hoc context-aware applications which can be used in real-time applications. The need to derive context in real-time drives the ability to create applications that can provide real-time user interaction in response to the dynamic nature of the users context. The growth of social communities on the Internet has raised the interest in social infrastructures in support of collaborative computing both mobile and stationary. The creation of services that can utilize ad hoc frameworks include, but is not restricted to, emergency systems, warning services, social gatherings, and general information sharing.

Visitors attending a museum in a city might be able to capture a small percentage of the multitude of photo opportunities that are presented at each time. However, collaboratively they provide an extensive overview of an environment. Each visitor exists as an ubiquitous island of information and potential social

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 177–187, 2011.

collaboration waiting to be exploited. Current approaches such as [2,3,4,5] use solutions that build systems around completely decentralized solutions which are implemented as general context exchange solutions, while aiming for minimal hop length towards each peers position. However these solutions still maintain a logarithmic hop count between users as the system grows in scale. Other solutions such as [6] use a centralized infrastructure, which has a constant low hop count, but does not scale well as the system increases.

The reliability of mobile communication cannot be guaranteed and subsequently, neither can dependent services. Decentralized solutions provide an alternative to this problem. Solutions such as [7] and [8] require that the mobile devices contain radios with broadcast capabilities and that there exists some indeterminate means of locating an initial neighboring user. There also exists solutions based on short range communication, such as [9] and [10] solutions, which mostly is based on multi-hop opportunistic communication via short range radios. However as mobile broadband infrastructure has near ubiquitous coverage, wide-area wireless technologies have become mainstream for providing network access. In face of these trends, the need for structuring of data dissemination in the network gains in importance, which is explored in this paper.

Therefore, there exists a need for a simple ad-hoc location-based protocol capable of solving the initial node problem as well as the radio beacon problem. In this paper we present such a model for discovering nodes and maintaining a personal peer-to-peer network enabling a platform for constructing location-based social networks and information exchange.

While we concur with [11] concerning what constitutes context information, our solution is focused on location, and employs a simple structure predicated on maintaining a single hop between a user and its interests, and remain simple enough to be deployed on resource constrained mobile devices.

2 Context Services

Context information is specific information tied to an entity, such as a user or location, which explains the surrounding conditions of that specific entity and is often acquired by using sensors systems such as GPS, or by manual input such as personal preferences. The concept of context networks has been developed to administer intelligent exchange of this context information between users. Three kinds of context networks exist today, centralized solutions, distributed solutions employing structured architectures, and distributed solutions that employ unstructured architectures.

Centralized solutions such as 3GPP's IMS Presence [12], build on using the SIP protocol, to establish sessions between end users. The SIP protocol is simple and real-time capable, but the architecture for the underlying IMS system is cumbersome and was not created for real-time context-aware applications, because of its control over the data flows and the presence architecture. Structured distributed systems such as DCXP [13] utilize a strict structure and layout of all nodes. Which is aimed towards scalability and real-time applications, but

require overhead messaging to maintain the underlying DHT. In addition to this, there is unstructured distributed systems such as Gnutella [14], which employs unstructured layouts which organizes nodes into an unoptimized, but much simpler topology. Requiring much less overhead to maintain the network, but has longer hop length.

The area of context networks have been revisited many times in research, by the CONTEXT [15], SenseWeb [16] and MobiLife [17] projects. Where the CONTEXT, and Mobilife project aimed towards using a centralized approach, using the SIP protocol in similary to 3GPP IMS. SenseWeb on the other hand uses an web-service oriented approach and the SOA oriented architectures used in that area, however still using a centralized approach.

Most research has been conducted in node traversal and communication optimization, which detracts from the problems associated with running the applications on mobile and limited devices, which have highly volatile connections, long communication delays, real-time demands, and a massive user base.

2.1 Context Architectures

Context-aware applications that can be enabled by personal self organizing networks are often aimed towards location-based services, such as group conversations (among people on a train) and social networks (finding people with similar interest), and requests to enable more location-based public services (location-based tourist information). However the possibilities of services based on this approach is almost endless, ranging from simple data sharing, such as location-based instant messaging, to complex services such as matchmaking services and location-based social networking systems employing multimedia delivery services.

General peer-to-peer networks do not natively support location-based services. To enable such services, message overhead is required in the form of flood searches or random walks. This mandates the need for an underlying structure which enables the location-based services natively and is lightweight enough to be run on a resource constrained device. Location-based grouping of information have existed for some time [18]. Such solutions point out the advantage gained by having a limited search area, resulting in decreased latency and bandwidth consumption.

Structured peer-to-peer networks, such as Chord [19], are often built using distributed hash tables and these networks enable many advantages such as optimized node searches. However structured networks, also introduce higher computational requirements for each node and increased signaling for maintaining the distributed hash table, especially with increased node volatility. Unstructured peer-to-peer networks are often flat architectures and without any globally structured connections among the nodes, which enables easy access for nodes to join and leave the network. Unstructured networks gain simplicity and low computational requirements, at the expense of bandwidth. Unstructured

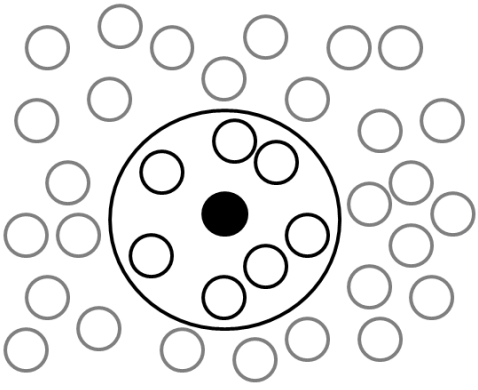


Fig. 1. Location-based personal network

peer-to-peer networks employ simple search algorithms, such as variants of flood or random walk searches. However such algorithms consume a considerable amount of network overhead even when optimized.

2.2 Context Proximity

Proximity can be defined as the physical or geographical nearness, i.e a kind of closeness or vicinity. However in regards to context, we want to utilize this concept of closeness to create the concept of context proximity. This concept means the proximity between users context, i.e. not only by location. For example, two people in two different towns, working at the same company, are not in geographical proximity to each other, but their context is in proximity, since they both working for the same employer. This proximity in regards to context has been seen in [20], although at that point it was quite simple. Although the focus in this paper is on location, the idea of location-based context proximity is used to create a form of interest between users.

3 The Model

This paper proposes a peer to peer model with limited signaling to enable a solution that natively support location-based services. The basis of the prototype comes from the fact that each node creates a personal network of nodes which are within its context proximity by gossiping with its peers. The interest area moves with the user, constantly changing when other users come within range, see Fig. 1 for an illustration of the context proximity interest area. The area can also be dynamically resized, based on the quality of service required by the applications. This quality of service can be determined on a per application basis, by choosing a radius of interest. This choice of quality of service comes from the fact that some applications might want to limit their interest area to a bus or

a train on which the user is currently traveling, while some applications might require a wider interest area such as an entire city block in order to provide an adequate service.

The network is built up by having one single persistent bootstrap node, called tracker. The bootstrap procedure is simple but centralized, therefore the amount of bootstrapping should be kept to a minimum. A node bootstrap by contacting the tracker and asking for a list of nodes, which contain interesting nodes for the bootstrapping node. As the node acquires this list it can start its gossiping, and given that the other nodes can provide adequate gossip information, the node will not need to bootstrap ever again. One unfortunate disadvantage of the centralized tracker, other than the obvious scalability problem, is that the tracker is required to be updated with context information for it to provide reliable bootstrapping information. Therefore the nodes has to contact the tracker once in a while to provide it with the nodes current context.

The nodes communicate and exchanges information with other nodes via wide area Internet access. Therefore, after the bootstrap procedure, they continually gossip with each other, exchanging context information. The information exchanged in the gossip is kept limited, as the request only includes a node's current context, and the response only contains a list of other nodes which might be interesting for the requester, in similarity to the bootstrap procedure. Although the gossiping procedure is simple, it requires some management and evaluation of the acquired information.

In detail, the gossip procedure is conducted in the following manner. A node evaluates it's own list of known nodes, chooses a remote node and gossip information. The answer from the remote node will update the list of current known nodes. However, each node also has to evaluate this list of nodes, removing obsolete nodes or uninteresting ones. Such a process is required due to the volatility of context with respect to location and size of the interest areas. Therefore, all nodes continuously evaluate their list of known nodes, maintaining an updated list of other known nodes. They also continually review this own list, removing nodes which are no longer required.

The created network is dynamic, since the nodes in the interest area are constantly changing to reflect their changes in context in the real world. It is also ubiquitous as it runs on mobile devices which can contact each other using mobile Internet access. The network is also highly volatile, which can be observed when a node leaves the network and then later rejoins, then the previous interest area is almost certainly outdated and of no use. This is because the personal network composition is never in a stable state for an extended period, since the nodes context is always changing.

The key point about this model is that there exists no need to maintain overlays and expensive routing protocols which can undermine the real-time characteristics. Instead all nodes relevant to a user are kept within one hop in the overlay. This differs from other solutions, because of the distributed approach, the gossiping method, the real-time properties, and the way context is handled.

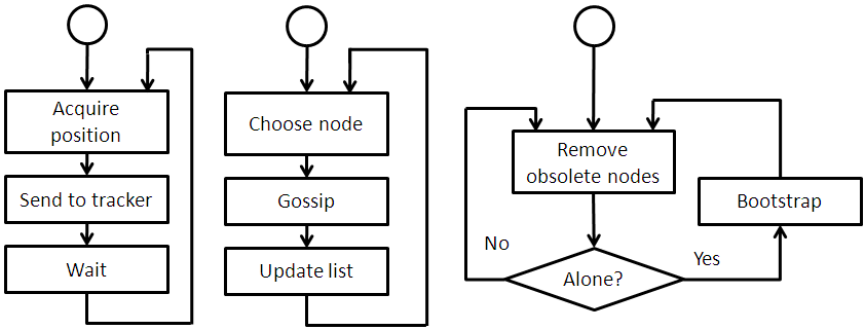


Fig. 2. Flowchart over the prototype

4 Prototype

A prototype has been developed, which resides on an implementation of the proposed model. It is however limited to only consider location as context, because of the difficulty in determining context proximity. The prototype is developed as a mobile application on the Android platform, but can also be run on a home computer. In addition to this, a prototype implementation of the tracker was implemented for home or server computers. The purpose of this application is to act as a proof-of-concept, that location-based peer-to-peer networks can run and provide suitable service even on limited mobile devices. The applications was also used to perform preliminary results and evaluations on how such a system will perform over mobile Internet and how it scales with multiple entities.

4.1 Implementation

Our prototype uses three independent algorithms, see Fig. 2. The first one, continuously updates the location of the user from available sensors, and updates it to the tracker. The second algorithm contacts a random peer chosen from a list of active nodes and queries that peer for more suitable nodes. The third algorithm manages the list of suitable nodes and removes obsolete nodes which are no longer within context proximity. It also manages the fall back situation where it will bootstrap again, which will happen if the user finds itself alone and unable to gossip.

In detail, the first algorithm acquires the GPS position of the mobile phone, and stores the location for later usage in the other algorithms, while also sending it to the tracker. The tracker will store the location, and use it when other users want to bootstrap. This algorithm is interrupt based, because of the location events created by the GPS. However the time between sending the location to the tracker can be varied, depending on the sought after quality of service. Although for the evaluation, the GPS was disabled and replaced with manual input for movement.

The second algorithm is the primary gossiping procedure. Which relies on a local list of current nodes which are in context proximity, which in the prototype is only based on physical location. The algorithm choose a random node in the list and tries to contact and gossip information with it. The request contains the current location of the user and the interest radius, which the user is interested in. As the remote node receives this information, it evaluates the request and compares to it's own local list. This process creates a new list of interesting nodes, including their location, which will be sent to the requesting user. The requesting user will receive this list and update it's own list, including the node's location. This algorithm can be performed as often as possible, depending on the quality of service required by the application.

The third algorithm relies on the local list of nodes and the current context of the user. This algorithm evaluate the context of all the nodes in the list, to determine if they still are in context proximity. As this prototype only takes into regards location as context, the evaluation for the context is quite simple. If a node is found to be outside of interest area of the user, it is removed from the list. However if the algorithm has removed all other nodes from the list, effectively making the user impossible to continue gossiping, the algorithm will contact the tracker again and bootstrap, to find new suitable nodes.

4.2 Preliminary Results

The application was deployed on Android mobile phones, but due to operator-side problems with peer-to-peer communication between mobile phones, the service runs best in emulated environments where all devices have public IP and without a firewall. Therefore the evaluation environment was setup with an active tracker with a public IP on the SUNET Internet backbone. In addition to this, a computer was setup with multiple nodes running, capable of operating up to 50 nodes at the same time. This computer operated these nodes just like a real node, and simulated their movement in the world. To this a single HTC Hero phone was added, a mobile Android device with access to the TeliaSonera 3G mobile Internet. This device runs the application as a real node, without any knowledge that the other nodes are simulated. Figure 3 shows a screenshot on how the application looks like, when running on the mobile phone. Where the node in the middle is oneself, which can be moved around on the screen in addition to the simulated movement of the other nodes.

The application also included a view with debug information with timers on how long the different actions in the algorithms have taken to be processed. These measurements was used to create an preliminary evaluation on how each of the algorithms operated. The actions which were available was bootstrap, position update, node management, and query. These evaluations are summarized in table 1, and was conducted in three different scenarios, with 10, 25, and 50 active nodes in the system. The values are measured in milliseconds using the current millis function in Java, and the values are averaged values over a small set of 10 samples each.

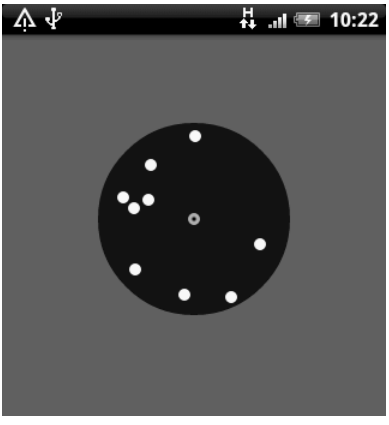


Fig. 3. Screenshot of the prototype on an Android phone

Table 1. Prototype response times

	10 nodes	**25 nodes**	**50 nodes**
Bootstrap	367ms	387ms	410ms
Position update	103ms	136ms	141ms
Node management	<1ms	<1ms	3ms
Query	310ms	313ms	335ms

Bootstrap is when a client bootstraps from the tracker. It is measured from the point when a peer begins to send a bootstrap request to the tracker, to the point in which it receives an answer, this time includes location evaluation on the tracker.

Position update is the operation performed by the clients when it sends its current location to the tracker. It is measured from the point that the node starts to send an update to the tracker, until it has received an acknowledgment that this position have been successfully updated on the tracker.

Node management is the evaluation that the client performs on the local list of nodes, which purpose is to remove obsolete nodes. It is the time it took for the client to go through the local list once.

Query is the gossiping operation, where a client asks another node for interesting nodes. It is measured from the point when the client starts to ask another node for other interesting nodes that is within vicinity, to the point in which it has received the list and updated it's own local list of interesting nodes.

4.3 Preliminary Analysis

With respect to the acquired values, one can observe that the preliminary response times is on par with normal TCP traffic over 3G Internet access, for the

values which include traffic over the mobile Internet connection. But in the steps which included contacting the tracker, i.e. *Position update* and *Bootstrap*, a noticeable scalability problem of the tracker is exposed. The small increase in *Node management* is probably accounted the increase in interesting nodes, because of larger set of available nodes. While the small increase in *Query* is probably accounted the simulation environment which is running multiple peer nodes on a single physical machine, in addition to the increased set of nodes.

Because of this, a more complete simulation environment is required, taking into account both the 3G latencies of the gossiping, and a more verifiable evaluation method. Such an environment can also include even more nodes and scalability evaluations.

But in comparison to other systems such as [2,5,3,4], whose approach follows the routing principle and therefore increasing the latency with the hop count, which in those cases increase logarithmic in regards to the amount of nodes. Our system maintains a single hop, and therefore the major drawback of the 3G mobile Internet system, the delay and latency, is circumvented.

5 Conclusions

This paper presented a peer to peer based model for organizing users into dynamic, unstructured, location-based per user unique groups. The network is ubiquitous and has highly volatile connections, but is lightweight enough to be handled by each user's mobile device. The model is aimed towards context services, running on limited mobile environments with Internet access. The model is an overlay but running under different applications, therefore the model is open-ended with respect to applications.

Therefore, this model addresses the need for a solution which is capable of discovering nodes and maintaining a personal peer-to-peer network, which enables location-based social networks and information exchange in real-time. The key point of this model is that there exists no need to maintain global overlays and expensive routing protocols which can undermine the real-time characteristics. Instead all nodes that are in context proximity are kept within one hop in the overlay. This model differs from other existing solutions, because of the distribution, the gossiping method, and the way context is handled in the model. Therefore, by utilizing this approach, one can enable location-based ubiquitous context exchange in mobile environments.

The models preliminary evaluation show that the system is capable of keeping as good as possible real-time properties, and that the major drawback of the system is the centralized tracker. However, even with the tracker, it is still comparable to other related solutions.

In the future, the possibilities of enabling more complex context-ware applications on this model is going to be explored, for example enabling a multi-dimensional context-aware database, which enables the possibility of complex context queries. Future work include scalability measurements and possibilities of increasing the performance. Other interesting areas which is being explored,

is to remove the tracker and see what kind of quality of service would be preserved. In addition to this, there is the possibility of arranging the members in each of the personalized networks to an alternative structure that increases the quality of service, without introducing overhead. Further future work include utilizing multiple dimensions of context to progressively build personalized networks of related sensors and context dependent entities. This will then enable more optimized real-time searching and browsing of context information sources and entities allowing more time critical applications and services to be deployed.

Acknowledgment

This research is a part of the MediaSense project, which is partially financed by the EU Regional Fund and the County Administrative Board of Västernorrland.

References

1. Kanter, T., Österberg, P., Walters, J., Kardeby, V., Forsström, S., Pettersson, S.: The mediasense framework. In: Proceedings of 4th IARIA International Conference on Digital Telecommunications (ICDT), Colmar, France, pp. 144–147 (July 2009)
2. Araújo, F., Rodrigues, L.: Geopeer: A location-aware peer-to-peer system. In: IEEE International Symposium on Network Computing and Applications, pp. 39–46 (2004)
3. Kaneko, Y., Harumoto, K., Fukumura, S., Shimojo, S., Nishio, S.: A location-based peer-to-peer network for context-aware services in a ubiquitous environment. In: The 2005 Symposium on Applications and the Internet Workshops, Saint Workshops 2005, pp. 208–211 (2005)
4. Sripanidkulchai, K., Maggs, B., Zhang, H.: Efficient content location using interest-based locality in peer-to-peer systems. In: DEF, vol. 3(3) (2002)
5. Asaduzzaman, S., von Bochmann, G.: Geop2p: An adaptive peer-to-peer overlay for efficient search and update of spatial information. In: CoRR, vol. abs/0903.3759, p. 13 (2009)
6. José Viterbo, F., Endler, M., Sacramento, V.: Discovering services with restricted location scope in ubiquitous environments. In: MPAC 2007: Proceedings of the 5th International Workshop on Middleware for Pervasive and Ad-Hoc Computing, pp. 55–60. ACM, New York (2007)
7. Li, X., Calinescu, G., Wan, P.: Distributed construction of a planar spanner and routing for ad hoc wireless networks. In: IEEE INFOCOM, vol. 3, pp. 1268–1277 (2002)
8. Gao, J., Guibas, L., Hershberger, J., Zhang, L., Zhu, A.: Geometric spanner for routing in mobile networks. In: Proceedings of the 2nd ACM International Symposium on Mobile Ad Hoc Networking & Computing, pp. 45–55. ACM, New York (2001)
9. Marsan, M., Chiasserini, C., Nucci, A., Carello, G., De Giovanni, L.: Optimizing the topology of Bluetooth wireless personal area networks. In: IEEE Infocom, vol. 2, pp. 572–579. Citeseer (2002)
10. Basagni, S., Conti, M., Giordano, S., Stojmenović, I.: Mobile ad hoc networking. Wiley-IEEE Press (2004)

11. Schmidt, A., Beigl, M., Gellersen, H.: There is more to context than location. Computers & Graphics 23(6), 893–901 (1999)
12. P. 3gp, Ts 23 228: Ip multimedia subsystem (ims); stage 2 (release 9). 3GPP (December 2009), http://www.3gpp.org/ftp/Specs/html-info/23228.htm
13. Kanter, T., Pettersson, S., Forsstrom, S., Kardeby, V., Norling, R., Walters, J., Osterberg, P.: Distributed context support for ubiquitous mobile awareness services. In: Fourth International Conference on Communications and Networking in China, ChinaCOM 2009, pp. 1–5 (August 2009)
14. Ripeanu, M.: Peer-to-peer architecture case study: Gnutella network. In: Proceedings of First International Conference on Peer-to-Peer Computing 2001, pp. 99–100 (August 2001)
15. Raz, D., Juhola, A.T., Serrat-Fernandez, J., Galis, A.: Fast and Efficient Context-Aware Services. Wiley Series on Communications Networking & Distributed Systems. John Wiley & Sons, Chichester (2006)
16. Microsoft. Senseweb, http://research.microsoft.com/en-us/projects/senseweb/
17. Klemettinen, M.: Enabling Technologies for Mobile Services: The MobiLife Book. John Wiley and Sons Ltd., Chichester (2007)
18. Li, M., Lee, W.-C., Sivasubramaniam, A.: Semantic small world: An overlay network for peer-to-peer search. In: IEEE International Conference on Network Protocols, pp. 228–238 (2004)
19. Stoica, I., Morris, R., Karger, D., Kaashoek, F., Balakrishnan, H.: Chord: A scalable peer-to-peer lookup service for internet applications. In: Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, vol. 31, pp. 149–160. ACM Press, New York (2001)
20. Holmquist, L., Mattern, F., Schiele, B., Alahuhta, P., Beigl5, M., Gellersen, H.: Smart-its friends: A technique for users to easily establish connections between smart artefacts. In: Abowd, G.D., Brumitt, B., Shafer, S. (eds.) UbiComp 2001. LNCS, vol. 2201, pp. 116–122. Springer, Heidelberg (2001)

Energy Efficiency of Dynamic Interface Selection Mechanisms in Wireless Ad-Hoc Networks

Luis Sanchez, Jorge Lanza, and Luis Muñoz

Laboratorios I+D Telecomunicación, Plaza de la Ciencia s/n, 39005 Santander, Spain
{lsanchez,jlanza,luis}@tlmat.unican.es

Abstract. Energy efficiency is critical to ensuring scalability, embedding, and portability of emerging computing and communication systems. It is of particular interest in the design of mobile computing systems because of the limitations in energy and power availability. This paper presents and compares in terms of energy efficiency two strategies for the dynamic selection of the outbound interface on multi-radio devices in wireless ad-hoc networks. Findings from the studies show that intelligent selection of communication interface in heterogeneous ad-hoc networks leads to more efficient use of the energy consumed while assuring the quality of service parameters necessary for the correct provision of applications running on top of wireless ad-hoc mobile networks.

1 Introduction

Although wireless networks have existed for many years already, explicit concern about their energy efficient operation has emerged only recently. It is quite evident that when the power source is either costly or in short supply, energy efficiency is of paramount importance. In some wireless network applications, energy is actually entirely non-renewable and is thus an overriding constraint for the design and operation of the network.

However, things are not that simple. First of all, if energy efficiency is the only concern in a communication system, one might as well transmit nothing. Energy reserves would thus remain intact perpetually. Clearly communication performance is also of paramount interest. Thus, the choice of how to incorporate energy efficiency in the overall design is far from clear. One approach is to try to minimize energy consumption subject to throughput (or delay) staying above (or below) a certain threshold. Alternatively, one can try to maximize throughput (or minimize delay) per joule of expended energy. Neither of these approaches led to simple precise formulations or easy solutions.

Different approaches have been proposed to provide more efficient energy consumption at different layers. At the link layer, transmissions may be avoided when channel conditions are poor, as studied in [1]. Also, error control schemes that combine Automatic Repeat Request (ARQ) and Forward Error Correction (FEC) mechanisms may be used to conserve power (i.e. trade off retransmissions with ARQ versus longer packets with FEC) as in [2]. Energy efficient routing protocols may be achieved by establishing routes that ensure that all nodes equally deplete their battery

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 188–197, 2011.

power, as studied in [3]-[4] or that avoid routing through nodes with lower battery power. More complex solutions such as [5] exploit cross-layer operation and control the network topology by varying the transmitted power of the nodes so that certain network properties are satisfied.

Cross-layer design is particularly interesting under energy constraints, since not only energy across the entire protocol stack must be minimized, but also system performance must be optimized. While layer-specific solutions might disregard valuable information residing in other layers, cross-layer solutions can exploit global knowledge of the system to provide sub-optimal solutions at each of the different layers that result in optimal solutions at system level.

In this paper, the framework over which dynamic interface selection mechanisms have been implemented is briefly presented and two different outbound interface dynamic selection strategies are evaluated from an energy efficiency point of view.

2 Universal Convergence Layer

The concept of isolating the upper-layers from the underlying wireless technologies and thus providing real multi-mode can be achieved by introducing a Universal Convergence Layer (UCL) within the protocol stack. The UCL can be seen as a twofold approach. It will mainly act as an enabler for backward and forward compatibility by defining a common interface towards the network layer while managing several different wireless access technologies independently of their PHY and MAC layers. On the other hand, UCL also enables the cross-layer optimisation paradigm. Its privileged location within the protocol stack gives the UCL the possibility to support the information flow both bottom-up (e.g. use of SNR information for enriching the decision-making process in an ad hoc routing algorithm) and top-down (e.g. tuning of MAC parameters depending on the battery status or QoS requirements).

Figure 1 presents the different pieces in which the UCL can be divided. Each of these modules is specialized in providing the different features the UCL offers. This approach allows the easy addition and removal of functionalities depending on the requirements and characteristics of the system on which it will run. For the sake of simplicity and due to the scope of this paper, the building blocks presented in Figure 1 are the ones that are involved in the dynamic interface selection strategies that are evaluated in this paper.

- **Multi-radio Management module.** The UCL hides the complexity of the available air interfaces and offers a unique interface to the upper layers. UCL aims at masquerading multihoming by aggregating the different network interfaces (one per access technology the node is equipped with) on a single interface. Multi-radio devices in wireless ad-hoc networks will therefore keep only one network identifier (i.e. IP address) while being able to exploit the opportunities offered by its multi-radio nature. The possibility of using different links to one destination allows the UCL to intelligently modify the output interface according to the requirements and needs of the system.

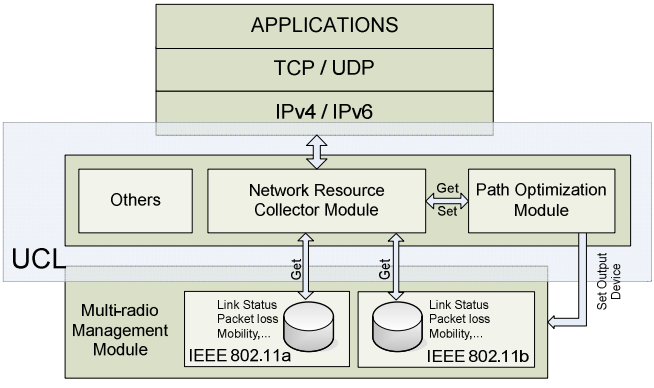


Fig. 1. UCL simplified high-level architecture

- **Network Resource Collector.** Data exported by the underlying technologies may offer important information to other layers in the stack. As this information has different meaning depending on where and by whom it is used, this module is specialized on collecting and translating data to make it available to the rest of the modules thus enabling the cross-layer optimization. As the different optimization rules are specific depending on the final purpose, this task is decentralized on different modules.
- **Path Optimization module.** Selecting the optimal output interface may depend on user preferences as well as the current status of the network. The outcome from the Network Resource Collector Module is analyzed and a decision is taken based on that.

Coordinated operation of these modules allows implementing communications management strategies as described in the following section.

3 Power-Aware Optimization Based on Dynamic Outbound Interface Selection

The UCL leverages Multiple Attribute Decision Making (MADM) algorithm in order to decide which of the available outgoing interfaces to use in order to guarantee system best possible performance. A utility function is evaluated for all the possible link layer interfaces and the one with the best result is selected.

Several policies and parameters might be used for implementing the utility function calculation. In this paper two of them are evaluated and compared. In both cases, the sender is the one that selects the interface on which it is transmitting the packets.

In the first of them, the UCL decides which interface to send the packet through taking into account the Signal to Noise Ratio (SNR) observed in each of the wireless channels available. Thresholds are set in advance for each wireless technology the device is equipped with. This way, the selection of the outbound interface maximizes the performance of the system given the SNR observed on each of the interfaces the

node can use for transmitting each IP datagram. It is important to note that SNR is a mean for assessing channel status. When the 802.11a channel becomes is bad (i.e. SNR goes below a pre-defined threshold), the status of the 802.11b channel is typically good.

The second adaptation technique is based on assessing the channel conditions using the number of lost packets. More specifically the UCL evaluates the bursts of correct and lost packets in order to know the actual status of the channel, and decides which radio interface to use correspondingly.

Wireless channels typically behave in a burst manner due to the fading processes they suffer. Hence, even when the SNR is relatively low we can correctly receive packets. This strategy attempts to take advantage of this behaviour and react quickly to these slots. Since it is not possible to get information about the packet los rate in the RAT that is not active, UCL takes decisions based on the bursts of erroneous or correct frames transmitted through the active RAT. In this sense, when a pre-determined number of frames are lost in a row, the UCL switches the transmission to a technology that is more robust. Of course, if channel conditions are so bad that even the most robust RAT suffers high FER, UCL cannot do anything. On the opposite direction, when a burst of frames are correctly sent, the UCL changes to the RAT that provides faster and/or energy efficient behaviour. To avoid as much as possible ping-point effects, the size of the bursts can be dynamically adapted so that it rises upon the detection of such abnormal behaviour.

The number of packets lost on which the threshold is set for swapping to use a more robust technology (downgrading threshold) can be adapted accordingly to have a trade-off between the throughput and the application loss. Similarly, the number of correct packets, received in a row through the more robust technology, with which we decide to use a faster technology (upgrading threshold) can also be tuned.

It is important to mention that the processing needed for this decision making is negligible since it is based on data exported by the underlying technologies and utility function calculation at the Path Optimization module can be greatly reduced. Additionally, since decisions are locally taken at the transmitter, no further signalling is necessary between the nodes involved in the communication.

4 Dynamic Interface Selection Mechanisms Power Efficiency Assessment

This section presents the results obtained from the evaluation analyses carried out in order to prove and validate the benefits introduced by the UCL solution for the selection of the most appropriate network interface for outgoing traffic.

The scenario selected for the analyses offered a range of situations from very good channel conditions (Location 1 in Figure 2) to poor ones that will produce a deep degradation of the communications (Location 3 in Figure 2). Basically, the charter of selecting this environment was to test the UCL on a real-world scenario which would allow us to extract conclusions that can be directly mapped on real user experiences. Two laptops were used, each of them equipped with one IEEE 802.11a and one IEEE 802.11b interface. Main rationale for this selection is that these technologies are

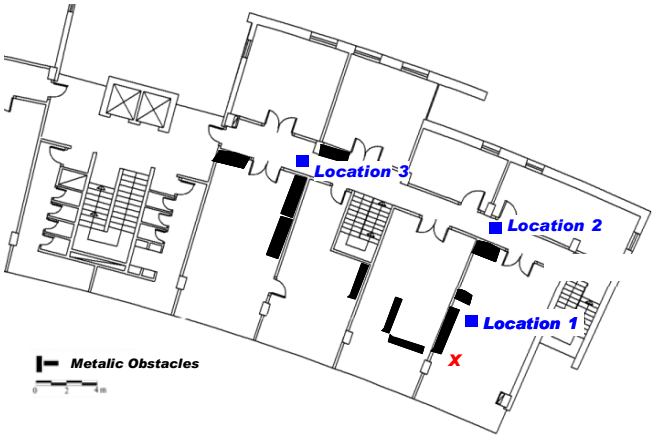


Fig. 2. Measurement campaign environment

nowadays the most widely used for Wireless Personal Area Networks (WPAN) and Wireless Local Area Network (WLAN) scenarios. Moreover, they operate in different frequency spectrum ranges, use different PHY and have slightly different MAC mechanisms, so they represent a good example of heterogeneity. The tests consisted on moving one of the laptops from Location 1 to Location 3 and back again while the other stays fixed in Location 1 (actually on the red cross). Transmission was done from the fixed laptop towards the mobile one. For both technologies we forced the transmission rate to be always the maximum possible, this is, 11 Mbps for IEEE 802.11b and 54 Mbps for IEEE 802.11a.

We simulated in Matlab® this moving scenario by generating a sequence of states which modelled the reception status of the transmitted frames. In our case we have used a Gilbert-Elliot model in which the parameters have been derived from [6] and tuned with our own experimental characterization of the office scenario shown in Figure 2. A total of 60000 link-layer frames were transmitted per simulation. In the scenario simulated, two thirds of the frames where transmitted over the best channel (Location 1) while the other third was equally transmitted over the other two channels (Locations 2 and 3). It is important to note that IEEE 802.11 MAC implements an ARQ scheme by which each frame is retransmitted a given number of times if an error occurs. Typically this number is fixed to 4. Thus, Frame Error Rate (FER) is not equivalent to Packet Error Rate (PER). This fact was also taken into account in the simulations. A Monte Carlo approach was taken so that simulations were run 1000 times.

Based on power consumption per bit values taken from [7] we can compare the efficiency of the different strategies in terms of power consumption and measure the optimization when weighting them against blind selection of the wireless interfaces.

As is shown in Figure 3, a blind selection of the 802.11a interface leads to slightly better use of power. This is due to its better natural energy efficiency (see Figure 4).

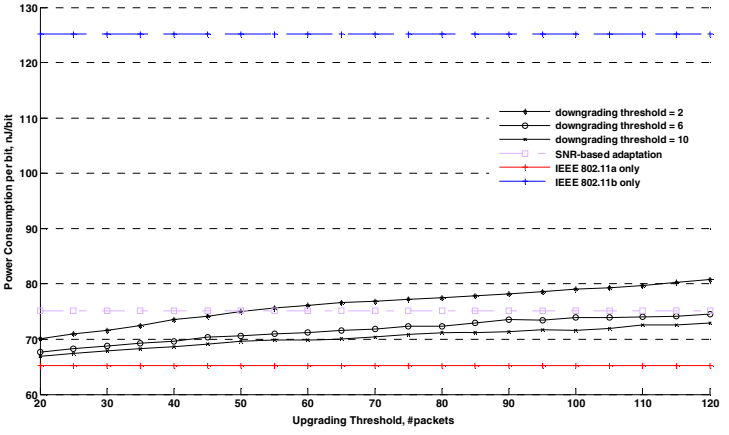


Fig. 3. Power consumption efficiency of UCL vs non-UCL approaches

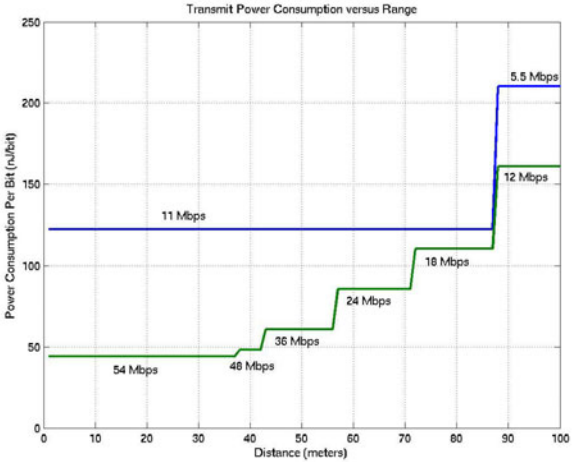


Fig. 4. Energy consumed per transmitted bit (802.11b vs. 802.11a) (source [7])

The consumed battery per unit of information is lower than the one of the other strategies. However, IEEE 802.11a is a less robust technology and suffers from higher FER under the same conditions. Thus, as it is shown in Figure 5, during the simulation a significant packet loss is experienced when using only the 802.11a. Packet loss shown in Figure 5 corresponds with loss at application level, this is, although the frame containing the packet was retransmitted 4 times at link level, none of these retransmissions arrived at the receiver error-free. Despite this packet loss, that would ruin the communication, and the associated number of retransmissions, inherent energy efficiency of 802.11a interface for transmission compensates it and results on slightly better results when looking only at the energy efficiency metrics.

This is also true because of the ratio chosen in the scenario between good and bad channel conditions (i.e. 2/3 vs 1/3). A different ratio would change the results in terms of energy efficiency of using IEEE 802.11a only. In contrast, when only using 802.11b the efficiency is much lower even though the low frame error rate experienced is much better on any of the channels that compose the moving scenario.

Having said this, it is not enough for making a correct comparison to look at only one metric but is necessary to assess the gain of the proposed solutions both from an energy efficient point of view and from a quality of service standpoint.

This is better seen in Figure 5 where the left Y-axis represents the power consumption per bit and the right Y-axis the packet loss at application level obtained in each of the cases. When selecting the appropriate parameters for the packet loss based adaptation strategy the UCL achieved relatively small power consumption per bit figures, but highly reduced the packet loss at application level. For example, using 6 packets as the downgrading threshold and 65 packets as the upgrading one we can obtain three times less packets lost at the application level, while still keeping the overall power consumption per bit only 10% higher.

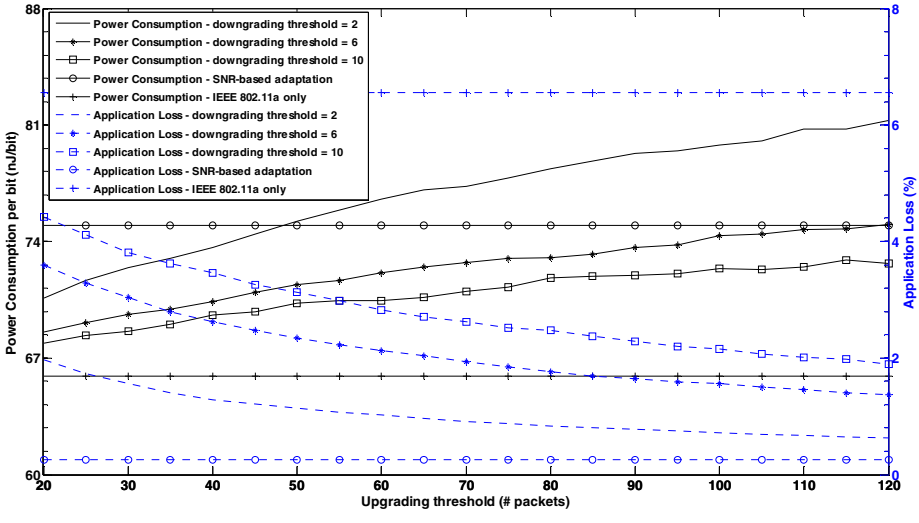


Fig. 5. Power consumption efficiency and packet loss of UCL and 802.11a only approaches

When the SNR-based approach is used the packet loss is reduced drastically (around 20 times less), while keeping the power consumption only 15% higher than when 802.11a interface is used all the time.

Making similar studies with the achieved throughput, we were able to obtain almost 2.5 times more throughput than using 802.11b only while still reducing the overall power consumption per bit by 40% as it can be seen in Figure 6 for the case when 6 packets are used as the downgrading threshold and 65 packets as the upgrading one. This means almost halving the total power consumed during the simulation. When the SNR-based approach is used we also double the throughput while the power consumption is still 40% less than when using the 802.11b interface only.

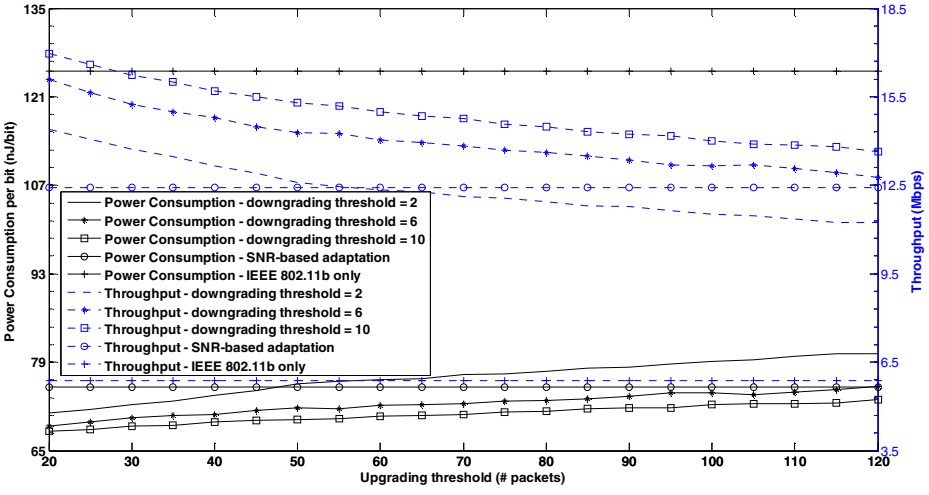


Fig. 6. Throughput versus power consumption efficiency of UCL and 802.11b only approaches

We can conclude that an intelligent use of the available interface leads to a sub-optimal power consumption, but does not jeopardize the system performance as would be the case with an 802.11b only approach, where throughput is reduced or with an 802.11a only approach, where application-level loss is much higher.

Only UDP transport protocol has been used during the assessment since the TCP congestion avoidance mechanisms would have stopped the transmitter when the channel conditions were poor, thus making it impossible for us to compare the energy efficiency of the different alternatives shown. UCL would keep performance on a reasonable level by using the most robust wireless technology while the device is facing bad channel conditions [8]. On contrary, when only IEEE 802.11a is used, as soon as the communication experience, TCP emitter is stopped [9]. Therefore, in terms of energy efficiency, this latter approach would be better, but at the cost of ruining the throughput and hence the services being provided.

5 Conclusions

It has been demonstrated that wireless ad-hoc networks energy efficiency can be enhanced by taking advantage of multi-radio capabilities of the devices. Opposite to other approaches, like rate adaptation mechanisms [10], that focus only on one technology, UCL allows seamless service provision in multi-radio mobile nodes. This enables larger versatility since the specific advantages of all the available WPAN and WLAN technologies can be exploited depending on the application and user requirements. Even though a vertical handover might be forced in order to use the most appropriate interface during the actual service provision, the session is not affected.

The solution proposed has proven to be able to offer a good range of possible outputs depending on the parameters chosen for each of the adaptation strategies. In this sense, the focus has not been so much on finding the optimal configuration but on assessing the optimization capacity of the solutions proposed. This is important since it would be able to fit the demands of a wide range of services, users and channel conditions.

Most interesting finding is that taking advantage of the global view that the convergence layer provides (access to information from multiple layers and access technologies) a complete map of the system can be inferred. After matching it with the user requirements, the UCL can take the appropriate decisions in order to best serve the final users' wishes, optimizing their quality of experience not only by providing enhanced bandwidth but also improving the power consumption efficiency or enhancing the service provision by lowering the packet loss due to wireless channel impairments. In this sense, the UCL selection strategies does not look for optimization of throughput only but the decisions taken depend on a multi-parametric matrix that include battery status, type of application or user preferences for example.

Implementation over real testbed has been done for the SNR-based adaptation approach. The results from experimental validation support the conclusions from this paper.

Acknowledgements

The authors would like to thank the Spanish Government (Ministerio de Ciencia e Innovación) for its funding in the project "Cognitive, Cooperative Communications and autonomous SErvice Management", C3SEM (TEC2009-14598-C02-01).

References

1. Zorzi, M., Rao, R.R.: Energy constrained error control for wireless channels. IEEE Personal Communications Magazine 4(6), 27–33 (1997)
2. Lettieri, P., Fragouli, C., Srivastava, M.B.: Low power error control for wireless links. In: Proceedings from the 3rd Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 139–150 (September 1997)
3. Woo, M., Singh, S., Raghavendra, C.S.: Power Aware routing in mobile ad hoc networks. In: Proceedings from the 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking, pp. 181–190 (October 1998)
4. Chang, J.-H., Tassiulas, L.: Energy conserving routing in wireless ad-hoc networks. In: Proceedings from the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 22–31 (March 2000)
5. Ramanathan, R., Rosales-Hain, R.: Topology Control of Multihop Wireless Networks using Transmit Power Adjustment. In: Proceedings from the 19th Annual Joint Conference of the IEEE Computer and Communications Societies, pp. 404–413 (March 2000)
6. Arauz, J., Krishnamurthy, P.: Markov modeling of 802.11 channels. In: Proceedings from the 58th IEEE Vehicular Technology Conference (October 2003)
7. Texas Instruments White Paper, Low Power Advantage of 802.11a/g vs. 802.11b (December 2003)

8. Sánchez, L., Lanza, J., Muñoz, L.: Experimental assessment of a cross-layer solution for TCP/IP traffic optimization on heterogeneous personal networking environments. In: Cuenca, P., Orozco-Barbosa, L. (eds.) PWC 2006. LNCS, vol. 4217, pp. 284–296. Springer, Heidelberg (2006)
9. García, M., Agüero, R., Muñoz, L.: On the unsuitability of TCP RTO estimation over bursty error channels. In: Niemegeers, I.G.M.M., de Groot, S.H. (eds.) PWC 2004. LNCS, vol. 3260, pp. 343–348. Springer, Heidelberg (2004)
10. Lacage, M., Manshej, M.H., Turletti, T.: IEEE 802.11 Rate Adaptation: A Practical Approach. In: Proc. MSWiM 2004, Venice, pp. 126–134 (June 2004)

Ubiquitous Computing by Utilizing Semantic Interoperability with Item-Level Object Identification

Janne Takalo-Mattila, Jussi Kiljander, Matti Eteläperä, and Juha-Pekka Soininen

VTT Technical Research Centre of Finland
P.O. Box 1100
FI-90571 Oulu, Finland
{janne.takalo-mattila,jussi.kiljander,matti.etelapera,
juha-pekka.soininen}@vtt.fi

Abstract. This paper presents a novel approach for utilizing item-level object identification in ubiquitous computing environment where the interaction between devices is based on semantic information interoperability. The paper also presents novel methods for human-machine interaction. In our approach we give a unique identifier for objects in the environment and combine the knowledge of the object identity with information from other sources by utilizing semantic information interoperability. The approach utilizes Smart-M3 interoperability solution for sharing semantic information between heterogeneous devices and RFID-technology for identifying physical objects in the environment. In order to demonstrate the use item-level tagging with semantic interoperability we have implemented a Smart Greenhouse demonstrator that consists of several smart devices and tagged objects.

Keywords: Ubiquitous computing; object identification; Smart-M3; interoperability.

1 Introduction

Computers have already spread everywhere in the world, and current trend is moving from one computer to multiple computers for each person. In the future it is expected that there can be up to one thousand tiny computers per each person [1][2]. The recent trend of ubiquitous computing will also make these computers and embedded devices far more invisible and natural part of our everyday life, so that we will use computers without even thinking about it. In this paper we use the term smart object for these devices that are capable to interact with each other and with the physical environment without human assistance.

Our homes and working places are already full of various electronic devices which provide the user with heterogeneous user interfaces. In addition most of these devices don't have ability to utilize information of other device as effectively as possible. The vision of ubiquitous computing is to make our lives easier by providing us with meaningful services. In order to realize this vision, typical ubiquitous computing environment requires large amount of different kind of devices and applications that are capable to exchange information seamlessly and interpret the meaning of the

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 198–209, 2011.

information similarly. Achieving information level interoperability between heterogeneous devices can be challenging task however. Our approach to share information in ubiquitous computing environment is based on Smart-M3, which utilizes ontology-based information presentation and semantic web technologies to provide information-level interoperability for applications and devices in the environment. Smart-M3 functional architecture consists of two main entities: Knowledge Processor (KP) and Semantic Information Broker (SIB). SIB is the storage of ontology and provides interface for semantic information for the KPs. The actual smartness of the environment lies in KPs that perform meaningful actions according to the information fetched from the SIB. The interaction between the SIB and KPs is specified in Smart Space Access Protocol (SSAP).

In addition to exchanging information meaningfully, sensing the environment and the ability to interact with physical environment are key issues in ubiquitous computing. Therefore it is necessary to have an ability to identify physical objects and locations in order to enable meaningful interaction in with the environment. Radio-frequency identification (RFID) is one of the most popular methods to identify objects and it has been used in large variety of applications. Lately the price of RFID-tags has also decreased and this has made the technology even more tempting for smart space applications. RFID has been the enabler of automatic identification of tagged objects since the beginning of the smart space research. For example retail applications utilize universal item-level tagging, where every object is tagged with unique identifier (ID). Currently we have seen new kind of applications in the field of ubiquitous computing such as ubiquitous learning [3]. Data repositories provided by EPCglobal [4], uID [5] and others have enabled usage of this kind of universal item-level-based applications. However in local scale applications it is more feasible to use local database that is more suitable for storing the rapid changes in the environment such as states of actuators and locations of objects.

In this paper we will present how item-level object identification and semantic interoperability can be combined in a novel way to achieve new kind of methods for human-machine interaction and how smart objects can exploit the information about tagged items in Smart Greenhouse environment. In the proposed method we have given each object a unique ID, which can be read using mobile device with RFID-reader. The necessary information about the environment is stored in local information storage SIB that conveys the information to the smart objects in the space. That is how we can achieve semantic interoperability between devices and enable the devices to improve their behaviour according to information about tagged items.

2 Background

2.1 Ubiquitous Computing

The inventor of term “ubiquitous computing” Mark Weiser has written that “The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it” This sentence reminds the purpose of the ubiquitous computing, which is to make computers so invisible and natural to use that people can use them without thinking about it. Other similar terms

for the vision behind ubiquitous computing are for instance pervasive computing, ambient intelligence and smart spaces. The key idea behind these terms is to achieve a technology that enables heterogeneous ubiquitous devices to communicate autonomously with each other in order to assist us in our everyday life. These ubiquitous devices can be divided to three groups:

- Sensors, which collect information about the environment, humans and other objects in space. Context-awareness is one the key issues ubiquitous computing.
- Actuators, which can make changes to environment.
- Processors, which read data and make decisions according that that data [6].

Even though devices can be divided to previous groups, many of real-life devices have some qualities of each of these groups. One example of ubiquitous device that has several qualities of previous groups is a mobile phone. Current mobile phones are already capable to sense the environment and make decisions according to that data, however, in order to exploit the true power of ubiquitous computing, different devices should be able to exchange information seamlessly with each other. In any case the mobile phones are probably the most interesting physical interface for ubiquitous computing environment [7].

2.2 Object Identification with RFID

One of the key issues in ubiquitous computing is how to identify objects, locations and people. RFID has been used in wide variety of applications to identify objects and it is considered to become one of the most pervasive computing technologies in history [8]. The basic idea behind RFID is marking items with tags, which contain transponder to send message to RFID-reader. Mostly this message is just an identification number, but additional information can be also stored to tags. Decreased prices of individual RFID-tags have enabled the usage of RFID-technology in novel applications such as retail business, for example. Basically RFID can be consider as an alternative for barcodes or other visual tags, but it also improves the functionality by enabling tag reading without visual contact, providing larger information density and also providing two-way communication ability [9].

Especially interesting subclass of object identification is item-level tagging, where objects and locations have unique ID. In the future printed electronics may also provide very low-cost passive RFID-tags, which could be used as a replacement of the barcodes [10]. This makes tagging individual objects with RFID financially reasonable. The best known examples of universal item-level tagging are EPCglobal and uID, where all objects are numbered uniquely and thereby physical objects can be connected to digital counterparts.

Figure 1 shows a simplified architecture of universal item-level tagging. Object is tagged uniquely with RFID and tag reader reads ID from tag attached to object. Digital counterpart for that object is stored in remote database, which can be accessed via internet using tag reader device. For example uID-architecture uses 128bit long unique ID number called ucode for every object and location. 128-bit ID can be assigned

practically limitless amount of objects. This ucode is stored in a tag, which can be either RFID-type or optical type such as barcode. Basic operation of uID architecture in simplified manner is that handheld computer called Ubiquitous Communicator (UC) reads ucode tag and retrieves corresponding data from remote database. UC and uID-architecture have been developed in YRP Ubiquitous Networking Laboratory (Tokyo, Japan), whose chairman professor Ken Sakamura is one of the pioneers in ubiquitous computing.

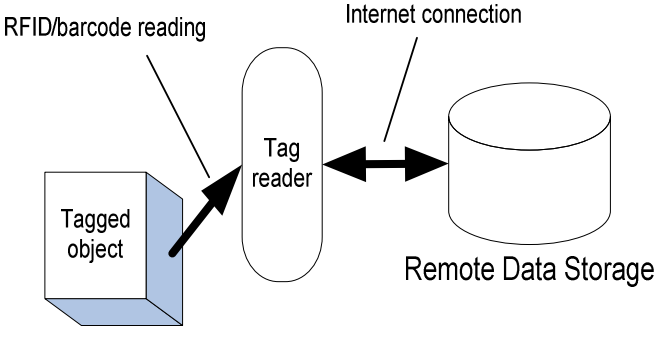


Fig. 1. Simplified architecture of universal item-level tagging

2.3 Semantic Interoperability and Smart-M3

The Semantic Web provides information level interoperability by presenting information as a collection of information called ontology [11]. It enables the computer applications to automatically interpret the meaning of the information and exploit it without human guidance. Several different technologies and design principles such as the Resource Description Framework (RDF), RDF Schema (RDFS), Web ontology language (OWL) and SPARQL have been developed in various Semantic Web working groups to enable information level interoperability between devices and humans. RDFS and OWL provides basic elements for describing ontologies and RDF is used to present ontologies in form of subject-predicate-object expressions, known as triples in RDF terminology. SPARQL is RDF query language and it is used to query RDF-type data.

Smart-M3 is an open source information level interoperability solution that can be implemented on top of any connectivity solution and it is device, domain and vendor independent [12]. It provides an architecture for sharing semantic information between software entities and devices. The information level interoperability in Smart-M3 is based on two core concepts: ontology-based information model and event-based functional architecture. The ontology-based information model of the Smart-M3 utilizes RDF and OWL technologies to present the information in ontology format. The Smart-M3 functional architecture defines how to access the shared information of the environment. Two main entities of the Smart-M3 functional architecture are KP and SIB. Figure 2 illustrates the core elements of Smart-M3 and their relations.

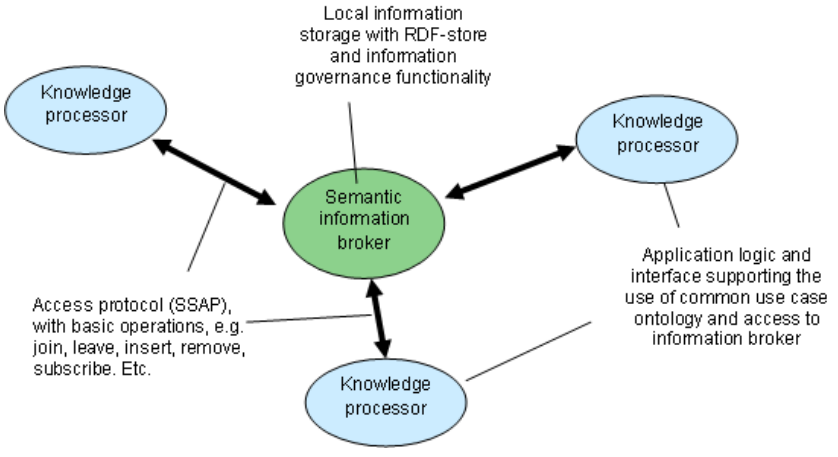


Fig. 2. Smart-M3 functional architecture

The Smart-M3 functional architecture is based on publish/query paradigm. SIB acts as a data repository, which stores semantic information as a RDF-type graph, which can be utilized by KPs. The actual smartness of the smart space lies in KPs, which can make smart decisions according to information that can be received from SIB. SSAP is the communication protocol of the Smart-M3 and it defines the rules and syntax of the communication between SIBs and KPs. Protocol supports all basic operations to publish and query information in ubiquitous environment.

3 The Use of Item-Level Object Identification with Smart-M3

In this section we present, how item-level object identification can extend the possibilities of Smart-M3 –based ubiquitous computing environment. We will combine item-level object identification with Smart-M3 to achieve:

- more natural human-machine interaction
- easier development of user interfaces for simple devices
- increased awareness of the environment and objects in the environment

3.1 System Model

System with heterogeneous devices that can share information using Smart-M3 interoperability solution is presented in figure 3. Objects in the ubiquitous computing environment are tagged with unique ID, which can be used to interact with object. In this approach we are using uID-based structure for presenting IDs. This structure is presented in section II. Reason for selecting the ucode structure is the possibility to utilize remote uID-databases in future applications.

Information of the tagged objects is stored in SIB, which is local database for storing the information in RDF-format. This information can be inserted to SIB using local database KP, which reads local item database and inserts information about

possible items or ubiquitous devices to the SIB. Another option is to use KP, which can read remote database, e.g. uID-database, and add required information to the local SIB. The usage of local database KP is the best suitable option for initializing information about ubiquitous devices such as actuators, sensors and processors to the SIB. Usually the status of the actuators and sensors are changing relatively quickly so it is reasonable to store this kind of information to local database that provides easier and faster methods for exchanging data via shared memory than remote databases. SIB provides KPs with straightforward methods to insert e.g. changed measurement result of the sensor to memory, so it is very natural to use it as a local database. The usage of remote database is more suitable for storing unchangeable information about items e.g. consumer products, food and plants.

Figure 3 contains all kinds of ubiquitous devices: actuators, sensors and processors, which all have their unique identifier. Besides these ubiquitous devices different locations and items in the smart space are assigned with unique identifier. KPs have to be implemented to every ubiquitous device in smart space in order to publish and query information in smart space. Because of the variety of different ubiquitous devices implementing KP can be challenging task, but utilizing ANSI-C based portable KP interface makes KP design simpler [13]. In our approach KPs utilize the SIB service to share information with each others. Tag reader is the mobile device with RFID communication capability and it also contains KP implementation with ability to publish and query information using SIB service. Figure 3 presents identifiers as a RFID, but it is also possible to use e.g. barcodes.

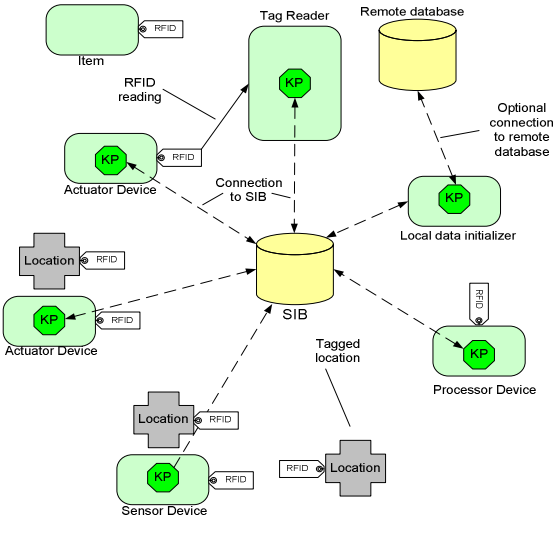


Fig. 3. Smart-M3 environment using item-level tagged objects

The basis of information presentation in the Smart-M3 information level interoperability solution is the usage of the ontologies. The most important task in implementing Smart-M3 –based ubiquitous computing environment is to design

common ontology and common data presentation format. Commonly accepted method for identify a name or resource on the Internet is a string of characters called Uniform Resource Identifier (URI), which is also exploited in OWL. In our approach we are using 128-bit unique identifiers as URIs that can be used to identify objects in smart space. This method decreases the amount of memory needed to store the information of the objects in SIB compared to method that uses separate URI and unique identifier. Figure 4 shows RDF graph of example ontology related to ubiquitous devices, where unique identifiers such as id_0 are inserted in ontology in order to act as URIs. Each of these unique identifiers represent individual object: actuators, sensors, processors and items. The class of instance has been presented using type-indicator.

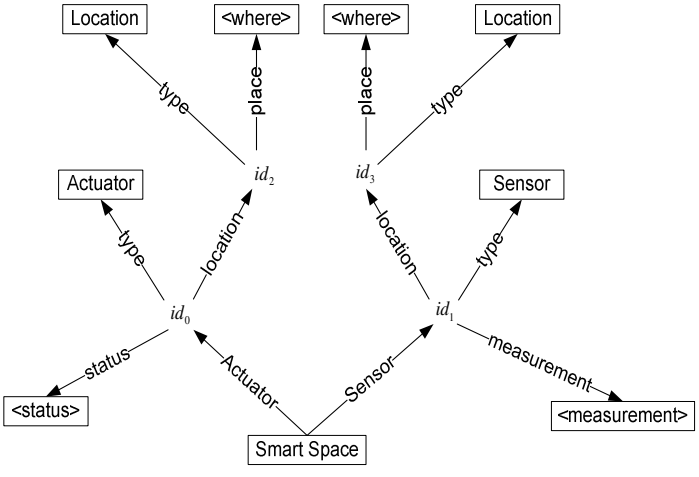


Fig. 4. RDF-graph of smart space with tagged items

3.2 Novel Applications Using Combined Item-Level Object Identification and Smart-M3

Identifying every object in ubiquitous computing environment with unique identifier brings novel approaches to interaction with objects. Mobile phones are expected to become to most important physical interface for ubiquitous computing environment so it would be very natural to use single mobile device with tag reader to control every ubiquitous device in the environment [14]. The same single device can be utilized also for viewing state information of the environment, which is produced by different sensors such as temperature-, humidity-, acceleration- and GPS-sensor.

Activating some function with mobile phone by selecting physical tag is an interesting vision for human-machine interaction [15]. Very clear examples of utilizing this idea are changing the status of an actuator, reading the measurements of the sensor or receiving information about item by touching these objects. Many simple ubiquitous devices can be controlled very similar ways, for instance basic operations for simple

actuators such as air conditioner, heaters and lights are usually almost identical: they can be turned on or off and their level can be adjusted. Also many sensors produce very similar kinds of results, which usually include the measurement value and unit. Utilizing these similarities we can build harmonized user interfaces for mobile device that acts as a physical interface for ubiquitous computing environment. That is how we can control different kinds of actuators very similar way, which can improve usability significantly. If we can use our mobile phones as controllers and sensor readers, other devices can be made simpler and it is not necessary to incorporate expensive screens and buttons for every device.

Another example of using this approach is increasing the awareness of the environment and objects in the environment by touching. User can e.g. insert new objects in the environment and change the location of objects by touching object and location tag. Information about new objects and their locations can be exploited by processors that can control actuators. Typical example is the greenhouse, where different kinds of plants need different kind of care. Using the knowledge about what kind of plants are inserted in greenhouse, processor can control actuators such as heaters and lights in different ways. This kind of approach provides also interesting possibilities for user to make configurable actions in smart space relatively easily. For example user could decide what happens if certain types of items are brought to certain location or what happens if certain location tag is touched.

In each of these cases, heterogeneous devices such as actuators, sensors and processors need to interpret exchanged information. Information can be for example the status of actuator, measurement of sensor or the location of object. Therefore Smart-M3 based semantic interoperability is suitable method for information exchanging. Using proposed approach each of these cases follows the same basic procedure: reading the unique identifier using RFID-reader, querying unique identifier related data such as type, status, measurement or location from SIB, showing this information on mobile devices screen and optionally inserting modified information such as status of the actuator to SIB. In our approach we can combine information from different sources including Internet, RFID-reader, sensors and actuators to same ontology and store the corresponding RDF-graph in SIB. That is how we can implement interesting Smart-M3 mash-up applications that utilize combined information from different sources.

4 Implementation

In order to demonstrate the usage item-level tagging with semantic interoperability we have used our Smart greenhouse demonstrator. Smart greenhouse consists of five smart objects and several different plants. Figure 5 presents five KPs in demonstration environment exchanging information using SIB as an information repository. Different plants have been tagged with RFID-tags and that is how unique ID has been given to each of the plants.

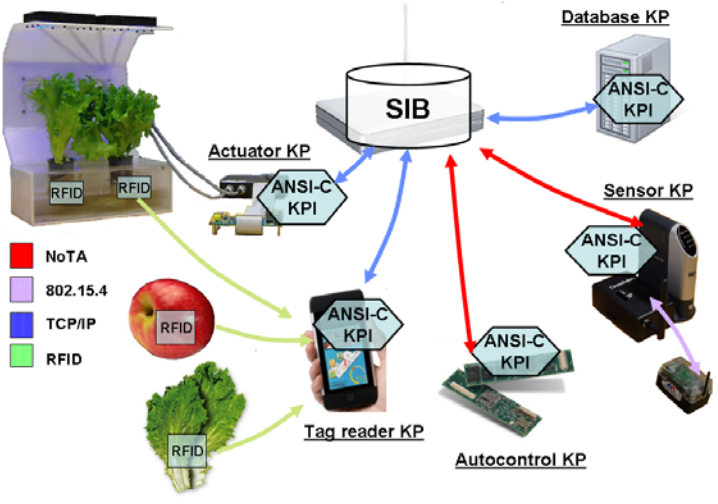


Fig. 5. Devices and plants in the Smart Greenhouse

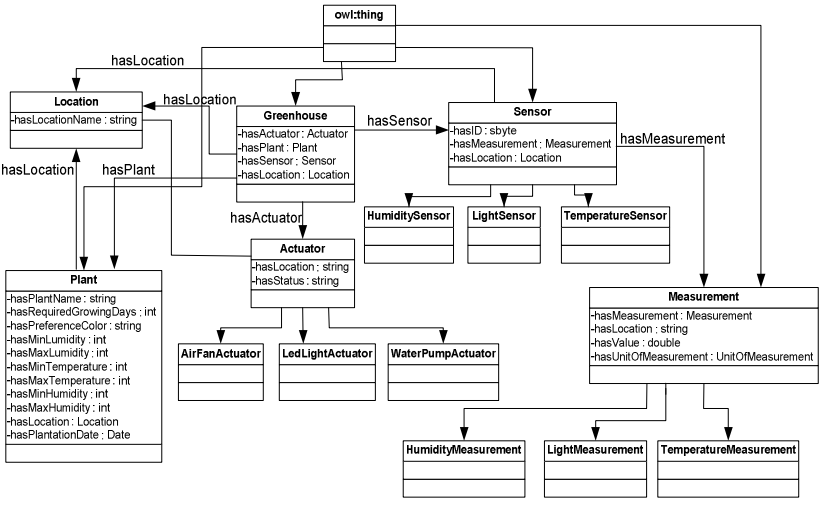


Fig. 6. Smart greenhouse ontology

In this demonstration case SIB is running in Asus WL500GPv1 Wireless Local Area Network (WLAN) access point (AP) with OpenWrt firmware. Demonstration environment has also five KPs for different purposes: Actuator KP for controlling greenhouse, Sensor KP for sensing the environment, Tag reader KP for detecting tagged objects, Database KP for initializing information about tagged objects to SIB and Autocontrol KP for controlling greenhouse when gardener is not present.

Smart greenhouse ontology is presented in figure 6. Plant, location, actuator, sensor, measurement and greenhouse are the main entities in the ontology, which is modelled using OWL. Every possible plant locations have been tagged with RFID-tags and we have also tagged several plants with similar kind of tags. RFID-tags contain unique 128-bit IDs, which are stored on the ontology using method presented in figure 4 in section 3.

Information about tagged objects is inserted to SIB using Database KP. Currently only local database is used, but in the future it could be possible to combine local database with remote databases e.g. uID database. In our approach information about every tagged object is initialized to SIB beforehand. We have ported our ANSI-C – based KP interface to UC with T-Kernel OS and implemented Tag reader KP on top of that KP interface. UC uses Bluetooth connected RFID-reader for tag reading.

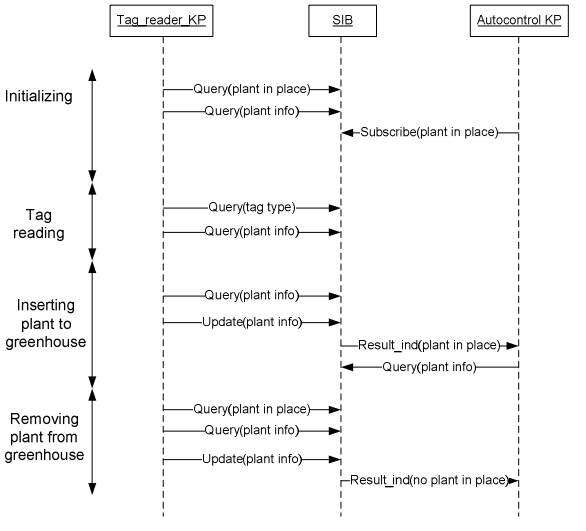


Fig. 7. Simplified sequential diagram of greenhouse activity

Figure 7 presents simplified diagram, how two different KPs operate in the greenhouse environment. Other KPs have been left off from the figure in order to clarify the case. At first the Tag Reader KP checks whether there are already some plants in the greenhouse and shows their information in the screen. Autocontrol KP also subscribes to the plants inserted to the greenhouse. When user touches uniquely tagged plant, Tag Reader KP first queries the type of the tag and then queries information related to that tag. User can also insert plants to greenhouse and in this case Tag Reader KP updates plant location to the SIB. When location is changed, Autocontrol KP gets indication about this event and queries plant information from the SIB. Each plant has different kind of preference conditions and Autocontrol KP tries to adjust temperature, humidity and light using these preferences that can be queried from the SIB. Autocontrol KP utilizes information about plants in the greenhouse and their preference conditions and compares these with current condition information produced

by Sensor KP. If current conditions such as humidity, temperature and lumidity differ from plant preference conditions, Autocontrol KP updates new status for virtual actuator responsible for that environmental condition. The SIB then indicates Actuator KP about this event and Actuator KP can modify the corresponding physical actuator.

The results of using this approach in smart greenhouse environment showed the easiness to bring new information to environment by touching. Autocontrol KP is an example of application that can improve its behaviour using the knowledge of tagged objects in the smart space, in this case plants in the greenhouse. The same approach could be utilized in smart buildings, where devices adjust their behaviour to be more suitable for different persons. Smart-M3 provided efficient way to share information between smart objects and using semantic interoperability completely new smart space applications can be implemented rapidly. Because item-level information is stored in the ontology, it is very easy to utilize this information for different purposes.

5 Conclusions and Future Work

This paper presented an approach to use item-level tagging with Smart-M3 to produce novel method for interactions with tagged objects in ubiquitous computing environments. Our experiences obtained from the work were positive and by using semantic interoperability between different devices it was easy to bring new information to environment by touching. This approach makes also possible to reduce costs by offering possibility to make simpler devices without own screens and buttons. Item-level tagging can however cause performance issues to communication between KPs, because that kind of approach requires large amount of data to be inserted in SIB that can slow down the operation of the SIB. One solution for that problem could be the usage of external database, which inserts information to SIB only when needed.

As for future work, we are planning to test more detailed user interactions with different kind of sensors, actuators and other objects in ubiquitous computing environment. Moreover we are going to connect remote databases e.q. uID database to our demonstration environment in order to reduce the problems related to large amount of data inserted to SIB.

Acknowledgments. This work has been funded by the Open Ubiquitous Technology (OPUTE), Device and Interoperable Ecosystem (DIEM) and Smart Objects for Intelligent Applications (SOFIA) projects. The author would like to express his gratitude also to professor Ken Sakamura and YRP UNL.

References

1. Uusitalo, M.: Global Visions for the Future Wireless World from the WWRF. IEEE Vehicular Technology Magazine (2006)
2. Weiser, M.: The computer of the 21st century. Scientific American, 94–100 (2001)
3. Sakamura, K., Koshizuka, N.: Ubiquitous Computing Technologies for Ubiquitous Learning. In: IEEE International Workshop on Wireless and Mobile Technologies in Education, WMTE 2005, November 28-30 (2005)

4. EPCglobal, http://www.epcglobalinc.org/home/
5. uID-architecture, http://www.uidcenter.org/index-en.html
6. Dhingra, V., Arora, A.: Pervasive Computing: Paradigm for New Era Computing. In: 2008 First International Conference on Emerging Trends in Engineering and Technology (2008)
7. Ballagas, R., Borchers, J., Rohs, M., Sheridan, J.G.: The smart phone: A ubiquitous input device. IEEE Pervasive Computing 5(1), 70–77 (2006)
8. Peris-Lopez, P., Hernandez-Castro, J.C., Estevez-Tapiador, J.M., Ribagorda, A.: RFID Systems: A Survey on Security Threats and Proposed Solutions. In: Cuenca, P., Orozco-Barbosa, L. (eds.) PWC 2006. LNCS, vol. 4217, pp. 159–170. Springer, Heidelberg (2006)
9. Weinstein, R.: RFID: A Technical Overview and Its Application to the Enterprise. IT Professional 7(3), 27–33 (2005)
10. Subramanian, V., Frechet, J.M.J., Chang, P.C., Huang, D.C., Lee, J.B., Molesa, S.E., Murphy, A.R., Redinger, D.R., Volkman, S.K.: Progress toward development of all-printed rfid tags: Materials, processes, and devices. Proceedings of the IEEE 93(7), 1330–1338 (2005)
11. Berners-Lee, T., Hendler, J., Lassila, O.: The Semantic Web. Scientific American Magazine (2001)
12. Soininen, J., Liuha, P., Lappeteläinen, A., Honkola, J., Främling, K., Raisamo, R.: DIEM project, White paper, http://www.tivit.fi/fi/dokumentit/64/DIEM%20whitepaper.pdf
13. Kiljander, J., Eteläperä, M., Takalo-Mattila, J., Soininen, J.: Opening information of embedded systems for Smart Spaces. In: WISES 2010, 8th IEEE Workshop on Intelligent Solutions in Embedded Systems, Heraklion, Crete, Greece, July 8-9 (2010)
14. Ailisto, H., Pohjanheimo, L., Välkkynen, P., Strömmer, E., Tuomisto, T., Korhonen, I.: Bridging the physical and virtual worlds by local connectivity-based physical selection. Personal and Ubiquitous Computing 10(6), 333–344 (2006)
15. Ailisto, H., Plomp, J., Pohjanheimo, L., Strömmer, E.: A Physical Selection Paradigm for Ubiquitous Computing. In: Aarts, E., Collier, R.W., van Loenen, E., de Ruyter, B. (eds.) EUSAI 2003. LNCS, vol. 2875, pp. 372–383. Springer, Heidelberg (2003)

Manager Selection over a Hierarchical/Distributed Management Architecture for Personal Networks

Jose A. Irastorza, Ramón Agüero, and Luis Muñoz

Department of Communications Engineering – University of Cantabria, Spain
{angel,ramon,luis}@tlmat.unican.es

Abstract. In spite of having been the focus of several works, traditional management architectures, usually based on a centralized model, are not suitable for the particular characteristics of personal networks and their underlying multi-hop topologies. A hierarchical/distributed approach is proposed in this work, which also analyzes different strategies to optimally select the nodes taking the manager role. In order to assess the benefits and drawbacks of these mechanisms, a proprietary simulator was developed, and different metrics were studied (probability for a node to take part on the management architecture, number of hops needed to reach a manager, and fairness of the distribution of the management burden). A novel heuristic is proposed to enhance one of the analyzed strategies, and it is shown to outperform the rest of algorithms.

Keywords: Personal Networks, Management Organization Model, Distributed-Hierarchical Models, Algorithmic assessment.

1 Introduction

The evolution and proliferation of wireless devices and peripherals have been as remarkable as the aim of applying network technologies to interconnect them. The resulting communication architectures are conditioned by some particular characteristics of wireless environments: dynamicity of nodes, heterogeneity of the devices and involved technologies, as well as energy and bandwidth constraints. A typical scenario can embrace multiple kinds of devices, ranging from modern laptops to low cost, low capacity sensors and actuators, interconnected via e.g. a wireless multi-hop network, which might be spontaneously deployed over a geographically limited area, which can be referred to as personal surface. This illustrative scenario has been given the name of personal network (PN) [1]. Multi-hop (or mesh) network topologies are usually adopted to implement the subjacent connectivity over these network deployments.

One key issue of any type of network, and especially a personal network, comes from its management. Hence, the management tasks associated to personal networks can be seen as a challenge that must be tackled in order to facilitate the operability of these deployments. Although different network management architectures and models have been widely studied over fixed networks, it must be considered that managing a personal network poses several new difficulties, taking into account the specific characteristics of such deployments. First, communication links within wireless multi-hop networks are, intrinsically, unreliable, dynamic (due to node movements) and

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 210–224, 2011.

showing varying capacity. Furthermore, nodes have several constraints, like limited battery and data storage capacity. The main consequence is that the resulting topology is unpredictable, and thus automatic on-demand reconfiguration procedures are usually required. These relevant facts pose some requirements to the network management task and, generally, to any service to be implemented over these networks. Questions like discovery protocols/procedures, autonomous topology and node reconfiguration, security and signaling overhead, just to name a few, must be resolved in order to efficiently manage a personal network.

The work presented in this paper can be focused on providing some answers to the "how to manage personal networks" question, by proposing an organizational model with an optimal distribution of the manager role to seek that a maximum number of nodes are managed, as well as balancing the management burden (fair distribution of agents between the available managers), so as to minimize the overhead introduced by the corresponding management traffic.

There are some other works which have analyzed the implications of multi-hop topologies over network management; two of the most referenced ones are the GUERRILLA framework [2] and the Ad Hoc Network Management Protocol [3], which propose a hierarchical approach. The former uses a clustering division among nodes, while the second one makes use of active probes which introduce some intelligence within the network. The work presented in [3] should be also mentioned, since it defines a complete information model, describing and implementing a prototype of a probe-based management architecture. One common aspect of these three previous works is that they analyzed both the information and the organization management models. There exist other works focusing their contribution on the organization model, paying special attention to management architectures based on a combination of both distributed and hierarchical approaches. Among these, some relevant works to be mentioned are [5], [6], [7], which are mainly based on specific clustering techniques, or [8], which analyzes how to optimally distribute management operations within the network.

The paper is structured as follows: Section II introduces relevant aspects about the organizational model for a management architecture, tailored to the specific characteristics of personal network environments. Section III discusses how manager and agent roles might be distributed between the network entities, identifying the parameters which should be taken into account so as to assess the appropriateness of the selection. In addition, different manager selection strategies are presented, ranging from rather pessimistic to almost optimum situations. Since the performance of those different strategies heavily depends on the particular characteristics of the network topologies, Section IV discusses the connectivity of random network deployments, paying special attention on the number of connected components (subgraphs) which might be expected for a particular network configuration. Afterwards, Section V, using an extensive analysis conducted over a proprietary simulator, evaluates the aforementioned strategies, discussing their benefits and disadvantages. Finally, Section VI concludes the paper, advocating some items which are left for future works.

2 Distributed/Hierarchical Management of Personal Networks

It is now believed that future personal networks will bring about some new requirements, including a much more relevant dynamicity, regarding their spontaneous configuration and self-managed topology.

Hence, in order to ensure an effective management of future personal networks, usually comprising multi-hop topologies and having a significant number of heterogeneous and mobile nodes, we propose using the adaptive and decentralized management paradigm. In this sense, the drawbacks of traditional management systems, characterized by centralized and static organizational models, are overcome.

Almost all current management frameworks are based on the traditional manager/agent model, which usually assumes a centralized control scheme, clearly inadequate to manage multi-hop or mesh topologies, since relying on a unique central entity to manage the whole network is not a sensible alternative, provided that nodes are intermittently connected. In such type of configurations, the manager would represent a single point of failure, and this is not admissible in future personal networks.

Another well known limitation of this centralized approach is its almost non-existing scalability; in this sense, having a single manager station might lead to heavy management traffic exchanged between this and the corresponding agents, as well as a high processing load on this particular node, which could lead to long execution times for management operations. This becomes even worse over personal networks; as it is assumed that the relevance of management operations (accounting, security, etc) would be quite high over such networks, and thus, the resulting overhead might heavily augment, decreasing the network performance. To sum up, we may say that the intrinsic characteristics of personal networks and multi-hop/mesh topologies, such as temporary links, sparse bandwidth and limited resources, impose a different approach rather than a traditional centralized architecture for the management framework.

There have been already some proposals aimed at overcoming these drawbacks, proposing some refinements to the basic centralized scheme. On of the most relevant ones is the so-called Management by Delegation (MbD) approach [9], which fosters the delegation of some tasks from the management station to the corresponding agents by means of dynamic downloadable scripts. Other extensions of this centralized scheme are based on the establishment of certain agent hierarchies, enabling direct interactions between agents. Nevertheless, these refinements (which are still based on a centralized approach) do not efficiently address all the previously enumerated shortcomings; on the contrary, the use of management architectures based on a decentralized operation and relying on self-management mechanisms could become a more sensible choice. Following this idea, in this work we present an organizational model based on a distributed and hierarchical model.

The proposed management framework is logically structured according to a three-level hierarchy, composed by a top level manager, which could be selected from a number of second-level managers. These take a localized manager role, controlling a set of nodes which could be seen as a cluster (characterized by some sort of connectivity between its components). The agents are thus the third level of the hierarchy. In spite of the three defined levels, there are only two management communication

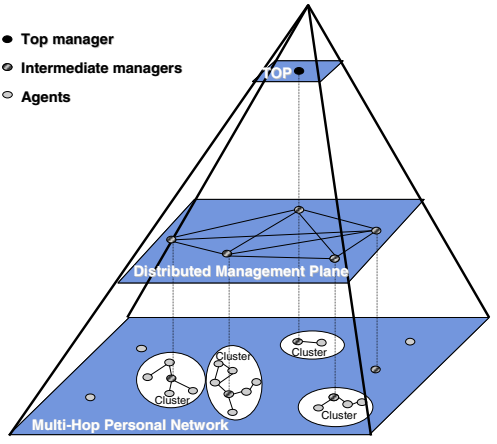


Fig. 1. Distributed/hierarchical management organization model for personal networks

planes: one established between the agents and their corresponding manager (second level); and another one which interconnect all the second level managers between them and with the overall manager (first level). As can be seen on Figure 1, this manager plane creates an overlay network on top of the PN which interconnects the second level managers by means of virtual links, each of which might correspond to a particular path, which could comprise several physical links, either over the underlying network or using another parallel communication facility.

The proposed architecture entails a distributed management plane (overlay network) with a number of nodes which take a manager role, each of them would control a subnetwork (or network component), communicating with the rest of managers in a peer to peer and collaborative manner. Using this distributed approach, the network management subsystem would achieve higher reliability and efficiency, as well as less overhead, on both communication and system resources.

In addition to that distributed plane, the management architecture also presents a hierarchical approach, since the manager role is distributed between two different levels: the top one represents overall network manager, while the 2nd level managers can be referred to as intermediate managers. Each of them controls its own domain (network component or cluster), collecting and processing information from the corresponding agents and forwarding such data to the upper level manager, if necessary. It also delivers management information from the top manager to its own domain nodes. As can be seen, the depicted management framework follows a distributed/hierarchical organization model.

3 Problem Statement

For the rest of this section we will assume that the network comprises a set of N nodes, M of which take the manager role while the resulting A (N-M) are agents. In addition, we will assume that the number of covered (able to access, at least, one manager) agents is A_C.

We have seen in the previous section that most of the sensible choices for the management architecture to be employed over a wireless multi-hop topology (as the one which we could map on a personal network) promote the balancing of the corresponding manager load between a set of nodes, thus bringing about hierarchical/distributed management architectures. Hence, the problem to be addressed is how to select this optimum set of managers. In order to be able to assess the suitability of the selection, we need to establish a number of aspects of merit. Their combination is what would yield the optimum selection strategy. Below we discuss three of the most relevant ones, which are the ones that will be later evaluated.

- **Covered probability.** this is the most obvious one, and refers to the percentage of nodes which are able to access a manager and, thus, can participate in the management architecture. The goal would be to reach a full coverage, meaning that all nodes are either managers or are covered by, at least, one manager.
- **Average number of hops.** one of the classical problems associated with multi-hop networks is the interference that communications can cause. In order to avoid a serious increase of the overhead brought about by the management traffic, it would be interesting having a small number of hops between each node and its corresponding manager.
- **Agent distribution.** the main goal of balancing the management burden is to avoid concentrating too much traffic into a single node; in this sense, the selection of managers should try to provide a fair distribution of the agents between them. In order to measure this aspect, we introduce the following parameter:

$$\beta = \frac{1}{M}\sum_{m}\frac{\left|A_m - \frac{A_C}{M}\right|}{\frac{A_C}{M}} = \frac{1}{A_C}\sum_{m}\left|A_m - \frac{A_C}{M}\right| \tag{1}$$

In the previous expression, AC accounts, as was already said, for the overall number of covered agents, while Am is the number of agents which are covered by the particular manager 'm'. Therefore the β parameter can be described as the relative difference between the optimum distribution (in which all managers have the same number of associated agents, AC/M) and the current one. The lower this parameter is, the closer it is with the fairest distribution.

In addition to the three parameters described before, we will also look into the combination of the first two, by studying the probability of being managed when a limit on the number of hops to be used to reach a manager is assumed.

The aforementioned figures of merit will be evaluated and studied using different manager selection strategies. Below, a description of each of them is provided.

- **Strategy 1 Ramdom manager deployment.** In this case, we assume that the M managers are randomly selected, without any kind of previous planning. This reflects a quite unlikeable situation since, depending on the network characteristics, there might be managers without any node within their coverage area. From an implementation point of view, this option poses no major difficulties.
- **Strategy 2: Topology-agnostic optimal manager deployment.** In this case, we assume that managers are placed in those points which ensure a maximum

"geographical" coverage of the whole area. This does not mean that the number of covered agents is maximized, since this would depend on their particular position in any particular network instantiation.

- **Strategy 3: Topology-aware optimal manager deployment.** In this case, the current topology of the network is used to optimally assign the managers. In order to accomplish this, we solve the p-median problem [10]. The p-median aims at finding a set of p managers from the overall deployed nodes which minimizes:

$$\sum_{j \in N} \{\min_{i \in M} d_{ij}\} \tag{2}$$

In our case, and without loss of generality, we have used a constant cost per link, so the overall cost between nodes i and j can be seen as the number of hops between them (provided that a path exists between them). Hence, the p-median problem would aim at minimizing the overall distances needed by all nodes to reach a manager; provided that all agents are covered (all the demand is satisfied).

- **Strategy 4: Topology-aware sub-optimal manager deployment.** As briefly introduced before, one of the main drawbacks which is usually attributed to the p-median problem is that its first goal is to cover all nodes [10] (i.e. all the demand should be covered). Depending on the network topology, there might be situations in which it could be more appropriate not managing some nodes, if that jeopardizes the rest of the assignment strategy. In order to better resolve this issue, we propose a simple heuristic, in which we do not account for those subgraphs having a reduced number of nodes, resolving the p-median problem only over the rest of the network. In this case, an additional design parameter is the size of the subgraphs to be discarded. A trade-off must be done between the nodes which are lost and the additional benefit which can be obtained by letting them out of the management architecture. Clearly, deleting isolated nodes is a sensible option, since those nodes are not able to communicate with the rest anyway. However, it has to be discussed whether it is sensible deleting subgraphs with a greater number of nodes.

Figure 2 aims at providing an illustrative example of the behavior of the four strategies. In this case, the network consists of 50 nodes, and 7 managers are deployed. Furthermore, Table 1 presents the particular figures which were obtained for each of the parameters which are being analyzed. The two first strategies (top part of the figure) are rather straightforward. First, the random deployment leads, in this particular case, to a relevant number of nodes which are left uncovered, while there are some managers which are unnecessarily close between them.

On the other hand, the second strategy, characterized by an optimal geographical placement of the 7 managers, ensures a high coverage, although it might lead to having some uncovered nodes, as those which are between the two lowest managers. In order to discuss the other two strategies, it is worth noting that the particular network example which has been used yields seven subgraphs, two of them consisting of isolated nodes. In this case, the advantages and drawbacks of the two latter manager selection approaches are clear: strategy 3 offers (for this particular network deployment) full coverage, since it devotes a manager to each of the existing subgraphs; however, the distribution is not very fair, since there is one manager which needs to handle a large number of nodes, which require, in some cases, a larger

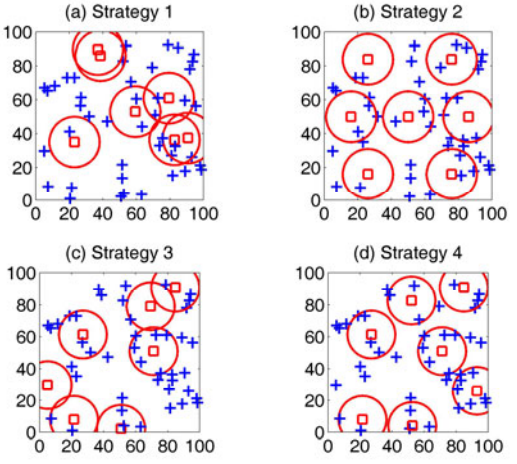


Fig. 2. Manager selection strategies. Squares represent managers (M=7) and crosses represent agent nodes.

Table 1. Summary of manager selection strategies results

Strategy	β	Avg_hops	Disconnected	2hops_covered
1	0.661	2.143	0.349	0.465
2	0.336	1.579	0.116	0.837
3	0.963	1.977	0.000	0.721
4	0.445	1.341	0.047	0.907

number of hops to reach the corresponding manager. Besides, strategy 4, in which the two isolated nodes are not considered, fails to provide a full coverage, but, on the other hand, it ensures a fairer distribution of the nodes between the available managers, thus leading to routes with a fewer number of hops. As can be seen on Table I, the connectivity is higher for strategy 3, but if we limited the maximum number of hops to reach a manager to 2, then the covered probability of strategy 4 clearly outperforms the p-median based one.

4 Connectivity of Random Network Deployments

Since one of the main goals of the analysis is to establish the number of managers which should be deployed (for the different analyzed strategies) in order to guarantee a certain connectivity degree, it is worth analyzing the connectivity of the subjacent network scenario. This would provide the information required to better understand the corresponding results. In this sense, we introduce a connectivity degree parameter (ζ), which accounts for the number of subgraphs (SG) of a particular deployment of N nodes.

$$\xi = \frac{N - SG}{N - 1} \tag{3}$$

From the above expression, it is straightforward that in a fully connected network ζ equals 1, while for a network in which there is not any connection amongst the nodes, ζ equals 0. E.g. the connectivity of the network ranges from 0 (all nodes are isolated) to 1 (there is, at least, one possible path between any pair of nodes within the network). Furthermore, and in order to be as generic as possible, we define the normalized coverage (ρ), as the ratio between the particular coverage of the Radio Access Technology and the side of the area under analysis (this allows us carrying out the most generic analysis as possible). In that sense, if we maintain the number of nodes (N) and ρ constant, the connectivity (as it has been defined) should not change, no matter the actual dimension of the area under analysis is. This means that it is possible establishing a function g so that ζ = g(N, ρ) (without depending on the particular area dimension). However, in many cases, network deployments are characterized by node density (D) and the coverage (R) of the subjacent technology. It would be also appropriate being able to establish a similar function, taking D and R as its arguments. Let's assume that we have a squared area (side Li), then:

$$\begin{aligned} D_i &= \frac{N}{L_i^{\,2}} \\ R_i &= \rho \cdot L_i \end{aligned} \tag{4}$$

By inspection, we can see that, provided we maintain N and ρ constants, the product of D and R^2 is also kept constant. In fact:

$$D_i R_i^{\,2} = \frac{N}{L_i^{\,2}} \rho^2 L_i^{\,2} = N\rho^2 = \Lambda \tag{5}$$

Hence, if we define the Λ parameter, as the product of N and ρ^2, or, equivalently, as the product of D and R^2, we should seek for a function f such that ζ = f(Λ).This result would be of outer relevance, since it would allow a quick estimation of the particular network configuration in order to reach a required connectivity degree; this might be quite beneficial, e.g. when deploying wireless sensor networks to cover a particular area.

Figure 3 shows the connectivity degrees which were obtained for the different (N, ρ^2) pairs, after a simulation analysis which consisted of 400 different scenarios, and 1000 independent runs per scenario. As it was expected, it is possible establishing the following relationship between the two parameters:

$$N\rho^2 = \Lambda \;\rightarrow\; \log(N) + \log(\rho^2) = K \tag{6}$$

Where Λ and K are constants and must be functions of ζ, which also linearly depends on log (N) + log (ρ^2). Hence, the problem is to find the relationship between Λ and ζ.

Before that, Figure 4 shows the standard deviation which was observed for the 1000 independent runs which were executed per scenario. As can be seen, the variation of the results is higher (around 0.15) for network deployments having fewer nodes; however, the standard deviation quickly decreases and the behavior for relatively large values of Λ is pretty much predictable.

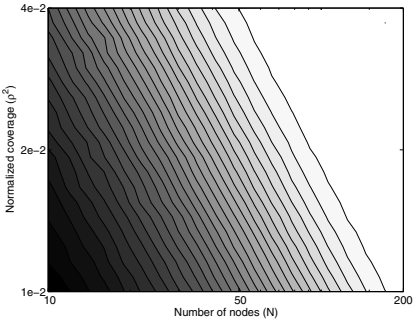


Fig. 3. Connectivity degree (ζ) for various (N,.ρ^2) pairs. ζ ranges from 0.13 (darker) to 1.00 (lighter).

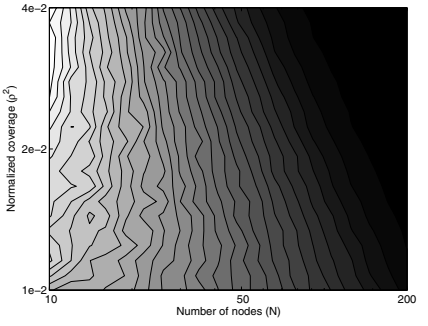


Fig. 4. Standard deviation of the connectivity degree (ζ) for various (N,ρ2) pairs. STD(ζ) ranges from 0.00 (darker) to 0.15 (lighter).

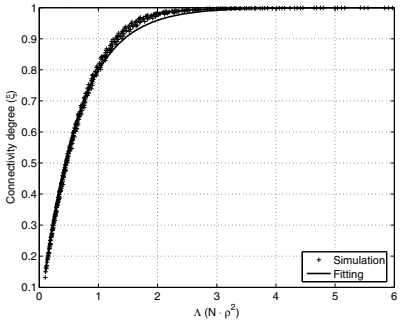


Fig. 5. Relationship between the network topology (Λ parameter) and its connectivity (ζ)

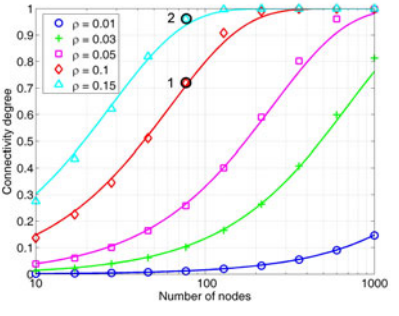


Fig. 6. Connectivity degree (ζ) for various node densities and normalized coverages (ρ)

The conclusion is that the estimation of the connectivity degree can be accurately estimated based on the particular characteristics of the network deployment (number of nodes, coverage of the radio access technology and dimensions of the area under analysis).Now the question is to find the expression that better fits the results which were obtained in the different simulation runs. Figure 5 shows the different (Λ,ζ) pairs which were observed during the Montecarlo procedure, as well as the exponential function which better fits them (LMSE is less than 0.02). The resulting parameter of such function is 1.6, so we can write:

$$\xi \approx 1 - e^{-1.6\Lambda} = 1 - e^{-1.6N\rho^2} = 1 - e^{-1.6DR^2} \tag{7}$$

Figure 6 shows the connectivity degree for various values of the normalized coverage. The lines have been obtained with Eq. [7], while the markers are empirically acquired with a Montecarlo simulation comprising 1000 runs per scenario. As can be seen the values which are provided by the proposed expression are rather close to the real ones, assessing the validity of such expression to estimate

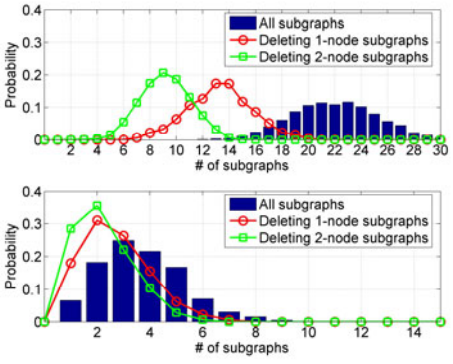


Fig. 7. pdf of the number of subgraphs for a network of 80 nodes over the two selected scenarios (top figure for ζ=0.7 and bottom graph for ζ=0.9)

the connectivity degree. This allows us to better select the parameters of the scenarios over which we will analyze the manager selection strategies. In this sense, in order to have a more thorough comparison of the different manager selection strategies, we select two illustrative network deployments (see Figure 6): (1) a sparse network, in which ζ is around 0.7, and (2) a high-connected scenario, in which ζ is higher than 0.95. We ensure these two scenarios by fixing N to 80 and using two normalized coverages: 0.10 (Λ = 0.8) and 0.15 (Λ = 1.8), respectively.

The last step before discussing the obtained results is to establish the design parameter of the fourth algorithm. In that sense, Figure 7 shows the probability distribution function (pdf) of the number of subgraphs which were encountered over the two scenarios which will be used during the evaluation of the different alternatives.

The number of subgraphs which are reflected on Figure 7 corresponds to those which could have been expected by using Eqs. [3] and [7]. First, in the case of the connected network, Λ equals 1.8, which yields a connectivity of ζ = 0.95, thus resulting in an average 4.95 subgraphs. On the other hand, for the sparse network, in which Λ equals 0.8, the connectivity is, according to Eq. [7], around 0.72, which yields 23.1 subgraphs, corresponding with the histogram which is represented in Figure 7. It has to be mentioned that, according to the differences between simulated and analytical values, the approximation is slightly better in the case of the sparse network.

Besides, the figure also permits inferring the potential advantages when discarding components of a particular size. As can be seen, by cutting components of 1 and 2 nodes, the additional benefit is rather notable, since with 5 managers the whole network will be covered in more than 90% of the cases for the high connected scenario. On the other hand, for the sparse case, since it is quite likely to have subgraphs of reduced size, there is a clear benefit when leaving out of the analysis those subgraphs (as can be yielded from the graph, the decrease on the number of resulting components is much more relevant than in the previous scenario). Furthermore, by cutting those network chunks, we ensure that the size of the remaining subgraphs would be much bigger and the p-median problem might provide better

results, especially in terms of the number of hops which are required to reach a manager, since more managers will be available for the remaining subgraphs (this is true for the two scenarios). Hence, we will assume that if the number of nodes which are reachable is fewer than 3, then a node would not take the manager role (in fact, it will not belong to the managed network).

5 Discussion of Results

This section discusses the results which were obtained when applying the previously defined algorithms to assign the manager role, based on an extensive simulation campaign. A Montecarlo method was applied, and the metrics which were previously described were analyzed.

The simulation setup was based on a proprietary simulator, while the p-median problem was solved with the popstar tool [9]. Since the main goal of this work is to analyze the manager selection strategies, an ideal propagation channel was assumed. Incorporating more realistic models would not add too much complexity to the simulator, since an average coverage might be used in such cases; in addition, by pursuing this approach, we ensure that the analysis could be easily extended to other type of scenarios, e.g. wireless sensor networks. We assume two simulation areas, over which we randomly (Poisson Point Process) deploy N=80 nodes. The first scenario (sparse network – $\zeta \approx 0.7$) is characterized by a normalized coverage of ρ=0.1, while in the second one (high connected network – $\zeta > 0.95$) the normalized coverage is ρ=0.15.

The next step was to select the managers, following the four previously discussed algorithms; afterwards, an analysis of how the agents would be distributed between them was performed. In this sense, we will study the additional benefit of increasing the number of managers for all cases, i.e. M will be increased from 1 to 10. Each individual setup was simulated 200 times, in order to get tight confidence intervals.

Figure 8 shows the probability for an agent to be covered by, at least, one manager. As can be seen, the worst manager deployment strategy leads to the smallest coverage probability. On the other hand, there is not a remarkable difference between strategies 3 (legacy p-median) and 4 (cutting the 1-node and 2-node network subgraphs), staying below 3% for all cases. The effect of deleting such nodes appears in the high connected scenario, as the coverage probability asymptotically leads to a value slightly smaller than 1 for the fourth strategy, while a full coverage is almost reached with 5 managers for the legacy p-median case. The results also show that the topology-agnostic approach does not reach the same coverage as the one which can be achieved by means of the p-median based algorithm. Last, it is worth referring to the benefit which is achieved over the sparse network with strategies 3 and 4; in this case, the difference between these two manager selection algorithms is even less relevant than the one which is observed for the high connected network.

One of the benefits that the proposed heuristic should have over the legacy p-median is that it should achieve a fairer distribution of the agents between the different managers. Figure 9 shows how the different strategies distribute the management burden between the selected managers, representing the β parameter for

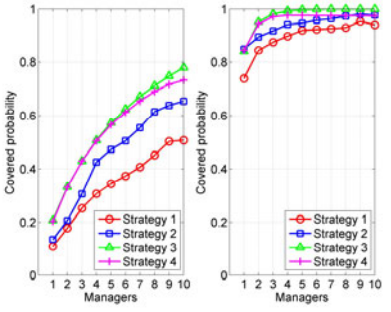


Fig. 8. Coverage probability for the four manager deployment strategies and the two connectivity scenarios (left ζ=0.7 and right for ζ=0.9)

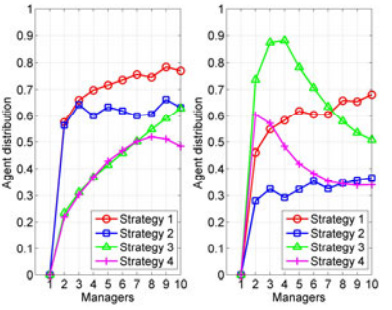


Fig. 9. Distribution of agents between managers for the four manager deployment strategies and the two connectivity scenarios (left ζ=0.7 and right for ζ=0.9)

the four algorithms. We shall observe an important result for the high connected network, since the third strategy (applying the p-median over the whole network) provides similar performance than the worst one (in terms of coverage probability, i.e. strategy 1), at least when the number of managers is reasonably low. This is the consequence of reserving managers to cover all subgraphs, no matter their size is. However, with the proposed modification (strategy 4), the distribution is strongly improved, and it almost reaches the behavior of the best solution (topology-agnostic, strategy 2) when the number of managers is greater than 4. For the sparse network, there is not a clear difference between strategies 3 and 4 (providing, both of them, a better performance than the others), although the graph shows that the tendency of the measured parameter tends to decrease for $M > 8$ for strategy 4, while for the legacy p-median approach the parameter still shows a growing trend.

The third aspect which was identified as a parameter to be optimized is the average number of hops which are required to reach a manager. The fewer the needed hops, the less overhead which would be caused by the management traffic. As can be seen in Figure 10, the heuristic presented in this paper again shows a very similar behavior than the second strategy for the high connected scenario, being slightly better for more than 4 managers. Obviously, the random selection approach provides the worst results, while for the p-median case, the graphs yields that for a small number of managers, the analyzed parameter is rather high (almost reaching the values obtained for the worst case); this is due to the fact that the p-median main goal is to cover all nodes (as mentioned earlier) and thus, it establishes managers in all subgraphs, with the consequence that in those ones in which the number of nodes is greater, more hops are needed to reach the selected manager, since there are fewer managers to serve those network components. For the sparse network, the proposed heuristic also shows the best performance, although in this case it is comparable with the legacy p-median. However, when the number of managers is higher than 8, there is a change in the corresponding trends (the benefit of the novel heuristic starts to be more relevant), due to the fact that the average number of resulting subgraphs is between 8 and 9, as was shown in Figure 7.

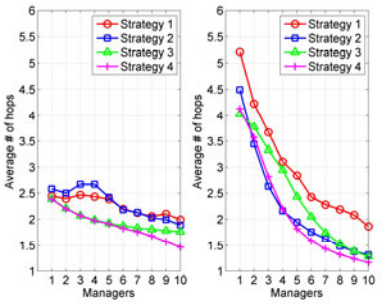


Fig. 10. Average number of hops required to reach a manager for the four manager deployment strategies and the two connectivity scenarios (left ζ=0.7 and right for ζ=0.9)

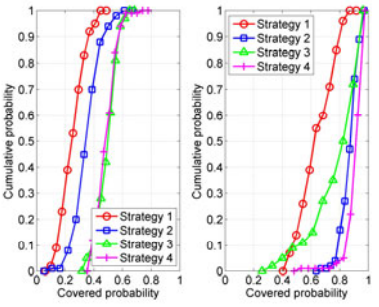


Fig. 11. cdf of the probability to reach, at least, one manager when the maximum number of hops is set to 3. The number of managers to deploy (M) is set to 5. Left figure: sparse network (ζ=0.7) and right figure (high connected network (ζ=0.9).

One conclusion which could be derived by analyzing the results of both Figures 8 and 10 is that, if a bound was set on the maximum number of hops which could be used to reach a manager, the behavior of the proposed scheme might be even much better than the rest of analyzed strategies. As can be seen in Figure 11, which shows the cumulative distribution function (cdf) of the probability of being connected to a manager using a maximum of three-hop routes (and fixing the number of managers to deploy, M, to 5), the proposed solution is the one which offers the best performance for the high connected network, since the probability of being connected with routes of either one or two hops is much higher with this strategy. In this case it is worth highlighting the improvement compared to the traditional p-median. Although this solution was able to provide a higher coverage (since it did not cut any node from the network), this benefit would not be perceived if a limit was established on the maximum number of hops which could be used to reach a manager. For the sparse network scenario, the difference is not that remarkable, since in this case we might need to deploy more managers to start noticing the improvements brought about by the proposed approach, as it has been discussed above.

The previous results show that the proposed heuristic provides a quite good performance for selecting managers on a hierarchical/distributed management architecture which could be used over forthcoming personal network scenarios. The tests carried out with two illustrative network deployments prove that this behavior can be extrapolated to other scenarios.

Another important aspect to be mentioned is that, although the analysis which has been performed is fundamental, we have kept an implementation perspective in mind, and e.g. the size of the subgraphs to be cut from the network would allow to bring the algorithm to a straightforward implementation, since the required information may be acquired by means of neighboring discovery processes, which are expected to be part of any personal network system.

6 Conclusions

This paper has tackled the management of personal networks. Opposed to the traditional centralized model, we have proposed a distributed/hierarchical architecture, in which the manager burden is shared between a number of nodes, thus making the whole system more scalable and alleviating the overhead associated to management traffic. These managers are together joined in an overlay network, used to interchange management related information. Furthermore, an optional higher level may be added, since it may be interesting that some individual nodes are able to gather all the information to be managed.

One of the most challenging issues which arise in the previously depicted architectures is how to carry out an appropriate selection of the manager roles. This work has specifically looked into this problem. We have identified different metrics which should be analyzed to assess the appropriateness of the selection, accounting for how a fair sharing of the management burden is fulfilled, the probability for any node to participate within the management subsystem and the resulting overhead (based on the number of hops which are required to communicate with the respective manager). Based on these metrics, different strategies have been analyzed: a rather pessimistic algorithm, in which the managers are randomly deployed, and two optimum approaches, based on the geographic position of the manager nodes and on the particular network topology. For this latter case, an enhancement, based on a novel heuristic, has been proposed, by eliminating those network components (subgraphs) with a relatively small number of nodes.

The results have yielded two main conclusions: on the one hand, it is of outer relevance to facilitate an appropriate selection of the managers, since there might be relatively large differences depending on the particular selection strategy; in addition, the proposed heuristic provides very interesting results, since it brings about much better performance in terms of agent distribution between the selected managers, as well as considering the number of hops which are used to reach the corresponding manager, while it does not severely degrade the coverage probability. In order to account for a wide range of potential scenarios, the analysis has been conducted using two illustrative network deployments.

The research which has been conducted in this work has been based on network graphs, but a major consideration has always been the possibility of mapping the proposed algorithms on top of real protocols and network deployments. In this sense, and thanks to the management simulation framework which was presented in [12], future work will analyze the implications of the proposed management architecture as well as the manager selection architectures on the communication and protocol performance. Regarding this future line of research, there are some interesting works available in the literature, mostly dealing with wireless sensor networks [13-14].

Acknowledgments

The authors would like to express their gratitude to the Spanish government for its funding in the following two projects: Mobilia - CELTIC Program (Avanza I+D TSI-020400-2008-82) and "Cognitive, Cooperative Communications and autonomous SErvice Management", C3SEM (TEC2009-14598-C02-01).

References

1. Lu, W., Gu, Y., Prasad, R.V., Lo, A., Niemegeers, I.: A Self-organized Personal Network Architecture. In: Third International Conference on Networking and Services, ICNS 2007, June 19-25, pp. 36–36 (2007)
2. Shen, C.-C., Srisathapornphat, C., Jaikaeo, C.: An Adaptive Management Architecture for Ad Hoc Networks. IEEE Communications Magazine 41(2), 108–115 (2003)
3. Chen, W., Jain, N., Singh, S.: ANMP: Ad Hoc Network Management Protocol. IEEE Journal on Selected Areas in Communications 17(8), 1506–1531 (1999)
4. Badonnel, R., State, R., Festor, O.: Management of Mobile Ad Hoc Networks: information model and probe-based architecture. International Journal of Network Management 15(5), 335–347 (2005)
5. Sivavakeesar, S., Pavlou, G., Liotta, A.: Stable Clustering Through Mobility Prediction for Large-Scale Multihop Intelligent Ad Hoc Networks. In: Proc. of WCNC 2004, Atlanta, USA (March 2004)
6. Fallon, L., Parker, D., Zach, M., Leitner, M., Collins, S.: Self-forming Network Management Topologies in the Madeira Management System. In: Bandara, A.K., Burgess, M. (eds.) AIMS 2007. LNCS, vol. 4543, pp. 61–72. Springer, Heidelberg (2007)
7. Badonnel, R., State, R., Festor, O.: A Probabilistic Approach for Managing Mobile Ad Hoc Networks. Transactions on Network and Service Management 4(1), 39–50 (2007)
8. Lim, K.-S., Adam, C., Stadler, R.: Decentralizing Network Management. KTH Technical Report (December 2005)
9. Yemini, Y., Goldszmidt, G., Yemini, S.: Network Management by Delegation. In: Second International Symposium on Integrated Network Management IM 1991, Washington, D.C., pp. 95–107 (April 1991)
10. Resende, M.G.C., Werneck, R.F.: A hybrid heuristic for the p-median problem. Journal of Heuristics 10(1), 59–88 (2004)
11. Canós, M.J., Ivorra, C., Liern, V.: Fuzzy p-median problem: A global analysis of the solutions. European Journal of Operational Research 130(2), 430–436 (2001)
12. Irastorza, J.A., Agüero, R., Muñoz, L.: Fostering the simulation-based evaluation of management architectures over multi-hop topologies. In: IEEE/IFIP Network Operation and Management Symposium (NOMS 2008), Salvador do Bahia, Brazil (April 2008)
13. Ruiz, L.B., Silva, F.A., Braga, T.R.M., Nogueira, J.M.S., Loureiro, A.A.F.: On impact of management in wireless sensors networks. In: IEEE/IFIP Network Operations and Management Symposium, NOMS 2004, April 19-23, vol. 1, pp. 657–670 (2004)
14. Ruiz, L.B., Nogueira, J.M., Loureiro, A.A.F.: MANNA: a management architecture for wireless sensor networks. Communications Magazine 41(2), 116–125 (2003)

OLSRp: Predicting Control Information to Achieve Scalability in OLSR Ad Hoc Networks

Esunly Medina[1], Roc Meseguer[1], Carlos Molina[2], and Dolors Royo[1]

[1] Dept. of Computer Architecture, Universitat Politècnica de Catalunya, Spain
{esunlyma,meseguer,dolors}@ac.upc.edu
[2] Dept. of Computer Engineering, Universitat Rovira i Virgili, Spain
carlos.molina@urv.net

Abstract. Scalability is a key design challenge that routing protocols for ad hoc networks must properly address to maintain the network performance when the number of nodes increases. We focus on this issue by reducing the amount of control information messages that a link state proactive routing algorithm introduces to the network. Our proposal is based on the observation that a high percentage of those messages is always the same. Therefore, we introduce a new mechanism that can predict the control messages that nodes need for building an accurate map of the network topology so they can avoid resending the same messages. This prediction mechanism, applied to OLSR protocol, could be used to reduce the number of messages transmitted through the network and to save computational processing and energy consumption. Our proposal is independent of the OLSR configuration parameters and it can dynamically self-adapt to network changes.

Keywords: mobile ad hoc networks, prediction, energy-aware.

1 Introduction and Motivation

A Mobile Ad hoc Network (MANET) is an autonomous and decentralized system formed by a collection of cooperating nodes that are connected by wireless links. They can dynamically self-organize and communicate between themselves in order to set up a network without necessarily using any pre-existing infrastructure.

Ad hoc routing protocols can be classified according to the combination of two different sets of characteristics: reactive or proactive combined with link state or distance vector. The MANET working group from the Internet Engineering Task Force (IETF) has proposed Optimized Link State Routing (OLSR) [2] as a standard link state proactive routing protocol for MANETS. In a link state routing protocol, a node periodically broadcasts the list of its neighbors over the network. Consequently, when operating normally, every node has information about all the other network nodes' neighbors. Therefore, a straightforward algorithm can compute the whole network topology, and thus we have all the routes and the shortest path to every destination. Proactive protocols maintain fresh lists of destinations and their routes regardless of whether data needs to be transferred or not.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 225–236, 2011.

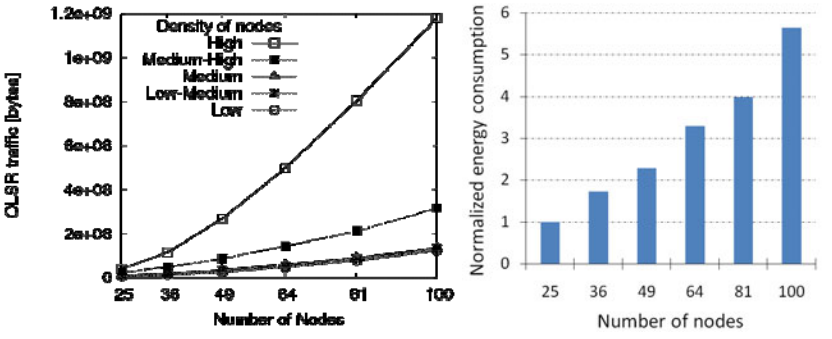
Link state proactive protocols allow lower latencies when sending data through the network because an optimized data path to the destination is already known. However, this comes at the cost of periodically flooding the routing information to all nodes in the network. When the number of nodes is large the amount of routing information to be sent is such as that it can overload the network, in this situation the system does not scale. Disseminating the routing information in order to reduce the overhead generated is essential to ensure that a protocol scales.

The overhead generated by sending the routing information follows the DQ principle [1], where Q stands for Queries and D for Data size. When applied to routing protocols, Q corresponds to the number of routing information packets that are sent to the network and D is the size in bytes of these packets. A system is perfectly scalable if DxQ remains constant when the number of nodes increases. However, when the number of nodes increases in a mobile ad hoc network, the DxQ coefficient also increases. In [7] and [8], the mechanisms described to make routing protocols more scalable focus on reducing Q, D or both. For instance, the FSR protocol decreases Q, sending the entire link state information only to neighbors instead of flooding it throughout the network; the OLSR protocol with Multi-Point Relays (MPRs) manages to reduce the number of ”superfluous” broadcast packet retransmissions (thus decreasing Q) and also to reduce the size of the link state update packets (thus decreasing D); the TBRPF protocol decreases D by sending periodically ”differential” messages that report only the changes of neighbors; and finally, the HOLSR decreases Q and D by proposing a dynamic clustering mechanism so that the OLSR can increase scalability.

This paper proposes a new mechanism that increases the scalability of link state proactive routing algorithms. In our proposal, all nodes responsible for disseminating the routing information have a very simple software predictor, so that if a message that is to be sent contains the same routing information that has just been posted in a previous message (i.e. if the network topology remains unchanged), then the message is not sent. If a node does not receive the packet with routing information, it assumes that the routing tables have not changed and does not recalculate paths, thus saving computational processing and energy consumption. It is important to notice that our mechanism is independent of the OLSR configuration (HELLO and TC emission intervals). That means that OLSRp does not modify the number of TC messages that are processed but it reduces the amount of TC messages transmitted through the network (those messages that are not transmitted are predicted by the receiver). Consequently, OLSRp dynamically self-adapt to network changes (OLSRp behaves exactly like OLSR but only if network changes occur).

Our proposal targets scalability by reducing Q. Whereas other proposals try to reduce Q by defining a hierarchy of nodes with different roles, only some of which send routing information to the network, we propose a mechanism where all the nodes have the same role, which simplifies network management. Moreover, in all the other mechanisms, the nodes involved in disseminating routing information always send routing information even when the network topology remains unchanged. Our approach only disseminates routing information if the network topology changes.

To evaluate the potential benefits of our proposal, we analyzed the degree to which the OLSR protocol repeated control packets and consumed node energy. Our proposal had two advantages:



(a) (b)

Fig. 1. (a) OLSR traffic and (b) Energy consumption versus number of nodes

- It reduces network collisions because the predictor only sends non-redundant routing control information, thus reducing the routing information traffic. Fig. 1.a shows clearly (for different node densities) that traffic generated by the OLSR protocol grows exponentially with the number of nodes. The following sections will show that a significant volume of this traffic contains redundant information.
- It reduces CPU processing time and energy consumption because fewer routing control packets are sent and received. This packet reduction is particularly interesting because the energy consumed by OLSR traffic increases with number of nodes (see Fig 1.b). Furthermore, the energy consumed by the OLSR protocol is a significant part of the overall energy consumption. For instance, our research in [15] shows that when commodity devices are used, the energy consumed by OLSR-protocol control traffic is a key concern. Moreover, in [3] a study of the energy consumption of several routing protocols shows that OLSR is one of the most energy-intensive consumers.

The results of this paper focus on the OLSR protocol, but we strongly believe that these results can be easily extrapolated to other protocols that need to deal with periodical control messages.

This paper makes the following contributions:

- It analyzes how much control information is repeated as a result of the OLSR.
- It proposes a transparent, cost-effective and energy-aware mechanism for reducing the control information produced by this protocol in order to achieve scalability.

2 Optimized Link State Routing Protocol

The OLSR [2] protocol is a well-known proactive routing protocol for ad hoc networks. It is an optimization of the Link State algorithm. The nodes in an OLSR network periodically exchange routing information to maintain a map of the network topology. The Multi Point Relays (MPRs) are the network nodes selected for propagating the topology information. The use of MPRs reduces the number and size

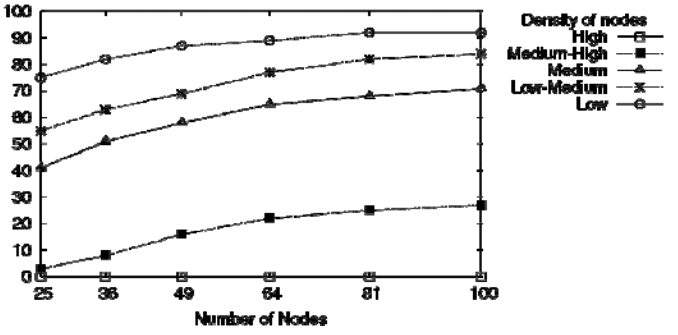


Fig. 2. Ratio of OLSR control messages corresponding with TC messages

of the messages to be flooded during the routing update process. In OLSR, there are two types of control messages: HELLO and Topology Control (TC).

HELLO messages allow each node to discover its neighboring nodes and to obtain information about the state of its links with them. In an OLSR network, every node periodically broadcasts HELLO messages to all its one-hop neighbors. By sending a HELLO message, a node identifies itself and reports its list of neighbors.

When an MPR receives HELLO messages, it records the list of nodes that have selected it as one of their MPRs (i.e. the Advertised Neighbor Set) and it generates a TC message, in which the MPR originator node announces its selectors. These routing update messages are relayed by other MPRs throughout the entire network, allowing every remote node to discover the links between each one of the MPRs and its selectors (note that the non-MPR nodes will receive and process the messages but will not retransmit them). Through this selective flooding mechanism, the MPRs retransmit and flood the whole network with TC messages. Fig. 2 shows the ratio of the total OLSR control messages corresponding with TC messages. When the distance between network nodes increases (i.e. low density), the percentage of TC messages also increases. It is also noticeable that the ratio of TC messages is very significant for network topologies with low node density. These results combined with the exponential growth trend of OLSR (shown in Fig.1.a) confirm that TC messages are an important part of the protocol traffic.

Each node maintains a routing table containing the information it receives periodically from the TCs and uses this to calculate the shortest path algorithm. In other words, a node calculates the shortest path to a given node using the topology map, which consists of all its neighbors and the MPRs of all other nodes and which it creates by means of the TC messages it receives. The routing tables of all nodes are updated every time a change in any link is detected. Fig. 3 shows the OLSR protocol operating in an ad hoc network with two MPRs. Every node periodically transmits HELLO messages to its one-hop neighbors and the nodes selected as MPRs are responsible for retransmitting the TC messages with the topology information.

A TC message field that is very significant for this research is the Advertised Neighbor Sequence Number (ANSN). This field is a sequence number that only increases its value if the Advertised Neighbor Set associated with a given MPR changes. Thus, every time the Advertised Neighbor Set of an MPR changes (i.e. when new nodes appear or existing nodes disappear), the MPR increases the ANSN value

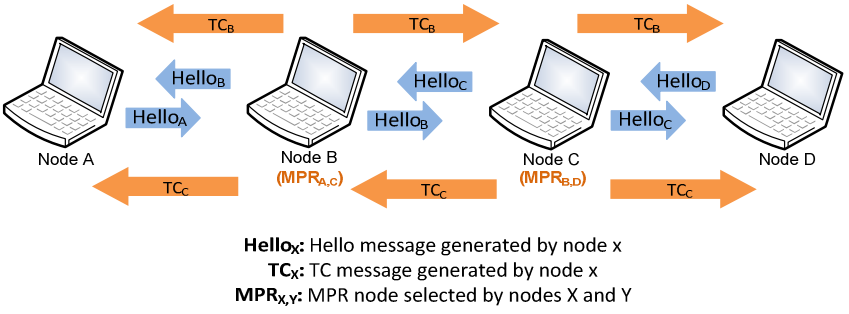


Fig. 3. MPR mechanism and control messages in OLSR

of its TC messages. When a node receives a TC message from an originator MPR, it can use this sequence number to determine whether the information about this MPR's advertised neighbors is more recent than the information that it already possesses. This mechanism allows a node to confirm whether the information it has received in the latest TC message is valid or not, that is, whether it has already received more a message with a higher ANSN value from the same originator node.

3 Experimental Setup

We have used ns-2 and ns-3 [12] simulators because this allows us to model several network scenarios and to collect statistics through the generation of PCAP files. Such simulation tools allow us, among others things, to define network topologies, configure wireless network interfaces and set node mobility patterns.

For our simulations, we assume an initial grid node distribution of N rows and N columns. This grid is initially set with nodes placed at a distance of D meters (delta distance) producing a box terrain of (N-1xD)x(N-1xD) meters. Fig. 4 summarizes the initial node distribution and the rectangle area assumed in our scenarios. Moreover, once a set of values for N and D has been obtained, all possible combinations can be evaluated. Finally, notice that we consider five delta distance values. That means that we assume, for a fixed number of nodes, five levels of node density (low, low-medium, medium, medium-high and high) that are derived from the size of the terrain in which they are deployed.

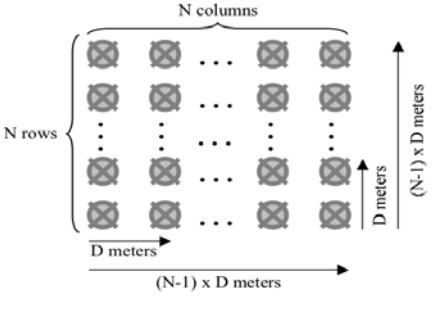


Fig. 4. Distribution of grid nodes

Each node is equipped with an 802.11b Wireless Network Interface Card operating at 2.4 GHz with a transmission rate of 1 Mbps and a coverage range of 500 meters. We also assume a Wi-Fi channel with a constant propagation delay and a Friis propagation loss model. Related to the OLSR protocol, we assume emission interval values of 2 and 5 seconds for HELLO and TC messages respectively.

The impact of node mobility is an important issue for our analysis of TC message duplication. We begin with a static (non-mobile) scenario and then assume a Random Direction 2D mobility model. This model deals with motion in random directions and forces nodes to reach the edge of the simulation area before changing their direction. Therefore, when a node gets to the boundary, it pauses and then selects a new direction and speed. We have considered scenarios with mobility and a fixed speed (meters/second) for all nodes involved in the simulation: 0.1 m/s (baby crawling speed), 1 m/s (walking speed), 5 m/s (running speed) and 10 m/s (car city circulating speed). We also fix the pause time of nodes to zero when they get to a boundary, because we are interested in the nodes moving continuously.

Finally, we generate application traffic that consists of several UDP packets transmitted every second, each of which is 100 bytes long. We also set half of the nodes to act as Echo servers and the other half to act as Echo clients.

4 Analysis of Control Information Repetition

In this section we quantify the amount of message repetition that is present in OLSR TC messages. We analyze this by considering the variables that we have already mentioned: mobility, number of nodes and delta distance.

The repetition that we want to quantify is based on which value was last observed. Consequently, we quantify the number of repeated TCs on the basis of whether the last message received is equal to the preceding one. To do so, every grid node observes the TC messages and quantifies the last value repetition (the overall results are presented in Fig. 5). Moreover, we distinguish messages on the basis of the generator node, that is, the node that creates the TC message. This means that every node has to store the last TC message sent by every neighbor to quantify repetition. Finally, in this study we focus on the ANSN field of the TC. If this field in the current TC matches the previous one, we consider that both messages are the same.

In static scenarios where all the nodes are always active, the results were as expected. We can state that 100% of TC messages are always the same. This changes for mobile scenarios. Fig.5 shows the percentage of message repetition observed in several mobility scenarios. From top-left to bottom-right, we present four figures that show behavior at four different speeds (0.1 m/s, 1 m/s, 5 m/s and 10 m/s). In each figure, the Y-axis shows the percentage of repetition, the X-axis shows the number of grid nodes and every line corresponds to a different node density. By looking at these figures, we can make the following observations regarding a mobility scenario.

The number of nodes does not affect the percentage of repetition. If we fix the speed of the node mobility and the node density, we observe that there are no significant differences when the number of nodes is increased. Notice that all the lines in each Fig. tend to be horizontal. That means that we can achieve the same percentage of repetition just by increasing the number of nodes. This result is also

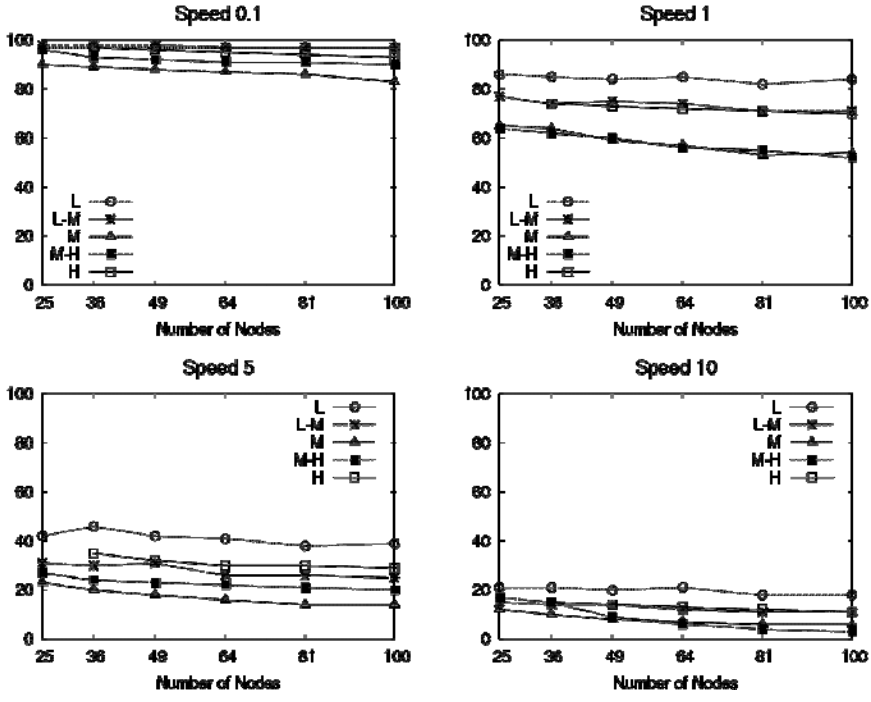


Fig. 5. Percentage of repetition under mobility scenarios

interesting in terms of scalability because our mechanism for reducing TC messages could be orthogonally applied independently of the number of nodes.

The percentage of repetition is significantly affected by mobility. We can observe that the percentage of last value repetition ranges from 80% to 98% when the speed is 0.1 m/s, from 40% to 80% when the speed is 1 m/s, from 20% to 40% when the speed is 5 m/s and, finally, from 5% to 20% when the speed is 10 m/s. This was expected because TC messages are generated every 5 seconds, which means, therefore, that when speed is increased the probability of topological changes during that period of time also increases.

The percentage of repetition is significant even with high node speeds. It was also expected that the percentage of repetition would remain high with low speed rates of mobility. In any case, this percentage of repetition is still significant at higher speeds (5% to 20% when speed is 10 m/s). This result is interesting because, as explained previously, the number of TC messages increases exponentially with the number of nodes (see Fig. 1.a). Therefore, even with low percentages of repetition, the amount of network congestion can be significantly reduced if we can provide a cost-effective mechanism to discharge the network of replicated TC messages.

The density of nodes affects the percentage of repetition. It can be observed in any of the four figures that when there is a given mobility speed and a fixed number of nodes, there are small differences between different levels of node density. This behavior is explained by the relationship between the number of MPRs and the

number of neighbors that a given node has. However, these results are interesting in terms of TC message reduction because they prove that our mechanism could be also applied to several scenarios independently of the density of nodes.

5 Our Proposal: OLSRp

We propose the implementation of a new mechanism for predicting OLSR control information: the OLSRp. This is a last-value predictor designed to be placed in every node of an OLSR ad hoc network. The purpose of this predictor is to prevent the MPRs from transmitting duplicated TC packets throughout the network. The OLSRp functions in the following manner:

A given MPR executes a prediction when it has a TC message to transmit. Because the OLSRp launches a Last-value predictor, the result of every prediction is always the last TC message generated by the MPR. Immediately after a prediction is made, the OLSRp compares the prediction result with the new TC message generated by the MPR. If both the predicted TC and the new TC message are the same, then the MPR does not transmit the new TC message. Because the OLSRp mechanism is installed in every network node and because all the nodes have the same Last-value predictor, the remaining nodes will also calculate the same TC message as that which was predicted by the original MPR. By making this prediction, we are able to reuse the same TC, thus preventing the transmission of duplicated TC messages and stopping changes from occurring to the network topology.

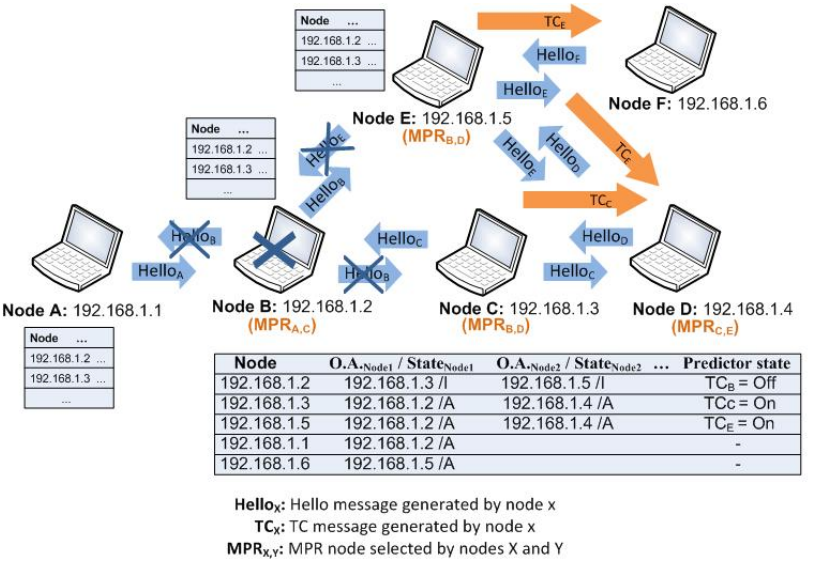
The OLSRp is 100% accurate because the prediction results are always correct (i.e. all the nodes expecting a given TC message will always predict the same TC message). When OLSRp can not make a prediction, a new TC message will be transmitted. However, it could be argued that although the proposed OLSRp is based on the certainty of its predictions, it does not take into account the fact that the destination nodes may not be properly working. In order to deal with this issue, the OLSRp uses the reception of the HELLO messages generated periodically for the network nodes as a validation method. Therefore, if an MPR implementing the OLSRp system does not receive a HELLO message from a given node, it will be aware that the node is inactive and that the topology has changed. Consequently, the OLSRp will deactivate the predictor and will send the real TC message.

The use of OLSRp means that every node keeps a table containing as many items as there are network nodes. Each entry in the table records the following information about the specific node:

- The node's IP address;
- A list of MPRs that announce the node in the TC message. This list includes the IP addresses of the MPRs (i.e. the originator addresses or O.A.) and the current state of the node, which is either active (A) or inactive (I). The state of a given node will be determined depending on whether or not the MPR has received HELLO messages from the node.
- A predictor state indicator for the MPR nodes (On or Off). This item will be activated when at least one of the TC messages that contains information about one MPR node is active, that is, when the MPR that generates the TC message in which the specific MPR is announced, has received HELLO messages from the specific

MPR. However, when the node is inactive in all the announcing TC messages, the predictor state indicator will be deactivated and the new TC message generated will be sent throughout the network.

Fig. 6 shows the execution of the OLSRp predictor in a network of six nodes where four of them were selected as MPRs. The figure shows the OLSRp table of node D. From the HELLO messages it has received this node detects that the MPRs C and E are active and so it starts the corresponding predictors. However, when the same nodes do not receive HELLOB (because node B is inactive) they generate a new TC message and send it throughout the network. In addition, when node D detects from the TC messages that node B is inactive, it deactivates the predictor of node B.



Node	O.A.$_{Node1}$ / State$_{Node1}$	O.A.$_{Node2}$ / State$_{Node2}$ …	Predictor state
192.168.1.2	192.168.1.3 /I	192.168.1.5 /I	TC$_B$ = Off
192.168.1.3	192.168.1.2 /A	192.168.1.4 /A	TCc = On
192.168.1.5	192.168.1.2 /A	192.168.1.4 /A	TC$_E$ = On
192.168.1.1	192.168.1.2 /A		-
192.168.1.6	192.168.1.5 /A		-

Fig. 6. OLSRp mechanism

Fig. 7 shows the interlayer communication of a node that is implementing the OLSRp system compared with that of a node that is only using the standard OLSR protocol. The OLSRp can be implemented as a transparent communication layer between OLSR and the lower communication layers. Notice that both approaches deal with exactly the same control traffic. The main difference is that the data sources for the OLSR layer are different. When the OLSR is used alone, all the information comes from the Wi-Fi, whereas when the OLSRp is used, the information can be provided by both the Wi-Fi and the OLSRp layer.

The OLSRp has several advantages. The most obvious one is the reduction of the control traffic that is transmitted and the consequent reduction in node energy consumption, network congestion, packet collisions and losses. This in turn increases the network's lifetime and has a positive impact on its performance and scalability.

On the other hand, implementing the OLSRp mechanism introduces some minimal additional costs. Each node executing the OLSRp has to maintain a table whose

dimensions depend on the number of network MPRs. In addition, the OLSRp consumes processing time of the node's CPU. However, OLSRp considerably reduces the overall cost involved in the transmission/reception and packing/unpacking processes. The cost in energy and processing time is higher than the additional cost introduced by the implementation of the OLSRp mechanism (it is widely known that a single packet transmission consumes the same energy as the execution of millions of instructions). Figures 8.a and 8.b show how the utilization and energy consumption of the CPU is affected by the number of TC messages transmitted in a 300 second test.

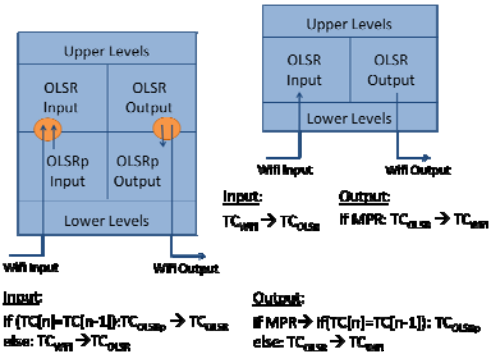


Fig. 7. OLSRp vs. OLSR layers

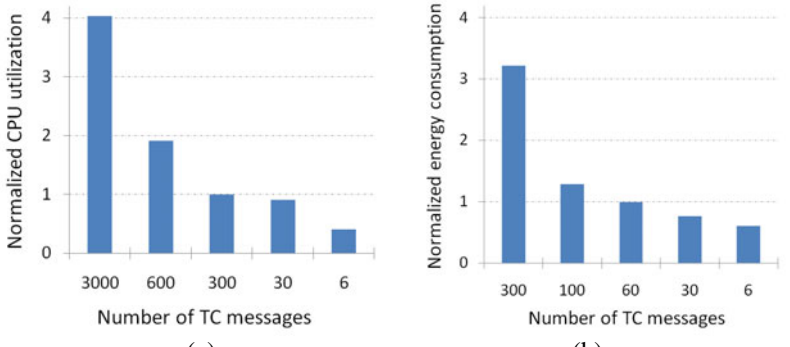


Fig. 8. (a) CPU utilization and (b) power consumption per node vs. TC emission interval

Finally, [13] states that energy consumption is correlated with mobility (the lower the speed, the higher the consumption). We also demonstrate (see Figure 5) that the repetition percentage is higher with lower speeds. Therefore, our mechanism fits better in high energy consumption environments.

6 Related Work

Prediction is a well-known and crucial technique in computer microarchitecture for achieving high performance and it has been applied successfully for years to several

parts of the processor. For instance, branch prediction [14] tries to reduce pipeline stalls by predicting the outcome of conditional branches, and value prediction [10] attempts to alleviate the serialization resulting from data dependences by predicting the results of arithmetic operations. Notice that prediction introduces additional complexity to the microprocessor because special hardware has to be devoted to predicting and then mandatorily validating the predictions. Moreover, there is an additional time penalty when there are mispredictions. However, on average if the percentage of predictions is high enough, the overall microprocessor performance is significantly improved at a reasonable hardware cost. The same concerns can be extended to ad hoc networks as these also benefit from prediction techniques.

Lifetime Prediction Routing (LPR) [11] is a routing protocol where each node tries to estimate its battery lifetime on the basis of its past activity. Hence, it is possible to increase the overall network lifetime by finding better routing solutions that take into account these predictions. The Kinetic Multipoint Relaying (KMPR) protocol [6] focuses on predicting mobility in order to improve routing. This approach selects relay nodes on the basis of the current relay configuration and the future network topology prediction. The Mobile Gambler's Ruin (MGR) algorithm [4] also applies mobility prediction. This predictive algorithm is developed under a cooperative scenario to identify nodes that are more likely to disconnect in the near future. Therefore, this prediction allows the coordination layer to reschedule the work among nodes in advance.

Finally, prediction is relatively easy to apply when there is a certain degree of redundancy in the network because it is normally based on the last value observed. Wireless sensor networks take advantage of this redundancy to reduce redundant communication, save energy and extend the battery lifetime [9], [5].

7 Conclusions and Future Work

OLSRp has been introduced as a scalable routing mechanism that focuses on eliminating redundant control information from the network and reducing computational processing and energy consumption. It is based on the observation that the probability of receiving a control message containing the same information as the previous one is very high. In fact, we have demonstrated that message repetition is only affected by mobility and remains almost constant when the number or density of nodes changes. Consequently, our proposal can be orthogonally applied to diverse scenarios where these parameters are different. Furthermore, we have also shown that, even with high speeds, the percentage of repetition is still significant.

Most previous studies have addressed the issue of routing protocol scalability in several ways but, to our knowledge, this is the first study that uses prediction to limit the increase in the number of control messages when the number of network nodes also increases.

In future research, we plan to implement the proposed mechanism in a simulation environment in order to experimentally demonstrate the potential of this technique. Furthermore, we want to investigate heterogeneous scenarios in which some nodes use the OLSRp predictor while other nodes use the standard OLSR protocol. This will also prove the adaptability of our mechanism as an additional transparent layer below

OLSR. We also want to extend this work to other proactive link-state protocols that control message flooding in a similar way. Moreover, we strongly believe that prediction can be also applied to proactive distance-vector protocols and even to reactive protocols in a similar way. Finally, the prediction accuracy and potential performance of the system could be improved with more sophisticated schemes than the one used in this study.

Acknowledgments. This work is supported by the Spanish Ministry of Science and Innovation (MCI) and FEDER funds of the EU under the contracts TIN2007-68050-C03-01, TIN 2007-68050-C03-03 and TIN 2007-61763.

References

1. Brewer, E.: Lessons from giant-scale services. IEEE Internet Computing 5(4), 46–55 (2001)
2. Clausen, T., Jacquet, P.: RFC3626 Optimized link state routing protocol, OLSR (2003)
3. Choi, J.-M., Ko, Y.-B.: A performance evaluation for ad hoc routing protocols in realistic military scenarios. Cellular and Intelligent Communications (2004)
4. De Rosa, F., Malizia, A., Mecella, M.: Disconnection prediction in mobile ad hoc networks for supporting cooperative work. IEEE Pervasive Computing 4(3), 62–70 (2005)
5. Gao, Y., Wu, K., Li, F.: Analysis on the redundancy of wireless sensor networks. In: Wireless Sensor Networks and Applications, pp. 108–114. ACM, New York (2003)
6. Härri, J., Filali, F., Bonnet, C.: Kinetic multipoint relaying: improvements using mobility predictions. In: Hutchison, D., Denazis, S., Lefevre, L., Minden, G.J. (eds.) IWAN 2005. LNCS, vol. 4388, pp. 224–229. Springer, Heidelberg (2009)
7. Hong, X., Xu, K., Gerla, M.: Scalable routing protocols for mobile ad hoc networks. Network 16(4), 11–21 (2002)
8. Iwata, A., Chiang, C.C., Pei, G., Gerla, M., Chen, T.-W.: Scalable routing strategies for ad hoc wireless networks. IEEE Journal on Selected Areas in Communications 17(8), 1369–1379 (1999)
9. Le, H.-C., Guyennet, H., Zerhouni, N.: Redundant communication avoidance for event-driven sensor network. J. Computer Science and Network Security 7(3), 193–200 (2007)
10. Lipasti, M.H., Wilkerson, C.B., Shen, J.P.: Value locality and load value prediction. In: Architectural Support for Programming Languages and Operating Systems, pp. 138–147. ACM, New York (1996)
11. Maleki, M., Dantu, K., Pedram, M.: Lifetime prediction routing in mobile ad hoc networks. In: Wireless Communications and Networking, pp. 1185–1190. IEEE, Los Alamitos (2003)
12. The Network Simulator, `http://www.nsnam.org/`
13. Pentikousis, K., Blume, O., Agüero, R., Papavassiliou, S., Puliafito, A.: Topology-aware hybrid random walk protolcs for wireless multihop networks. In: Mobile networks and management. LNICST, vol. 32, pp. 107–118. Springer, Heidelberg (2010)
14. Smith, J.E.: A study of branch prediction strategies. In: 25 years of the International Symposia on Computer Architecture, pp. 202–215. ACM, New York (1998)
15. Guifi network, `http://www.guifi.net/`

Maximum Sum-Rate Interference Alignment Schemes for the 3-User Deterministic MIMO Channel

Óscar González and Ignacio Santamaría

Dept. of Communications Engineering
University of Cantabria
39005 Santander, Cantabria, Spain
{oscargf,nacho}@gtas.dicom.unican.es

Abstract. Closed-form solutions exist for the interference alignment (IA) problem in the multiple-input multiple-output (MIMO) interference channel when there are exactly $K = 3$ users. Specifically, when each user wishes to send d streams and is equipped with $N = 2d$ antennas at both sides of the link, a finite number of IA solutions exist. Exploiting this observation, in this paper we find the maximum sum-rate solution by exhaustive search over the finite set of IA solutions and evaluate its performance. As an alternative, the solution that maximizes the received sum-power in the interference free subspace is also considered. Simulation results show the improvement achieved by both IA strategies in comparison with the conventional scheme proposed by Cadambe and Jafar, which randomly picks one of the IA solutions. Furthermore, the impact of channel correlation in these interference management techniques has also been studied.

Keywords: sum-rate, interference management, interference alignment, interference channel, multiple-input multiple-output (MIMO).

1 Introduction

Interference alignment (IA) is a recently proposed technique to achieve the maximum degrees of freedom (DoF) for K-user interference channels [1]. The degrees of freedom (also referred to as network multiplexing gain) approximates the capacity of a network as

$$C(SNR) = d\log(SNR) + o(\log(SNR))$$

where d is the number of degrees of freedom and $C(SNR)$ represents the capacity of the network as a function of the signal to noise ratio (SNR). At high SNR, the $o(\log(SNR))$ term becomes negligible in comparison to $\log(SNR)$. Therefore, d, represents the asymptotic slope of the $C(SNR)$ vs $\log(SNR)$ curve. By studying wireless networks at high SNR, the degrees of freedom approach de-emphasizes noise and explicitly addresses the effects of interference in a wireless network.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 237–244, 2011.

The DoF also indicate the number of simultaneous data streams for which the interference can be completely cancelled.

For the K-user interference channel, IA has been shown to achieve almost surely $K/2$ DoF per time, frequency or spatial dimension. Basically, IA schemes jointly design the signals transmitted by all users in such a way that the interfering signals at each receiver fall into a reduced-dimensional subspace. The receivers can then extract the projection of the desired signal that lies in the interference-free subspace.

In this paper we consider IA schemes based on linear precoding for the 3-user multiple-input multiple-output (MIMO) interference channel with constant coefficients, where every transmitter and receiver has an even number of antennas ($N = 2d$) and each user wishes to send d streams of data. For such systems, whose closed-form solutions are known, it follows that each user achieves $d = N/2$ DoF [1]. When N is odd, a two time-slot symbol extension is required. These systems[1] are usually denoted as $(2d \times 2d, d)^3$.

The first contribution of this paper is to show that for these MIMO interference channels there is a finite number of IA solutions. Although this result can easily be derived from the IA scheme described in [1] and [2], the existence of a finite number of IA solutions has not been widely acknowledged in the literature. The second contribution is to exploit this fact to find the IA solution that maximizes either the sum-rate or the sum-power.[2] For the 3-user case and a reasonable number of transmitted streams, the max sum-rate and max sum-power solutions can be obtained by exhaustive search over the finite set of IA solutions. The performance of the max sum-rate and the max sum-power IA solutions is studied in this paper by means of simulation results. In comparison to the conventional IA solution proposed by Cadambe and Jafar in [1], the relative advantage of the proposed schemes increases with the number of transmitted streams. The impact of channel correlation has also been considered.

This article is organized as follows: the system model, the principle of interference alignment and the closed-form solution for the 3-user channel are reviewed in Sections 2 and 3. The max sum-rate and max sum-power solutions are described in Section 4, while Section 5 presents simulations and performance comparisons.

2 Interference Alignment in the K-User Deterministic MIMO Channel

Consider the K-user interference channel, comprised of K transmitter - receiver pairs (links) that interfere with each other. We assume that each user wishes to achieve d degrees of freedom and is equipped with N antennas at both sides of

[1] The notation $(n_T \times n_R, d)^K$ means that every transmitter has n_T antennas, every receiver has n_R antennas and each one of the K users wishes to achieve d DoF. These interference channels are called symmetric.

[2] The term sum-power refers to the total power received in the interference-free subspaces by all the users.

the link. Also, let $\mathbf{V}^{[k]} \in \mathbb{C}^{N\times d}$ be an orthonormal basis of the transmitted signal space for user k. The discrete-time signal received at receiver k at a given time instant is the superposition of the signals transmitted by the K transmitters, weighted by their respective channel gains and affected by noise. It can be written as

$$\mathbf{y}^{[k]} = \mathbf{H}^{[kk]}\mathbf{V}^{[k]}\mathbf{s}^{[k]} + \sum_{l\neq k} \mathbf{H}^{[kl]}\mathbf{V}^{[l]}\mathbf{s}^{[l]} + \mathbf{w}^{[k]}, \tag{1}$$

where $\mathbf{H}^{[kl]} \in \mathbb{C}^{N\times N}$ is the flat-fading MIMO channel from transmitter l to receiver k, $\mathbf{s}^{[l]} \in \mathbb{C}^{d\times 1}$ is the signal transmitted by the l-th user and $\mathbf{w}^{[k]}$ is the additive and spatially white Gaussian noise at receiver k.

In this context, interference alignment is achieved when we are able to find a set of unitary precoding matrices $\mathbf{V}^{[k]}$ and unitary interference filtering matrices $\mathbf{U}^{[k]}$ such that, for $k = 1, \ldots, K$

$$\mathbf{U}^{[k]H}\mathbf{H}^{[kl]}\mathbf{V}^{[l]} = 0, \quad \forall l \neq k, \tag{2}$$

and

$$\operatorname{rank}(\mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}) = d. \tag{3}$$

When an IA solution exists, the signal received at user k after projecting $\mathbf{y}^{[k]}$ onto the orthogonal subspace of the interference, using $\mathbf{U}^{[k]}$, yields

$$\begin{aligned}\mathbf{r}^{[k]} &= \mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}\mathbf{s}^{[k]} + \sum_{l\neq k} \mathbf{U}^{[k]H}\mathbf{H}^{[kl]}\mathbf{V}^{[l]}\mathbf{s}^{[l]} + \mathbf{U}^{[k]H}\mathbf{w}^{[k]} \\ &= \mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}\mathbf{s}^{[k]} + \mathbf{n}^{[k]}.\end{aligned} \tag{4}$$

According to (3) and (4), the effective channel $\mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}$ is now $d \times d$ dimensional. In summary, by applying IA the MIMO interference channel has been transformed into a set of Gaussian parallel $d \times d$ MIMO channels.

3 Closed-Form IA Solutions for the 3-User Case

In this paper, we focus on the case where the number of users is $K = 3$, the number of antennas is $N = 2d$ at the transmitter and receiver sides, and each user wishes to send d streams. All the interference alignment solutions (as proposed in [1]) can be obtained as follows:

1. The precoder for the user 1, $\mathbf{V}^{[1]}$, is formed by taking any subset of d eigenvectors of the following $2d \times 2d$ matrix, not necessarily the main eigenvectors

$$\mathbf{E} = (\mathbf{H}^{[31]})^{-1}\mathbf{H}^{[32]}(\mathbf{H}^{[12]})^{-1}\mathbf{H}^{[13]}(\mathbf{H}^{[23]})^{-1}\mathbf{H}^{[21]}. \tag{5}$$

2. The precoders for users 2 and 3, $\mathbf{V}^{[2]}$ and $\mathbf{V}^{[3]}$, are obtained respectively as

$$\mathbf{V}^{[2]} = (\mathbf{H}^{[32]})^{-1}\mathbf{H}^{[31]}\mathbf{V}^{[1]}, \tag{6}$$

and

$$\mathbf{V}^{[3]} = (\mathbf{H}^{[23]})^{-1}\mathbf{H}^{[21]}\mathbf{V}^{[1]}. \tag{7}$$

Since $\mathbf{E}$ is a full-rank $2d \times 2d$ matrix, there are $C_{2d}^{d} = \binom{2d}{d}$ ways to choose the d-dimensional unitary precoder for the first user. Each one of these precoders for user 1 yields a distinct IA solution.

An interesting fact of the 3-user interference channel is that it induces a permutation structure which makes that starting the process described above with a different user yields exactly the same set of the IA solutions. In other words, the IA solutions do not depend on which user is picked first.

To clarify this point let us, for example, start the procedure with user 2 instead of user 1. Now, $\mathbf{V}^{[2]}$ is formed by taking any subset of d eigenvectors of the matrix

$$\mathbf{E}' = (\mathbf{H}^{[12]})^{-1}\mathbf{H}^{[13]}(\mathbf{H}^{[23]})^{-1}\mathbf{H}^{[21]}(\mathbf{H}^{[31]})^{-1}\mathbf{H}^{[32]} \tag{8}$$

and the precoder for the user 1, can be calculated as

$$\mathbf{V}^{[1]} = \mathbf{H}^{[31]}(\mathbf{H}^{[32]})^{-1}\mathbf{V}^{[2]}. \tag{9}$$

The following relationship holds between the eigenvectors of $\mathbf{E}$ and $\mathbf{E}'$ (this is due to the permutation structure induced by the 3-user channel)

$$\nu(\mathbf{E}) = \mathbf{H}^{[31]}(\mathbf{H}^{[32]})^{-1}\nu(\mathbf{E}'), \tag{10}$$

and therefore it is clear that the same solutions are obtained starting with user 1 or 2. Obviously, this also occurs when we start the alignment procedure from user 3. In conclusion, the number of IA solutions for the $(2d \times 2d, d)^3$ systems is exactly $C_{2d}^{d} = \binom{2d}{d}$.

The main implication of this result is that for reasonable values of d (e.g., from 1 to 5), it is feasible to find the best IA solution (in terms of sum-rate, for instance) by exhaustive search over the finite set of $\binom{2d}{d}$ solutions. The additional computation cost is moderate and, as we will show in Section 5, a significant improvement can be achieved.

4 Max Sum-Rate and Max Sum-Power IA Solutions

The most straightforward figure of merit for measuring the system performance is the sum-rate, which is given by

$$R = \sum_{k=1}^{K} \log \left| \mathbf{I}_N + \left(\sigma^2 \mathbf{I}_N + \sum_{l \neq k} \mathbf{Q}^{[kl]} \right)^{-1} \mathbf{Q}^{kk} \right|, \tag{11}$$

where $\mathbf{Q}^{[kl]}$ denotes the $N \times N$ covariance matrix of the signal from the l-th transmitter to the k-th receiver, and σ^2 is the variance of the additive white Gaussian noise.

When the interference is perfectly aligned (i.e. (2) and (3) are satisfied), the interference channel is decoupled into a set of parallel Gaussian MIMO channels and the sum-rate in (11) reduces to

$$R = \sum_{k=1}^{K} \log \left| \mathbf{I}_N + \frac{1}{\sigma^2} \mathbf{U}^{[k]H} \mathbf{H}^{[kk]} \mathbf{V}^{[k]} \mathbf{V}^{[k]H} \mathbf{H}^{[kk]H} \mathbf{U}^{[k]} \right|, \tag{12}$$

which simply adds up the achievable rates in each interference-free $d \times d$ MIMO channel given by $\overline{\mathbf{H}}^{[k]} = \mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}$, for $k = 1, \ldots, K$. Moreover, it is clear that after the channel has been block-diagonalized better throughputs can be obtained by using non-unitary precoders and decoders, or by applying power waterfilling among the eigenmodes of the equivalent non-interfering $d \times d$ MIMO channels. Maximizing the sum-rate in (11) subject to the unitary constraints $\mathbf{V}^{[k]H}\mathbf{V}^{[k]} = \mathbf{I}_d$, $\mathbf{U}^{[k]H}\mathbf{U}^{[k]} = \mathbf{I}_d$ and also to the zero-interference constraint (2) is a challenging problem that has not been solved yet.

As an alternative to the sum-rate criterion, in this paper we will also consider the maximization of the power received in the interference-free subspaces, which is given by

$$P = \sum_{k=1}^{K} \mathrm{tr}\left(\mathbf{U}^{[k]H}\mathbf{H}^{[kk]}\mathbf{V}^{[k]}\mathbf{V}^{[k]H}\mathbf{H}^{[kk]H}\mathbf{U}^{[k]}\right), \tag{13}$$

where $\mathrm{tr}(\mathbf{A})$ denotes the trace of matrix $\mathbf{A}$.

As it has been shown in Section 3, for $K = 3$ users, a finite number of closed-form interference alignment solutions exist. Therefore, in principle it is possible to choose the solution that maximizes (12) or (13) by exhaustive search over all IA solutions.

In networks with $K > 3$ users, closed-form IA solutions are yet to be found. For this reason, one has to resort to iterative algorithms which alternatively optimize the precoders at the transmitters and the decoders at the receivers [3], [4]. These alternating minimization algorithms obtain an IA solution that do not optimize neither the sum-rate nor the sum-power. Although they can be modified to maximize certain cost function (i.e. the sum-rate in (12)), this requires the computation of quite complex derivatives [5]. The advantage of maximizing the sum-power is that it requires less complex derivatives. In Section 5 the usefulness of using the sum-power as a proxy for the sum-rate is assessed.

5 Simulation Results

In this section, numerical simulation results are presented to evaluate the performance of the maximum sum-rate (MaxSR-IA) and the maximum sum-power (MaxPower-IA) solutions. These solutions are also compared with the conventional IA solution (proposed in [1]) which randomly selects a subset of d eigenvectors of the matrix $\mathbf{E}$ in (5). All the results presented in this section consider a $(2d \times 2d, d)^3$ systems with $d = 1, \ldots, 4$.

5.1 Rayleigh Interference MIMO Channel

For this example, the matrices $\mathbf{H}^{[kl]}$ represent independent and identically distributed (i.i.d.) Rayleigh fading MIMO channels with unit-variance entries. The average sum-rate has been evaluated for 10000 channel realizations for different values of d and signal-to-noise ratio (SNR) values ranging from 0 to 40 dB. The obtained results are depicted in Fig. 1.

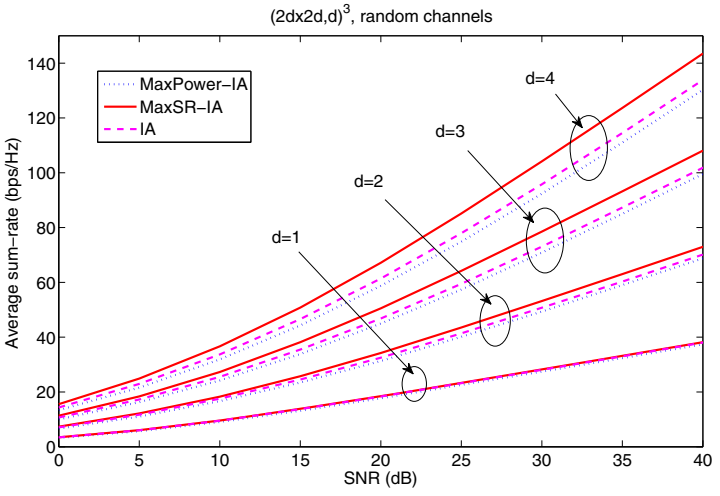


Fig. 1. Average sum-rate achieved by the MaxSR-IA, MaxPower-IA and the conventional IA solutions

As it can be seen in Fig. 1, the MaxSR-IA solution provides a considerable improvement with respect to the other two IA solutions for all values of d. MaxPower-IA seems to be a good approximation of the MaxSR-IA for low values of d. However, as d becomes larger its performance gets considerably lower than the MaxSR-IA solution.

In turn, the average sum-rate improvement with respect to the conventional IA solution normalized by the number of streams is depicted in Fig. 2. It shows that the relative sum-rate difference between the MaxSR-IA and the conventional IA solution is increased when the value of d increases.

5.2 Correlated MIMO Channel

In this subsection we evaluate the impact of MIMO channel correlation using the well-known Kronecker model

$$\mathbf{H}_c^{[kl]} = \mathbf{R}_{rx}^{[k]1/2} \mathbf{H}^{[kl]} \mathbf{R}_{tx}^{[l]1/2} \tag{14}$$

where $\mathbf{H}_c^{[kl]}$ is the correlated MIMO channel matrix for transmitter l and receiver k, $\mathbf{R}_{rx}^{[k]}$ and $\mathbf{R}_{tx}^{[l]}$ are the receiver k and transmitter l antenna correlations. All transmitter and receiver correlation matrices are assumed equal, that is $\mathbf{R}_{tx}^{[l]} = \mathbf{R}_{rx}^{[k]} = \mathbf{R} \quad \forall\, l, k \in \{1, 2, 3\}$. The ij entry of $\mathbf{R}$ is given by the Jakes model [6] which assumes that the antennas are uniformly arranged along a line (each pair is separated a distance L),

$$\mathbf{R}_{ij} = J_0 \left(2\pi \frac{L|i-j|}{\lambda} \right), \tag{15}$$

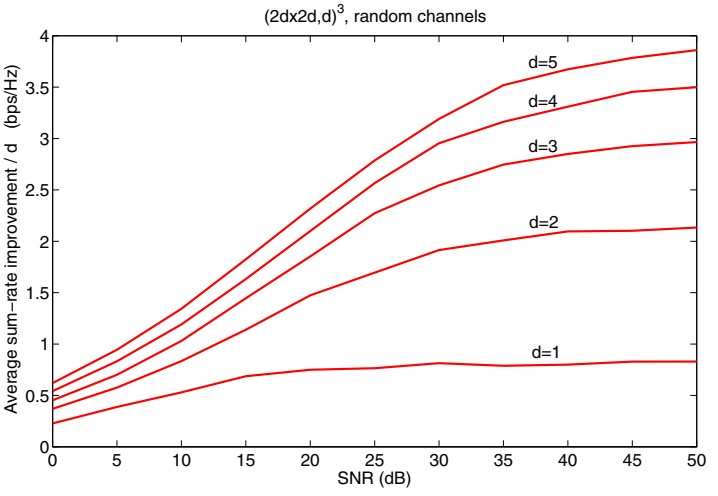


Fig. 2. Sum-rate improvement per stream achieved by the MaxSR-IA solution over the conventional IA solution from $d = 1$ to $d = 5$ streams per user

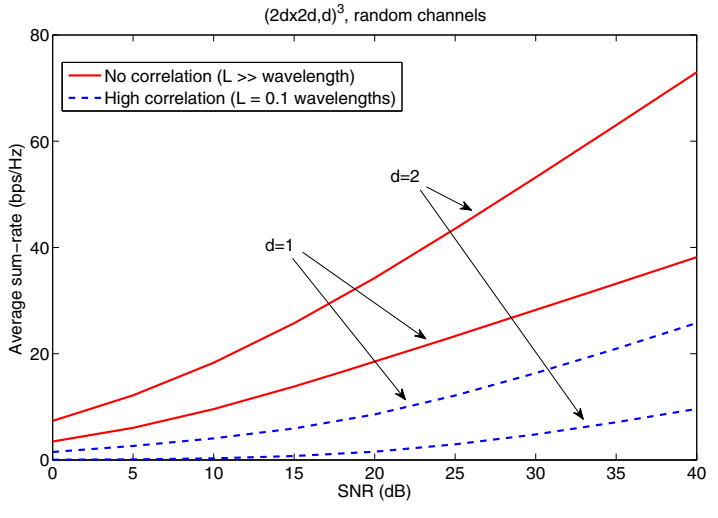


Fig. 3. Impact of correlation in the average sum-rate for 1 and 2 streams per user

where $J_0(\cdot)$ is the zeroth order Bessel function. The correlation effect has been studied by assuming an antenna separation $L = 0.1\lambda$. Obviously, such a small separation gives a highly correlated channel. Fig. 3 shows the MaxSR-IA solution performance for $d = 1$ and $d = 2$.

As it can be seen in Fig. 3, an increasing correlation is detrimental for the sum-rate. However, it is interesting to notice that the average sum-rate for $d = 1$

is better than for $d = 2$. This effect is due to the fact that highly correlated scenarios result in almost rank-deficient MIMO channels, which do not allow to transmit more than one stream per user. In other words, with strong correlation it is more difficult to exploit the different eigenmodes of the MIMO channels.

6 Conclusion

In this paper the number of interference alignement solutions for the 3-user symmetric channel has been analyzed. Exploiting the fact that a finite number of solutions exist, we have derived the maximum sum-rate and the maximum sum-power IA solutions for this scenario. The obtained results have shown that the improvement of the sum-rate solution increases when the number of streams per user increases. As expected, both proposed solutions overcome the conventional IA solution (in terms of sum-rate). This justifies the need to find algorithms that optimize the sum-rate while cancelling the interference leakage.

Acknowledgment

This work has been supported by the Spanish Government (MICINN) under project TEC2007-68020-C04-02/TCM (MultiMIMO).

References

1. Cadambe, V.R., Jafar, S.A.: Interference alignment and degrees of freedom of the K-user interference channel. IEEE Trans. Inf. Theory 54, 3425–3441 (2008)
2. Yetis, C.M., Gou, T., Jafar, S.A., Kayran, A.H.: On the feasibility conditions for interference alignment. In: Proc. IEEE Global Telecommunications Conference (GLOBECOM), Honolulu, HI, USA (2009)
3. Gomadam, K., Cadambe, V.R., Jafar, S.A.: Approaching the capacity of wireless networks through distributed interference alignment. In: Proc. IEEE Global Telecommunications Conference (GLOBECOM), New Orleans, LA, USA (2008)
4. Peters, S.W., Heath, J.R.W.: Interference alignment via alternating minimization. In: Int. Conf. Acoust. Speech and Signal Processing (ICASSP), Taipei, Taiwan (2009)
5. Santamaría, I., González, O., Heath, R.W., Peters, S.W.: Maximum sum-rate interference alignment algorithms. In: Proc. IEEE Global Telecommunications Conference, Miami, FL, USA (2010)
6. Jakes, W.C.: Microwave Mobile Communications. IEEE Press, New York (1993)

A Novel LTE Wireless Virtualization Framework

Yasir Zaki[1], Liang Zhao[1], Carmelita Goerg[1], and Andreas Timm-Giel[2]

[1] Communication Networks, TZI, University of Bremen,
Otto-Hahn-Allee, NW1, 28359 Bremen, Germany
{yzaki,zhao,cg}@tzi.de
[2] Institute of Communication Networks/TUHH, Hamburg, Germany,
Schwarzenbergstr. 95E, 21073 Hamburg, Germany
timm-giel@tuhh.de

Abstract. Network virtualization is one of the topics that recently have been receiving attention in the research community. It is becoming evidently clear that network virtualization will be a major player in the shaping of the Future Internet. Many research projects around the world are studying different aspects of network virtualization: some are focusing on resource virtualization like Node, Server and Router virtualization; while others are focusing on building a framework to setup virtual networks on the fly based on the different virtual resources. In spite of all that work, we still think that one very important piece of the puzzle is still missing that is "Wireless Virtualization". According to the best of our knowledge, the virtualization of the wireless medium has not yet received the appropriate attention it is entitled to, and there has been very small work done in that field. This is why this paper is proposing a framework for the virtualization of the wireless medium. This framework is proposed to virtualize mobile communication systems so that multiple operators can share the same physical resources. We mainly focus on the Long Term Evolution (LTE) but the framework can also be generalized to fit any other wireless system.

Keywords: network virtualization, wireless virtualization, LTE simulations, Future Internet.

1 Introduction

Virtualization is a well known technique that has been used for years, especially in computer science like the use of virtual memory and virtual operating systems. What is new though is the idea of using virtualization to create complete virtual networks. One option in the Future Internet is the possibility of having instead of only one multiple co-existing architectures, in which each of the architectures is designed and customized to fit one specific type of network and satisfies its requirements. That is why Network Virtualization will play a vital role since it helps diversifying the Future Internet into separate virtual networks that are isolated from each other and can run different architectures within.

This paper is organized as follows: section 2 gives a short introduction to network virtualization as well as wireless virtualization and the motivation behind it. Section 3 introduces the Long Term Evolution, and our proposal for mobile system

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 245–257, 2011.

virtualization is discussed. A network simulator that is developed in OPNET for the mobile system virtualization is described in section 4. At the end, section 5 and 6 shows the simulation scenarios and results, as well as the conclusions and the outlook.

2 Network Virtualization

Due to the congenital flaws, e.g. inadequate supporting of security, mobility and multi-homing, the current Internet architecture is being criticized and it seems that many people believe that the current Internet will soon break and will not be able to cope with all the requirements coming in the future, although many patching-up works have been done for years. In order to satisfy the boom of new services running on the networks, many research activities on the Future Internet architecture have been launched throughout the world, for example 4WARD [4] in Europe, VINI [2] and GENI [1] in the U.S. and AKARI [5] and AsiaFI [6] in Asia. VINI is a virtual network infrastructure based on PlanetLab on which researchers can deploy, run and test their own protocols and services in a large scale. GENI (Global Environment for Network Innovations) is a novel suite of architectures which support a range of experimental protocols, and virtualization is one of the most important features of it. The AKARI Architecture Design Project aims to implement the basic technology of a new generation network from a clean slate, and network virtualization has been seen as one of the principles. Asia Future Internet Forum (AsiaFI) was founded to coordinate research and development on Future Internet where network virtualization is also one of the research topics. By observing these projects, one tendency can be perceived that network virtualization is an attractive technique which receives more and more research attention, and it will be a key area in the future network development.

In principle, network virtualization enables multiple virtual networks to coexist on a common infrastructure. Each virtual network is running similar to a normal network and does not necessarily have the awareness of the underlying virtualization process. Individual virtual networks can contain operator-specific protocols and architectures which could be totally different from other co-existing virtual networks.

2.1 Motivation of Wireless Network Virtualization

Mobile system virtualization can obtain several advantages and also have impacts on different aspects:

- For the infrastructure providers: deploying and operating a mobile system such as GSM or UMTS needs enormous investments. The maintenance and upgrade of the hardware need high operational expenditure. The big companies have to play both roles: system operators as well as infrastructure providers. With virtualization, they can only concentrate on the maintenance of the physical equipments and save the manpower for running the networks.
- For the virtual mobile system operators: infrastructure sharing is very attractive, where the huge investment on the hardware and fundamental construction could be saved. It also enables the small companies to get into the

market without huge investments. The deployment, maintenance, migration and upgrade of the virtual mobile systems will be flexible and with short time frame, even on-the-fly.

- For the end users: the increased number of operators will bring a diversity of services to end users. Which mean more options to satisfy the user's personal demand and subsequently enjoy the services with reasonable pricing.

2.2 State-of-the-Art of Mobile System Sharing/Virtualization

One of the interesting commercial products of soft radio is VANU [3], that is a wireless infrastructure solution that enables individual base stations to simultaneously operate GSM, CDMA, iDEN and beyond. It is also announced that the virtualized base station and RAN are also supported by their products. Spectrum sharing is one crucial part of our proposal on the mobile system virtualization, through which multiple operators can be introduced into one band. Other than the traditional DSA (dynamic spectrum allocation) model, [14] proposed a centralized spectrum broker that has the responsibility to coordinate the frequency allocation among different wireless networks. Two optimization problems, maximized requirements and minimized interference, are formulated and algorithms are designed to solve the efficiency problem.

One thing has to be mentioned here that MVNO (Mobile Virtual Network Operator) has been available for long time in the current mobile systems. However, from the definition of MVNO we know that they don't own any physical resources and have no impact on the network configuration or algorithms. Actually, MVNOs act more like a service provider. Our proposal introduced in the next section is trying to decouple the physical system from the network, and the Virtual Operator has the overall control of its own virtual network.

The mobile system infrastructure sharing also exists along with the deployment of the network, especially at the beginning. The reason is to reduce the cost and to roll out the infrastructure quickly. As referred to in [15], the infrastructure sharing has mainly three areas: *Passive component sharing, Active element sharing* and *Geographical sharing*. To reduce the operating cost, different operators will share the fundamental establishments like roof locations, tower frame, equipment houses and power supply. This kind of sharing doesn't involve network virtualization at all and is already accepted by most of the operators because of the economical benefits. Geographical sharing is simply dividing the whole area into several regions and each operator will be in charge of one of them. In this way the full coverage can be achieved in short time by the federation of operators on different regions.

3 LTE Virtualization Framework

Long Term Evolution (LTE) has been seen as one of the major solutions of next generation mobile systems. According the specifications from 3GPP, LTE will provide much higher data rates than UMTS, low latency and enhanced QoS especially for the users at the cell edge. Due those promising system performances, we choose

the LTE system as a case study to show the advantages of (wireless) network virtualization. The idea is to virtualize the LTE base station or what is also known as the enhanced Node-B (eNodeB). This is the physical hardware that is responsible to send and receive the date to and from the LTE users. Virtualizing the eNodeB is similar to any other node virtualization, where an entity called "Hypervisor" is added on top of physical resources, and is responsible for scheduling these resources between the different virtual instances running on top of it. Figure 1 shows the LTE eNodeB virtualization architecture.

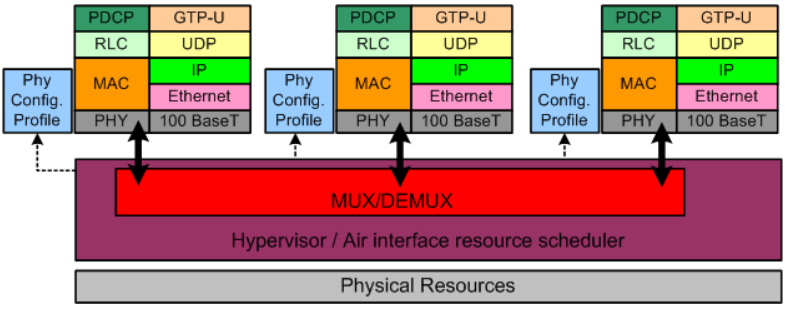


Fig. 1. Virtualized LTE eNodeB protocol stack

The hypervisor is also responsible for scheduling the air interface resources or the LTE spectrum between the different virtual eNodeBs (different virtual operators). The virtual operators share the spectrum based on different criteria. The Hypervisor collects information from the individual virtual eNodeB stacks, like user channel conditions, loads, priorities, QoS requirements and information related to the contract of each of the virtual operators. This information is used to schedule the air interface resources between the different virtual operators. LTE uses OFDMA in the downlink, which means that the frequency band is divided into a number of sub-bands each with a carrier frequency. The air interface resources that the hypervisor schedules are actually the Physical Radio Resource Blocks (PRB); this is the smallest unit that the LTE MAC scheduler can allocate to a user.

Scheduling the PRBs between the different virtual eNodeBs actually means splitting the frequency spectrum between the different eNodeBs of the different operators. The hypervisor can make use of apriori knowledge (e.g. users channel conditions, virtual operator contract, load ... etc.) to schedule the PRBs. OFDMA scheduling has been studied extensively in the literature [10] [11] [12] [13], but what is new here is that the frequency spectrum among the different operators has to be scheduled. This is even more challenging because of the additional degree of freedom that has been added. A number of possibilities and degrees of freedom exist here, where the scheduling could be based upon different criteria's: bandwidth, data rates, power, interference, pre-defined contracts, channel conditions, traffic load or a combination of them. At the end the hypervisor has to convert these criteria into a number of PRBs to be scheduled to each operator, but the challenge is to make sure that the allocated PRBs would be fair and enable the operators to satisfy their

requirements. This means that some mechanisms/contracts guidelines has to be defined to guarantee the resources to the operators which could be done by different options for example setting a guaranteed amount for each operator and leaving the rest of the resources to be shared. What is also important here is the time frame that the hypervisor operates with in order to guarantees the pre-defined requirements.

In our paper, we only concentrate on two different types of the scheduler algorithms, a static and a dynamic one. In the static algorithm the spectrum is divided between the virtual operators beforehand, and each operator gets his operating spectrum and keeps it for the whole time. This is similar to today's network, where each operator has his own frequency spectrum and no other operator is allowed to use it. In the dynamic algorithm, the resources are allocated to the different operators during runtime, and the amount and allocation can be changed over time depending on the operators traffic load. The latter algorithm is an example of what can be gained by applying network virtualization into wireless mobile communication systems, where not only the operator will share the physical infrastructure but will also share the frequency spectrum and this in turn will lead to better resource utilization.

4 Simulation Model

The LTE simulation model used in this paper was developed using the OPNET simulation tool [7]. The model is designed and implemented following the 3GPP specifications.

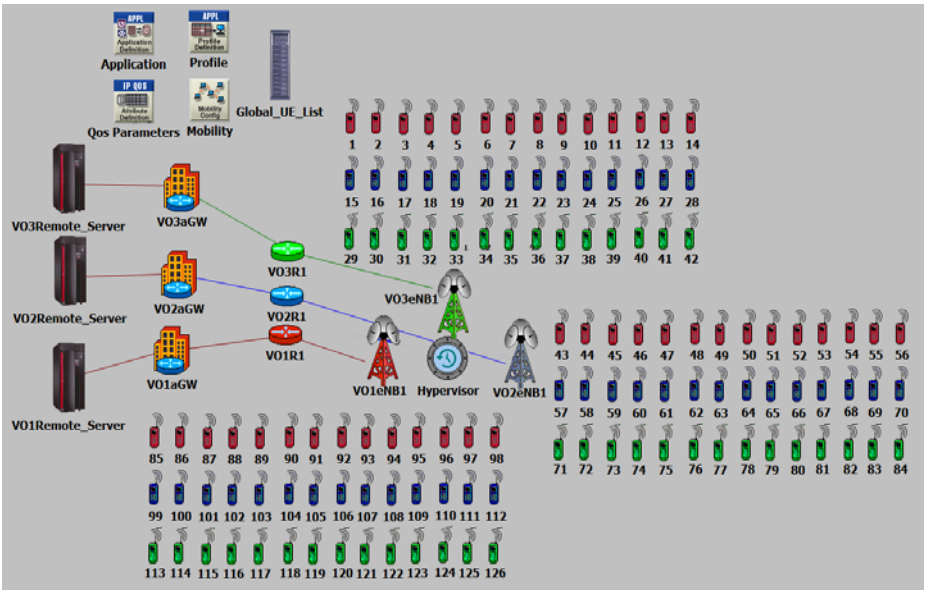


Fig. 2. Three Virtual Operators example of a virtualized LTE network

The focus of this work was not on the node or link virtualization, but rather on the air interface virtualization and how to schedule the air interface resources among the different virtual operators, no node/link virtualization was simulated, instead the assumption was that we have a perfect node/link virtualization. Figure 2 shows an example scenario, it can be seen that an additional entity between the virtual eNodeBs that is the "Hypervisor". This entity is responsible for scheduling the air interface resources (frequency spectrum or PRBs) among the different virtual eNodeBs. The Hypervisor also has direct access to the MAC layers of each of the LTE virtual eNodeBs to collect the required relevant information to be used to base the scheduling on, like each operator's users channel conditions, and operator traffic load.

In our implementation, two versions of the hypervisor exist based on the previously discussed hypervisor types. One is the static version, where the hypervisor allocates the PRBs among the different virtual operators just once at the beginning of the simulations, the number of the allocated PRBs for each operator will be equal, where each virtual eNodeB will get the exact same amount of PRBs and keeps it regardless if it is being actually used or not. The second version of the hypervisor is a dynamic one, where the PRBs are allocated to the different virtual operators in a dynamic manner at equal time intervals. The amount of the allocated PRBs will depend on the load that each operator is experiencing during the last time instance. In this way, each operator will only get his required share of the PRBs and no waste of resources will occur.

5 Simulation Configurations and Results

In this paper, we mainly aim to show the effects and advantages of using network virtualization in mobile communication systems. Specifically we investigate on additional benefits that could be gained if the mobile operators actually share the air interface resources (i.e. frequency band), which would be possible through virtualizing the mobile system infrastructure. From that, the simulations scenarios investigated in this paper are divided into two different setups:

- Static hypervisor configuration: which could be viewed similar to today's mobile network setup apart from sharing the infrastructure which is an additional benefit
- Dynamic hypervisor (load based) configuration: which aims to show how the mobile network operators can share the spectrum and what are the benefits of the LTE virtualization

5.1 Simulation Configurations

As discussed earlier, two different scenarios are investigated within this paper; besides of the hypervisor scheduler algorithm the rest of the configuration is exactly the same for both scenarios and is shown in subsequent Table 1.

Table 1. Simulation Parameters

Parameter	Assumption
Number of virtual operators	3 virtual operators
Number of virtual eNodeBs	3 eNodeBs (one per virtual operator)
eNodeB coverage area	Circular with Radius = 375 meters, number of cells = 3 with 120°
Total Number of PRBs	99 (which corresponds to about ~ 20 MHz), Reuse factor = 3
Mobility model	Random Way Point, users are initially distributed uniformly
Users speed	5 km/h
Number of active users	<table><tr><td></td><td>Cell1</td><td>Cell2</td><td>Cell3</td></tr><tr><td rowspan="2">VO 1</td><td>10 VOIP</td><td>10 VOIP</td><td>10 VOIP</td></tr><tr><td>4 Video</td><td>2 Video</td><td>1 Video</td></tr><tr><td rowspan="2">VO 2</td><td>10 VOIP</td><td>10 VOIP</td><td>10 VOIP</td></tr><tr><td>4 Video</td><td>2 Video</td><td>1 Video</td></tr><tr><td rowspan="2">VO 3</td><td>10 VOIP</td><td>10 VOIP</td><td>10 VOIP</td></tr><tr><td>4 Video</td><td>2 Video</td><td>1 Video</td></tr></table>
Path loss model	128.1 + 37.6 log10(R) dB, R in km [8]
Slow Fading model	Lognormal distributed with Mean value = zero Standard deviation = 8 dB and Correlation distance = 50 meters
Fast Fading model	Jake's model
CQI reporting	Ideal
Downlink Low traffic model	Voice Over Internet Protocol (VOIP) Silence/Talk Spurt length = neg. exponential with 3 sec mean Call duration = 10 sec Inter-repetition time = uniformly distributed between 5 and 10 sec throughout the whole simulation
Downlink Peak traffic model	Video conferencing application Incoming/Outgoing stream inter arrival time = Const (0.01 sec) Incoming/Outgoing stream frame size = Const (80 Bytes) Duration = Const (300 sec) (see Table 2)
Hypervisor resolution	1 sec (this is only for the dynamic scenario)
Simulation run time	1000 sec

In order to show the effect and benefits of the wireless virtualization and the air interface resource sharing, a peak traffic model has been introduced to the simulation scenarios, where, as it can be seen from the table above, this peak traffic model is used for only 300 seconds to emulate a sudden peak in the load of the operator. The place where this peak traffic model is used for each operator's users within the simulation time is configured as follows:

Table 2. Peak traffic setup for each virtual operator

	Cell 1	Cell 2	Cell 3
Virtual operator 1	0 – 300 sec	600 – 900 sec	300 – 600 sec
Virtual operator 2	300 – 600 sec	0 – 300 sec	600 – 900 sec
Virtual operator 3	600 – 900 sec	300 – 600 sec	0 – 300 sec

5.2 Simulation Results

As discussed earlier the main focus of this paper is the LTE air interface virtualization, and in order to show the performance gains from virtualizing the air interface and thus sharing the resources i.e. frequency band between the different virtual network operators two different scenarios are compared against each other.

The scenarios are configured as stated in the earlier section, one scenario which is similar to today mobile network operator setup apart from sharing the infrastructure, where three different operators are operating in the same region, each operator uses his own frequency band; these bands are being pre-allocated in the beginning of the simulations. Since we have 99 PRBs in total, each operator will get one third of the PRBs that is 33 PRBs. the second scenario can be looked at as the futuristic approach where the three operators are actually sharing the frequency band dynamically depending on their traffic load and requirements, each second the hypervisor tries to calculate from the previous time instance what the traffic load of each operator is and how the channel conditions experienced by the operator's users look like, and then assigns the PRBs among the different operators based on these calculations. The first scenario is called "Static" and the second one "Dynamic".

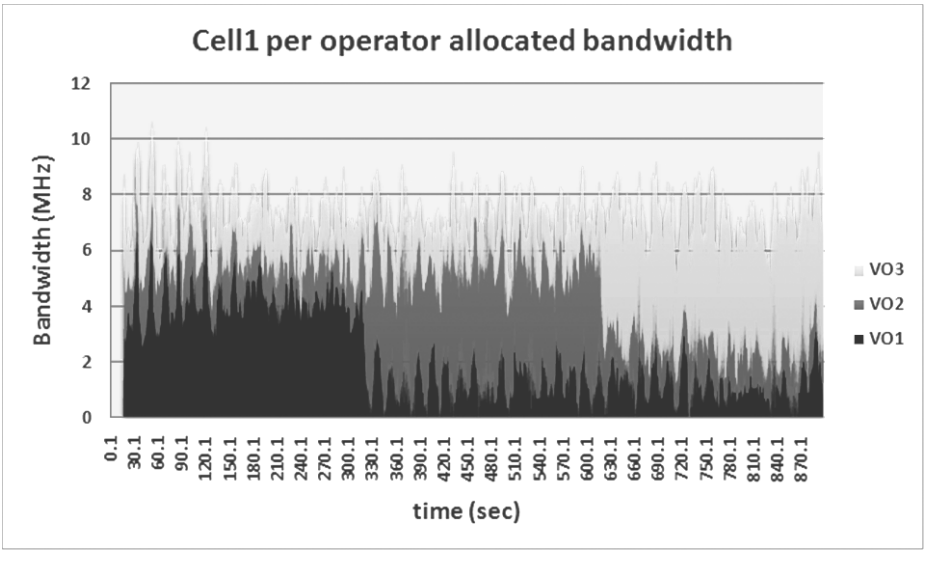


Fig. 3. Cell1 per virtual operator (VO) allocated bandwidth (MHz)

Figure 3 shows the bandwidth of each of the virtual operators that the hypervisor has allocated. It can be noticed that the allocated bandwidth changes with time depending on the traffic load and the users channel condition of each operator. In the first 300 seconds of the simulation run time it can be seen that operator 1 has been allocated a much higher bandwidth compared to the other two operators. This is due to the scenario configuration, where as it was previously configured in Table 2, operator 1 will have four additional users with video applications causing an increase in that operator's traffic. Similarly there are additional users in the second 300 seconds in the virtual network 2 (VO2) and the last 300 seconds in the virtual network 3 (VO3).

The average per user air interface throughput for operator 1 users in cell number 1 can be seen in Figure 4. The left side figure shows the air interface throughput for the

VOIP users whereas the right side figure shows the video users throughputs. What can be noticed from these figures is that the dynamic scenario achieves an expected better throughput than the static case. The reason for that is mainly the fact that additional resources have been used for the dynamic case, where in the static case these resources were allocated for the other operators but were not used.

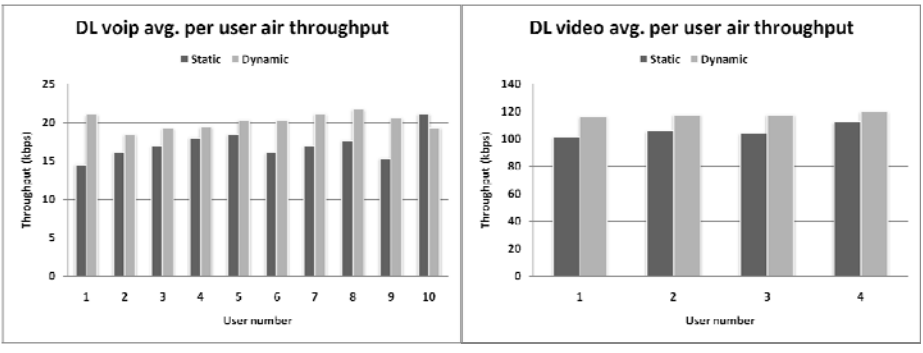


Fig. 4. Cell1 virtual operator 1 per user air interface throughput (kbps)

The average user application end-to-end delays can be seen in Figure 5. It can be noticed that the static scenario suffers from higher application delays as compared to the dynamic case especially for the video users. The reason is the limited air interface resources.

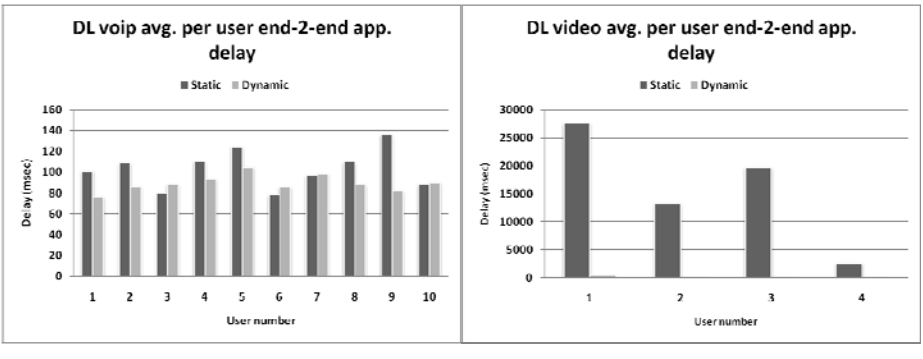


Fig. 5. Cell1 virtual operator 1 per user end to end application delay (m sec)

In order to check how the VOIP application performs, the average delay values shown in the previous figures are not sufficient. So in addition Figure 6 shows the probability when the end-to-end delays of the VOIP packets were greater than 300 msec (which is the QoS limit for the VOIP packets). It can be seen that the static case has a higher probability for not satisfying the QoS threshold.

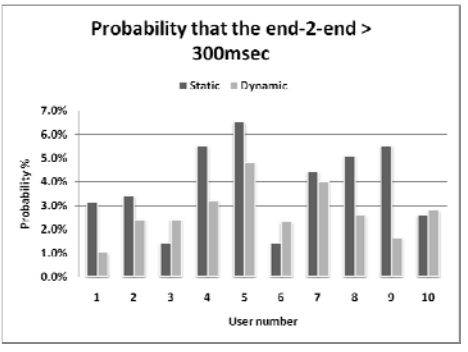


Fig. 6. Cell1 virtual operator 1 end to end delay probability > 300 m sec

Figure 7 to Figure 12 show results of cell 1 obtained for the other two operators. Similar observations can be made from these results, where mainly the dynamic case performs better than the static one especially when it comes to the video traffic. The video users are sending a continuous stream of data for a larger time span as compared to the VOIP bursty traffic.

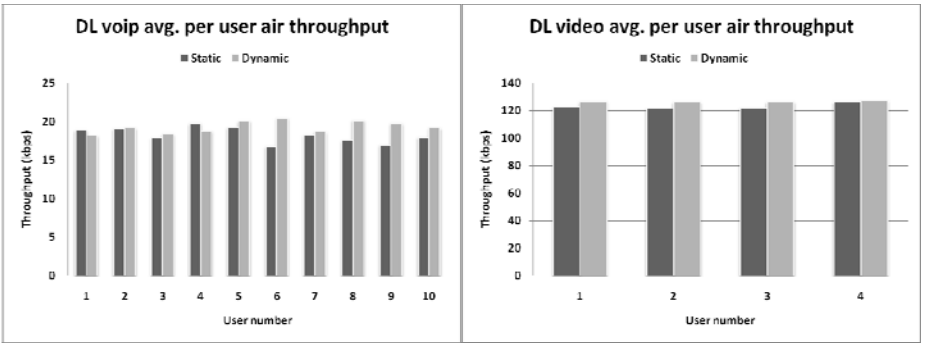


Fig. 7. Cell1 virtual operator 2 per user air interface throughput (kbps)

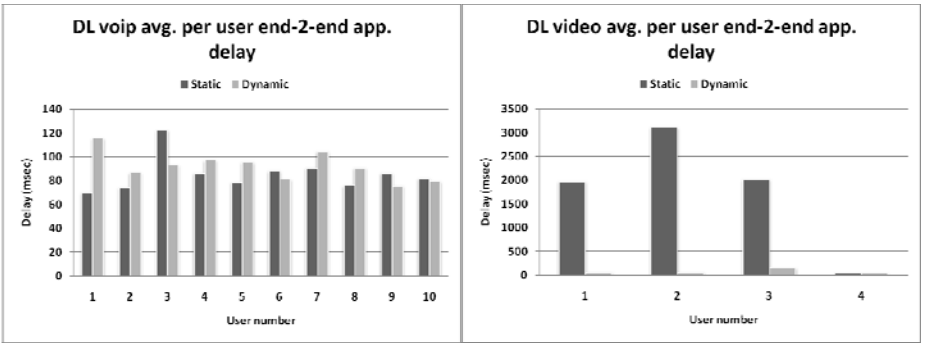


Fig. 8. Cell1 virtual operator 2 per user end to end application delay (m sec)

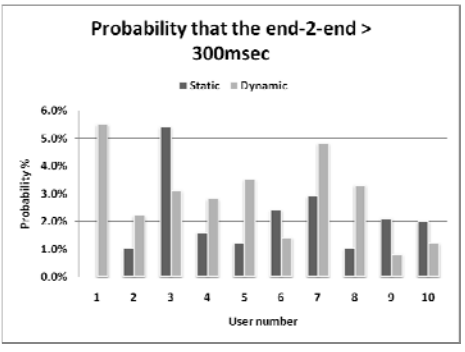


Fig. 9. Cell1 virtual operator 2 end to end delay probability > 300 m sec

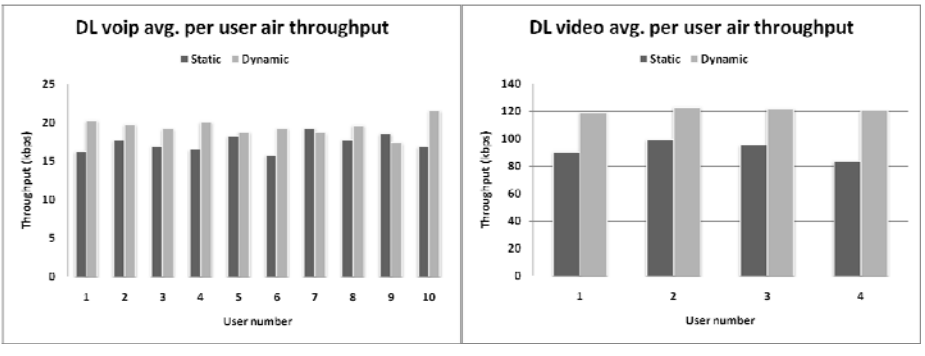


Fig. 10. Cell1 virtual operator 3 per user air interface throughput (kbps)

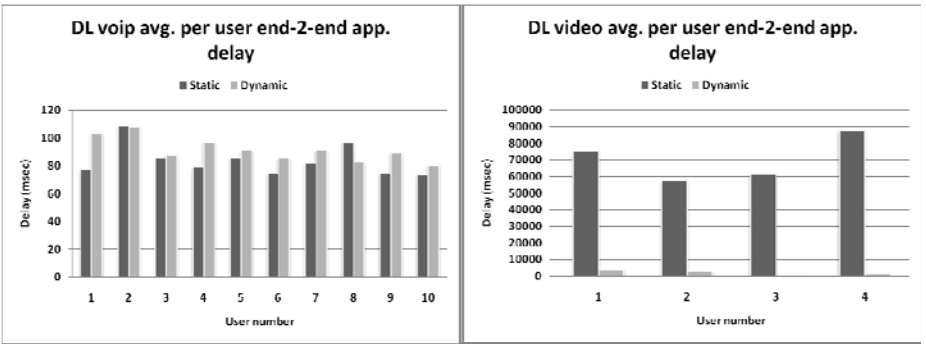


Fig. 11. Cell1 virtual operator 3 per user end to end application delay (m sec)

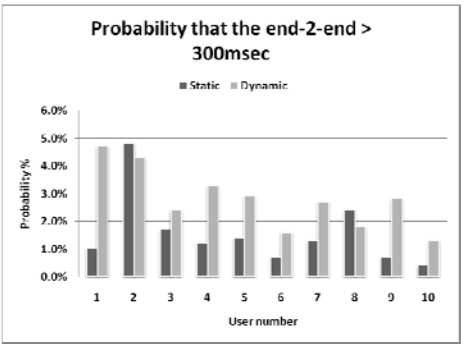


Fig. 12. Cell1 virtual operator 3 end to end delay probability > 300 m sec

As a summary, the simulation results show that using the dynamic spectrum sharing hypervisor compared to the static case is of advantage. This mainly shows in the better quality of service of the users and in particular video users receive. In the dynamic case the video users do not suffer from the huge end-to-end delay experienced in the static scenario. Keeping in mind that both scenarios used the same total amount of air interface resources and same configuration, the only difference is that in the dynamic case a better overall resource allocation has been achieved and no resources have been wasted as in the static one.

Some of the results (especially the ones related to the VOIP users) showed that the static scenario has a slightly better performance mainly when it came to the end-to-end delays. The reason this is that in the static case there are not sufficient overall resources to serve all of the video and VOIP users. Since the video traffic was put in a higher MAC priority class with higher data rates to be served first, and due to not having enough resources to fulfill all of the video user requirements in all of the TTIs, more free resources were left available to serve the VOIP users which explain the better performance. Whereas in the dynamic case, there are sufficient resources to serve all of the video users, this means fewer resources for the VOIP users.

6 Conclusion and Outlook

From the introduction of the LTE virtualization, it can be expected that better system performance can be achieved by dynamic spectrum sharing, and the simulation results confirmed this expectation. Based on the instantaneous traffic load each of the virtual operators can require more resources (PRBs) from the other virtual operator when they still have spare resources. In this way, the overall resource utilization is enhanced and in turn the performance of both network and end-user perspective is better. Although the simulation results are quite scenario-specific, the basic findings are representative and show the advantages that can be achieved by applying network virtualization to the LTE system. In addition for some cases especially in rural areas with low density of population and traffic using the dynamic spectrum sharing it is a much better choice than in today's static spectrum allocation.

This work is a starting point of the LTE virtualization, as there are more issues to be addressed, e.g. interference coordination among multiple virtual operators, signaling overhead due to the hypervisor in charge of the resource allocation, defining guidelines and scheduling disciplines for the hypervisor based on different criteria/contracts and more diverse simulation scenarios. Nevertheless, with LTE wireless virtualization operators can expect not only lower investment for flexible network deployment but also lower costs for network management and maintenance, meanwhile the end-user can expect better services with lower prices in the future.

Acknowledgments

We would like to thank all members of the 4WARD project, especially the virtualization work package.

References

1. GENI Planning Group, GENI: Conceptual Design, Project Execution Plan, GENI Design Document 06-07 (January 2006), http://www.geni.net/GDD/GDD-06-07.pdf
2. Feamster, N., Gao, L., Rexford, J.: How to lease the Internet in your spare time. ACM SIGCOMM Computer Communications Review 37(1) (January 2007)
3. Chapin, J.: Overview of Vanu Software Radio (2009), http://www.vanu.com
4. Niebert, N., Baucke, S., El-Khayat, I., et al.: The way 4WARD to the creation of a Future Internet. In: ICT Mobile Summit, Stockholm (June 2008)
5. AKARI Architecture Conceptual Design for New Generation Network [1.1], http://akari-project.nict.go.jp/eng/concept-design/AKARI_fulltext_e_translated_version_1_1.pdf
6. Asia Future Internet (AsiaFI), http://www.asiafi.net
7. OPNET website, http://www.opnet.com
8. Anas, M., Calabrese, F.D., Mogensen, P.E., Rosa, C., Pedersen, K.I.: Performance Evaluation of Received Signal Strength Based Hard Handover for UTRAN LTE. In: IEEE 65th of Vehicular Technology Conference, VTC 2007-Spring (2007)
9. Westman, E.: Calibration and Evaluation of the Exponential Effective SINR Mapping (EESM) in 802.16. Master Thesis. Stockholm, Sweden (2006)
10. Agrawal, R., Berry, R., Jianwei, H., Subramanian, V.: Optimal Scheduling for OFDMA Systems. In: Fortieth Asilomar Conference of Signals, Systems and Computers, ACSSC 2006 (2006)
11. Agarwal, R., Majjigi, V., Vannithamby, R., Cioffi, J.: Efficient Scheduling for Heterogeneous Services in OFDMA Downlink. In: IEEE Globecom 2007, Washington D.C. (2007)
12. Einhaus, M., Klein, O.: Performance Evaluation of a Basic OFDMA Scheduling Algorithm for Packet Data Transmissions. In: ISCC 2006, Cagliary, Italy (2006)
13. Einhaus, M., Klein, O., Walke, B.: Comparison of OFDMA Resource Scheduling Strategies with Fair Allocation of Capacity. In: 2008 5th IEEE Consumer Communications & Networking Conference (CCNC 2008), Las Vegas, NV (January 2008)
14. Subramanian, A.P., Gupta, H., Das, S.R., Buddhikot, M.M.: Fast Spectrum Allocation in Coordinated Dynamic Spectrum Access Based Cellular Networks. In: DySPAN 2007, Dublin, Ireland (April 2007)
15. Village, J.A., Worrall, K.P., Crawford, D.I.: 3G Shared Infrastructure. In: 3G Mobile Communication Technologies (May 2002)

Accurate Modelling of OFDMA Transmission Technique Using IEEE 802.16m Recommendations for WiMAX Network Simulator Design

Marco Miozzo and Faouzi Bader

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Parc Mediterrani de la Tecnologia (PMT)
Av. Carl Friedrich Gauss 7 - 08860 - Castelldefels (Barcelona) - Spain
{marco.miozzo,faouzi.bader}@cttc.es

Abstract. Worldwide Interoperability for Microwave Access (WiMAX) is the name selected by WiMAX Forum for referring to the standard defined by the IEEE 802.16 task force. The standard introduces several interesting novelties both from PHY and MAC perspective which lead to a complex architecture. In order to understand and investigate its potentialities, analysis is needed. Due to its intrinsic complicated architecture nature, mathematical models may be only applied to portion of the whole system. The same has done for simulation with link level, system and network simulators. However, the new research requirements impose that the model has to be more comprehensive as possible, in order to take care of all the interactions, from physical to application layer. In this paper we propose a novel library for the Miracle extension for ns2 simulator in which, by means of link-to-system mapping (LSM) techniques, the level of details in the PHY layer to be simulated is tuneable in order to take in consideration its important phenomena in a network simulator.

Keywords: IEEE 802.16, WiMAX, OFDMA, ns2.

1 Introduction

In the last years the market for Broadband Wireless Access (BWA) started to grow and nowadays seems to be very attractive for the future, In fact, we are assisting to the proliferation of several new applications and services which requires high quality of service (QoS) and large bandwidth connections. Up to now, the technology which dominates the market is the High-Speed Downlink Packet Access (HSDPA), an extension of Universal Mobile Telecommunications System (UMTS) [1], due to the easy installation (i.e., via USB key) and to the widespread of UMTS accesses. However nowadays, many new technologies are under evaluation in order to obtain better performance thanks to the recent innovations in transmission techniques. Worldwide Interoperability for Microwave Access (WiMAX) and Long Term Evolution [2] (LTE) are the two most interesting ones. The latter one is a standard defined by 3GPP to convey the UMTS radio technology towards the 4G network view, i.e., the next generation of high data rate access network totally based on IP

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 258–269, 2011.

flows. These goals are also the ones of WiMAX and it seems that it will be an enthralling challenge on which one will win on the market.

The IEEE 802.16 standard [3] defines the specifications for the WiMAX radio technology related to the lower three layers of the OSI protocol stack. The network is composed by base stations (BSs) in charge of providing connectivity to subscriber stations (SSs) over superframe according to the compulsory procedures defined within the standard. Recently, relay stations (RS) have been introduced to support the BS in the service provisioning. The first functional WiMAX wireless air interface was defined by the IEEE 802.16-2004 [3] targeting only the fixed wireless broadband access systems. Later, with 802.16-2005 [4] (also called 802.16e), it was introduced the support for mobile functionalities thanks to the OFDMA radio transmission technique. With the IEEE 802.16j [5] standard released in the 2006, relay technology has been introduced in the WiMAX architecture. At the time of this writing, the IEEE 802.16 task force is working on the 802.16m draft in the definition of an advanced air interface for high data rates (e.g., 100 Mbit/s mobile and 1 Gbit/s fixed). However, many part of the architecture have been left open to vendor specific implementation (e.g., the scheduling of the packets for the different class of users, the hybrid ARQ schemes, the physical allocation to the users and many others). This means that there is a considerable need of instruments to correctly test the actual performance of the existing solutions and to investigate new enhancements. According to the specific topic under research, this need can be limited to a specific protocol or layer of the OSI stack; for instance, a link level simulator is the best solution for evaluating the coding and modulation performance and more in general PHY aspects. The problem becomes more argued in the upper layers, where the common practice is the adoption of system level simulators or network level ones. The former one models mostly the hardware physical constrains and the link level; in fact, it is commonly exploited by industry during the development of new equipments. Since this kind of solution is usually designed for a specific implementation, it models the hardware constraints and generally does not consider the whole protocol stack interactions. Finally, network simulators account for the whole OSI stack allowing the definition of realistic network topologies and they do not consider hardware aspects. This allows the definitions of scenarios with base stations, servers, routers and clients which exchange flows of data among them, making possible to collect end-to-end performance. We want to note that, network and system level simulators usually rely on the performance obtained by link level simulators in order to model the PHY layer. On this matter, the level of abstraction adopted may differ a lot and we consider that this assumption is becoming a fundamental feature in simulation tools considering the level of details required by most of the new research fields. For instance, the interference and in general the channel state information represent crucial parameters in optimization algorithms at several layers. A few samples are: SINR adaptive coding scheme, TCP optimization and in general smart routing and radio resource management (RRM) schemes.

Democles® project [6] (dynamic resource management for advanced multiple carrier system platform) started with the aim of developing a framework for simulating next generation of wireless networks (e.g., 4G) which exploit multi-carrier techniques. This project also focuses on RRM functionalities in order to better exploit all the PHY and MAC layer characteristics and satisfy the QoS needs of the services

by means of cross-layer algorithms. In order to address these goals and considering the discussions above, we decided to implement a simulator for WiMAX networks in the context of this project starting from the Miracle [7, 8] extension for ns2 [9]. This work is called WiMAX for Democles® (WiDe). The rest of the paper is organized as follows. Section 2 introduces the IEEE 802.16 standard. In Section 3 we propose a brief overview of other 802.16 networks simulators. In Section 4 we discuss on WiDe module and we detailed out the implementation process. We note that, we do not present specific implementation details, but rather argue on our simulation methodology in order to highlight the advantages and the limitations due to the complexity adopted. In Section 5 we present interesting research scenarios which can be opportunely described by means of WiDe. Finally, Section 6 concludes the paper.

2 Standard Overview

In this section we summarize the principles of the IEEE 802.16 standard, for a comprehensive view of the architecture the reader may refer to [4, 5]. WiMAX supports both point-to-multipoint (PMP) architecture and mesh topology. The two types of communication systems are similar from PHY layer point of view; they differ mostly in some MAC procedures in order to enable the support to RSs. These changes will be highlighted in this section. The standard defines multiple PHY layers according to the application environment, the most used are: WirelessMAN-OFDM exploiting orthogonal frequency division multiplexing technique (OFDM) and WirelessMAN-OFDMA using orthogonal frequency division multiple access scheme (OFDMA). The PHY layer of WiMAX is organized in frames of fixed length. According to the TDD mode, each frame is divided into two subframes to guarantee the bidirectional communications (i.e., downlink and uplink). In case of relay mode, each subframe is divided into one or more access zones and relay zones. The downlink/uplink access zones are allocated to the transmissions between SS and their access point. The MAC layer is divided into three sublayers: the convergence sublayer (CS), the common part sublayer (CPS) and the security sublayers. CS is in charge of the classification and the mapping of the incoming packets from the upper layers and their transmission to the CPS where classical MAC procedures are applied. The main CPS functionalities are: connection establishment and management, generation of MAC signalling, service flow management and scheduling. One of the main roles of MAC signalling is the negotiation of the bandwidth. This can be done with stand alone bandwidth request messages or piggybacked in data packets. The uplink scheduler at BS side decides which SSs among the ones have requested bandwidth can transmit in the next uplink subframe. Similarly, BS scheduler picks up the packets to transmit in the downlink subframe according to the scheduling services (or QoS classes).

The MAC protocol is connection-oriented: all traffic is mapped onto connections which are uniquely identified by the connection identifier (CID). The registration phase is a two way handshake procedure called *initial ranging*. The downlink map (DL-MAP) and uplink map (UL-MAP) are broadcasted each frame by BS in order to indicate how the accesses have to be managed in the current frame. The downlink channel descriptor (DCD) provides the burst profiles (physical parameter sets) that

can be used by a downlink physical channel during a burst, in addition to other useful downlink parameters. The uplink channel descriptor (UCD) does the same as the DCD for the uplink subframe.

In order to transmit and receive data of the service requested, the SS has to establish a connection. Each service flow defines a unidirectional flow of data traffic and is characterized by a set of QoS parameters. In mesh mode, from MAC layer perspective an RS has two operative modes. A transparent RS (T-RS), in which RS does not have to transmit any control messages. In this case, SSs receive broadcast signalling from BS and they are not aware of the RS (i.e., there is not logical connection established), which are in charge to only relay data traffic. In case of non-transparent RS (NT-RS), the RS has to broadcast management messages in its relay zone, that is: the relay DL-MAP (RDL-MAP) and relay UL-MAP (RUL-MAP). This implies that SSs are logically connected to their RS instead of the BS.

3 Related Work

In this section we present a brief overview of the existent WiMAX modules developed for network simulators. We would like to stress that, in this section we are not going to detail the features and gives a practical comparison among them, rather then we are more interested to examine and highlight their approach. The NIST module [10] is one of first developed as extension of the ns2 simulator. It provides functionalities for 802.16 MAC, handover and scheduling. The main drawbacks are that it provides only a simplistic OFDM PHY layer and the absence of an ARQ scheme. The NDSL module [11] focuses mostly on MAC functionalities referring to OFDMA PHY layer, which however is model in a very high level fashion. There are many other extensions provided for ns2, a more accurate exposition can be found in [12]. Recently has been released also a module for the brand new NS3 simulator [13]. The module presents many interesting MAC functionalities but it does not implement any packet error model and it models only the OFDM transmission scheme.

All the modules presented above have a common characteristic: they implement the disk propagation model. In respect to this, many works on IEEE 802.11 have just demonstrated that this model is far from addressing sufficient PHY aspects. In fact, it is designed for a single carrier case and it models the error distribution with a threshold on the power received. However, when we consider modern transmission techniques, such as OFDM, we are referring to multi-carriers systems, where the data is spread over several subcarriers and this assumption might be simplistic. Since these subcarriers are placed at different frequencies, they experience different propagation behaviours and frequency selectivity which implies different degradations.

One of the first modules that try to relax this assumption is the one developed by WiMAX Forum starting from NIST module which, however, is available only to consortium members. Finally also WINSE [12] module is aware of this problematic. WINSE seems to be one of the most complete modules for IEEE 802.16. It has most of the functionalities counted by the standard for the MAC layer, it accurately models both the OFDM and OFDMA PHY layer by exploiting propagation traces generated with dynamic system simulators. However, this solution implies that only scenarios pre-simulated by the system simulator can be then simulated.

4 WIDE

4.1 General Aspects

As done by WiMAX Forum, we implemented WiDe starting from NIST module as extension of the well known network simulator ns2. This choice is based on the fact that NIST extension represents a good solution from MAC layer point of view and, moreover, it has just been integrated in Miracle framework by University of Karlstad [14]. Thanks to the latter feature, this module inherits all the functionalities provided by Miracle. This enables the developing of a fully integrable module without make any changes to the ns2 core, besides it can coexist with other radio technology modules and therefore it allows to simulate scenarios where node are equipped with many of them simultaneously. In fact, WiDe is a simple library which has to be dynamically loaded in the Tcl simulation script (see [7] for more info on dynamic libraries). Another important feature inherited is the support for the PHY layer modelling, which we opportunely adapted to multi-carrier techniques. Finally, a cross layer message engine provides an efficient way to exchange info between MAC and PHY layer, but also to potentially all layers, enabling the easy implementation of RRM schemes. A diagram of WiDe implementation architecture is given in Figure 1.

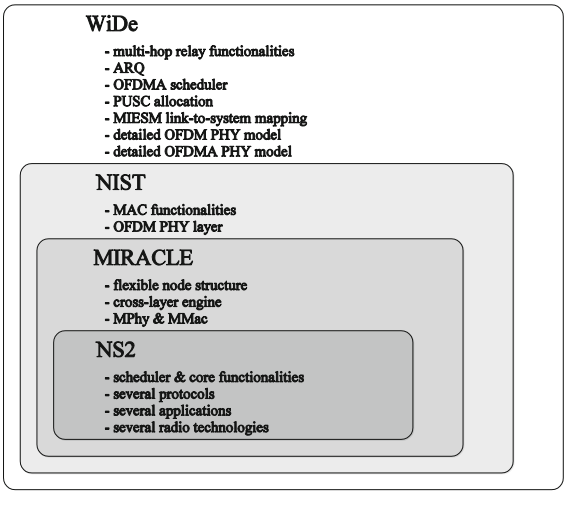


Fig. 1. Diagram of WiDe implementation architecture

4.2 WiDe PHY Layer

In this subsection we explain the solutions we adopted in the implementation of the PHY layer in WiDe. As anticipated in Section 3, OFDM transmissions are affected by complex propagation phenomena due to the time-frequency selective channel nature. This implies that subcarriers may experience frequency selective fading and therefore different channel gains one from each other. We have also to consider that we need to

find a trade off between the computational complexity of the simulator and the abstraction level in order to be able to adapt the framework to the specific research scenario we are considering. For instance, in link adaptation schemes and channel aware scheduling algorithms is fundamental to have a clear and detailed view of the channel conditions; while in load balancing scheme or, more in general, in large networks evaluations, these aspects can be neglected (or, usually, have to be relaxed to reduce the computational complexity due to the intrinsic complexity of the scenario). This led us to develop a PHY layer in which the level of complexity is tuneable in order to satisfy different research needs with a reasonable simulation time. This has been obtained by modelling the channel with a different number of logical subchannels, where with modelling we mean the evaluation of the SINR per packet level considering both channel propagation phenomena and the interference perceived counting for all the on-going transmissions. We identified two main different levels of complexity:

- BASIC (standard ns2 behaviour): one single channel simulates all the subcarriers.
- FULL: all the data subcarriers are modelled.

The first model mimics the standard ns2 behaviour where all the transmissions are simulated as performed in the same carrier. This model can be still considered a valid approach in fixed WiMAX, where standard OFDM schemes are applied and therefore all the subcarriers are used simultaneously for the same transmission with the cost of neglecting the frequency selectivity phenomenon. The latter model, called FULL, allows modelling all the aspects involved; in fact, interference and channel gain is counted for each subcarrier. Thanks to the SINR evaluated through the process described above, we implemented an error distribution model which estimates the errors according to the actual SINR frequency profile perceived by the radio during the packet reception. This is done interfacing WiDe with a link level simulator [15] by means of link-to-system mapping (LSM) technique. This allows the relaxing of one of the strongest assumption usually adopted in network simulators: the disk propagation model. This model has a critical limitation: all the transmission schemes have the same performance from error distribution as function of the received power point of view. This behaviour is due to the fact this error reception model marks as corrupted all the packets which have the SINR under a certain threshold, unique for all the transmission types. For instance, the NIST module adopts this solution; in fact, it defines a unique reception threshold below which all the packets are considered corrupted. This implies that the system models all the modulation schemes with the same energy robustness; therefore, for instance, it does not make difference to transmit with the 64-QAM respect to the QPSK from error distribution perspective. We would like to note that, this is the approach adopted by WiMAX Forum [17].

Finally, in order to carefully model the propagation phenomena, the system accounts for fading, shadowing and path loss. Fading is modelled thanks to the Jakes Simulator [17]. Shadowing is modelled according to the Gudmunson model [17], and path loss according to the Hata model [18].

The link level curves exploited are generated by a WiMAX OFDM link level simulator assuming a frequency flat response at given SINR, therefore, in order to

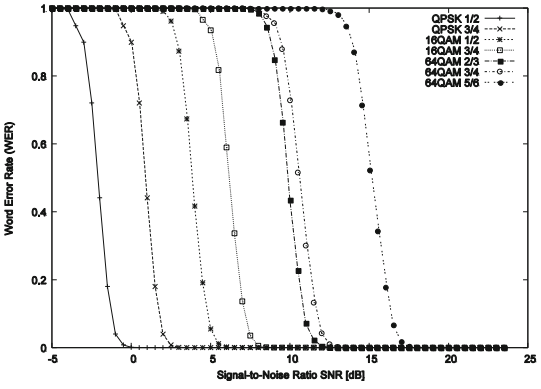


Fig. 2. Example of Word Error Rate curve (example for codewords of 14 bytes)

map SINR values coming from the WiDe PHY layer onto that curves, we need a specific mapping scheme which takes into account of this. This technique is called *effective SINR mapping (ESM)*.

In practice, ESM maps the vector of the N SINR values received $\{\gamma_1 \ldots \gamma_N\}$ into a single effective SINR (ESNR) value which can be further used to estimate the block error rate (BLER) according to the coding performance curves; in formula, ESNR is approximated with functions of the type

$$\gamma = \theta^{-1}\left(\frac{1}{N}\sum_{n=1}^{N}\theta(\gamma_n)\right)$$

where γ is the ESNR, N is the number of subcarriers, γ_n is the SINR perceived by the n and the particular function θ depends on the wireless technology being used. Several ESM schemes has been proposed to model the link performance, the most interesting ones are: exponential-effective SINR mapping (EESM) [19] and mutual information effective SINR mapping (MIESM) [20]. According to [20], MIESM is the one with the best performance from PER prediction accuracy point of view and therefore we have selected it in our implementation. The idea behind this solution is to compute mutual information metric based on the samples of SINRs of the different subcarriers as function of the specific modulation, in our case it is called received bit mutual information rate (RBIR). Thanks to RBIR value we may have the word error rate (WER) through the curves of the coding simulator [15]; a sketch of the MIESM approach is given in Figure 3. An example of WER curve is given in Figure 3 where it is clearly depicted how the modulation and coding scheme strongly impact on the performance in the reception and justify our concerns on the single threshold model, previously adopted. This is clearly demonstrated by Figure 4, where the standard disk propagation model is adopted for simulating the data rate perceived by the application in a downlink connection with the OFDMA PHY layer. In Figure 4 we may observe that all the modulation and coding schemes have the performance as function of the distance between the SS and the BS, this is due to the fact that the reception model is a single threshold on the received power unique for all the transmission profiles. In Figure 5 we plot the results for the same simulation scenario with the LSM scheme

implemented and the impact of the introduction of the WER curves is not negligible in the in the end-to-end performance; in fact the curves follow the same behaviour of the respective ones with the flat response presented in Figure 2.

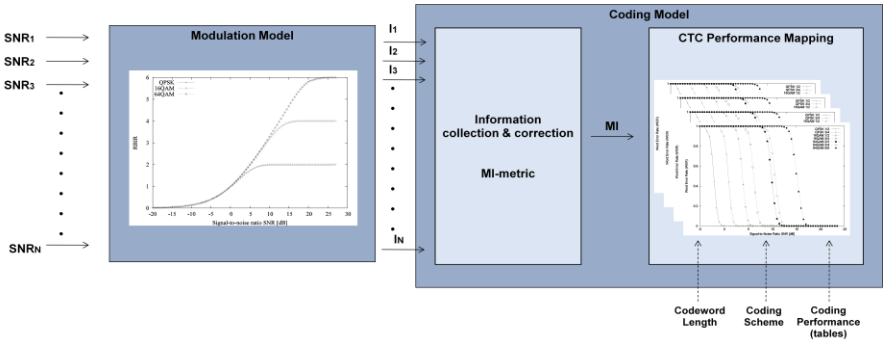


Fig. 3. Diagram of the MIESM link-to-system mapping procedure

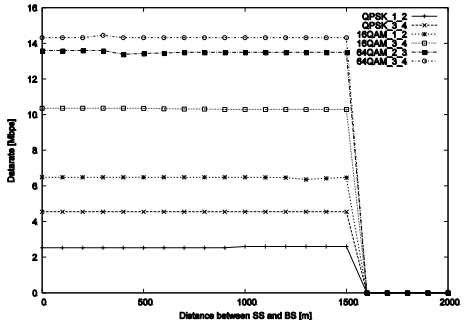


Fig. 4. Data rate perceived in the downlink connection with disk propagation model

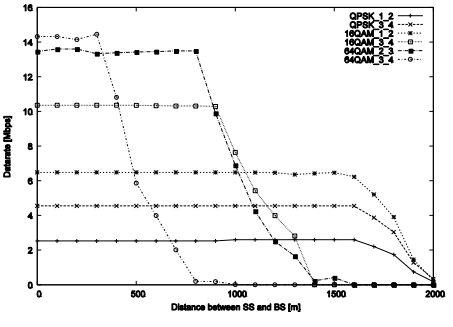


Fig. 5. Data rate perceived in downlink connection with LSM approach

The OFDM transmission scheme is the first real implementation of WiMAX and is the one which nowadays is exploited in all the fixed WiMAX commercial equipments. It supports both TDD and FDD, but in our model only TDD is implemented. It exploits 256 subcarriers and they can be used only simultaneously by the transmission entity. In this case, starting the NIST implementation of the OFDM PHY layer, we improved error model according to the LSM technique described above. In Table 1 we report the main OFDM parameters of WiDe. The OFDMA layer model is the standard adopted by mobile WiMAX and is the model we adopt as reference in our implementation. In case of TDD combined with OFDMA techniques allows to duplex transmissions both in time and frequency. The latter one is obtained by allocating different set of subchannels to each transmission. The standard defines several combination of number of subcarriers supported, in order to provide scalability features. In the particular version of the standard we are referring to (i.e., IEEE802.16m) the mandatory number of the subcarriers is fixed to 1024. From the

1024 subcarriers, we have to remove the ones exploited in the guard bands and the one reserved as central null subcarrier. The result is that OFDMA guarantees 840 subcarriers for actual information transmissions (i.e., data and pilot subcarriers). In order to manage the allocation of the subchannels, several sub-channelization schemes have been defined. In the following we concentrate in the partial usage of the subchannels (PUSC), which is the one adopted in WiDe. PUSC defines 30 logical suchannels, each of one composed of 24 subcarriers for the data and 2 reserved for pilot signalling. The subcarriers involved in the subchannels are selected in a non adjacent fashion with a technique of two levels of permutation and grouping. This is due in order to obtain incorrelation among the subcarriers exploited in the same subchannel.

In our implementation we have modelled PUSC sub-channelization and we are able therefore to track the transmissions on subcarrier level. In this case, we implement a third level of complexity called PUSC which models 30 channels as the 30 logical channels defined by PUSC. This implies that the interference can be accounted per subchannel level but still the frequency selectivity is not modelled. This model can be useful for simulating scheduling scheme in single cell scenarios where the subchannels are common for all the devices.

In Figure 6 we provide statistics on the computational complexity of this technique considering the OFDMA PHY layer. We plotted the simulation time as function of the number of the users for a single simulation run in a single processor Intel Pentium 4 machine. Consider that, the curves refer to a simulation of 30 seconds of simultaneous transmissions of the users involved. Since the rate of the data to be transmitted is set to saturate the channel, we may consider this as a worst case scenario. From the curves we can see how the computational complexity grows with the number of subchannels/ subcarriers simulated as expected. With the FULL model, the simulation time increases very rapidly and therefore it might be difficult to use it in large simulation scenarios. However, it can be still useful to test small scenarios in which a high level of details simulated is required; for instance in interference limitation scenarios such as: compatibility tests between radio transmission technologies and channel aware packet scheduling algorithms. Thanks to the PUSC model we instead have, at a reasonable simulation time overhead, enough information to correctly model the interference in the different subchannels.

4.3 WiDe MAC Layer

In the MAC layer, we worked mostly on addressing the limitations of the NIST library. Due to the realistic error model introduced, we implemented an ARQ scheme in order to mitigate its effect on the end-to-end performance. The standard ARQ implementation in WiMAX, considers dividing the flow into blocks of variable in order to identify the portion of the flow which lacks at the receiver side and subsequently asks for their retransmission. Retransmissions are requested by the receiver according to different policies, the actual algorithm implementation is left to vendor definition. In our implementation a receiver can acknowledge set of adjacent blocks received correctly in a cumulative fashion or specify the sequence of blocks received correctly but non adjacent thanks to the sequence maps.

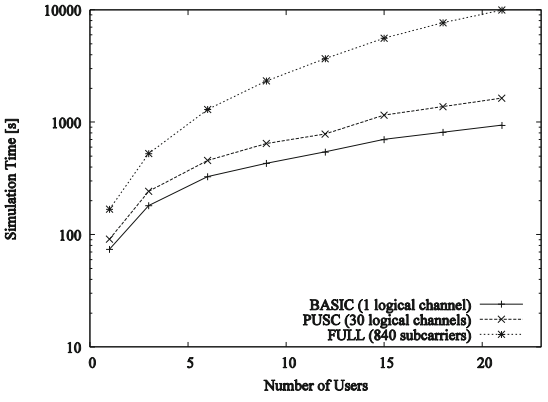


Fig. 6. Computational complexity for OFDMA PHY layer

Finally, we introduced all the functionalities to manage the relay node in mesh mode. In this case we designed two new entities: the relay stations and the subscriber station connected to a relay station. This implies also the extension of the functionality of scheduling from the BS also to the non-transparent RS in order to allow the transmissions both in access and relay zone. In Table 1 we report the main MAC parameters of WiDe.

Table 1. Wide principal features

PHY Layer
SINR traced per each packet in fly at subcarrier grade
MIESM link-to-system mapping
Packet Error Model with modulation and CTC
OFDM with 1 channel
OFDM with 256 simulated subcarriers
OFDMA with 1 channel (BASIC mode)
OFDMA with 30 PUSC subchannels (PUSC mode)
OFDMA with 1024 simulated subcarriers (FULL mode)
MAC Layer
MAC management messages: DL-MAP, UL-MAP; DCD, UCD
Relay functionalities
ARQ scheme (feedbacks and transmission window)
ARQ tunable size blocks
Bandwidth request: standalone and piggy-backed
Best Effort BS scheduler
SS scheduler
Connectivity Service Network
Network entry procedures (initial ranging and registration)
Connection establishment messaging: dynamic service management
Handover

5 Research Scenarios

Thanks to Miracle framework, WiDe inherits a very flexible definition of the node architecture. For instance, it is possible to define mobile device equipped with

WiMAX and also other radio technologies. On top of the network level, several transport protocols can be used, such as Transport Control Protocol (TCP), User Datagram Protocol (UDP) and Real Time Protocol (RTP). Regarding the latter, a set of applications are just ready to be connected to transport layer, such as: VoIP and Video codec (both of them incorporate instruments to evaluate the quality perceived after the decodification) with constant bit rate (CBR) and conversation-like behaviour. For instance, we have integrated in Wide the support for the Evalvid tool [21], which is a set of applications designed to manage video streaming flows, simulate their transmission and collect end-to-end statistics, such as picture signal to noise ratio (PSNR), packets losses and jitter.

Cross-layer messaging represents a suitable framework for the improvement of several algorithms, especially when combined with a detailed view of the link and network condition, as described in [22].

From users' point of view several research topics can be carefully investigated. For instance, scheduling algorithms can take advantage of the information on the actual conditions of the channel at subcarriers grade in order to implement more efficient allocation schemes. In fact, WiDe PHY layer models the channel at single subcarrier level both from propagation phenomena and interference perspectives. This paradigm can be exploited also by upper layers, where RRM modules, transport protocols and applications can adapt their algorithms to the channel or radio conditions (e.g., smarter TCP transmission windows updating, variable bit rate video codec and triggers to RRM handover decision making policies).

Finally, thanks to WiDe we may now evaluate relay architectures by carefully taking in account for the interference in the whole system and exploiting the relay functionalities of IEEE 802.16j. One of the big challenges is to find how to optimally split the intelligence between cognitive terminals and cognitive networks. From network perspective, schedule transmissions both in time and frequency [23] among the entities is another interesting research topic by considering a more flexible partitioning of the frequency bands thanks to admission of a tolerant interference. Both of the last points allow further to consider also energy saving problem, part of the emerging research field known as *green communication*, an evergreen topic in wireless systems due to the intrinsic battery limitations of the handleable devices.

6 Conclusions

In this paper we presented a novel implementation of WiMAX for network simulation called WiDe. According to the trend of the research community, we carefully implemented the PHY layer with a tuneable level of detail specification in order to take into account for interference and propagation phenomena up to subcarriers grade. We demonstrated that this does not introduce too much computational complexity. In fact, with a full level of details it is still possible to simulate intra cell scenarios with a reasonable simulation time. Thanks to these features, WiDe allows the accurate simulation of several new research scenarios, such as cross-layer optimization schemes and green communication, where the knowledge of PHY conditions at simulation run-time is a crucial aspect.

References

1. Holma, H., Toskala, A.: HSDPA/HSUPA for UMTS. John Wiley & Sons, Chichester (2006)
2. Sesia, S., Toufik, I., Baker, M.: LTE – The UMTS Long Term Evolution - From Theory to Practice. John Wiley & Sons, Chichester (2009)
3. IEEE Std. 802.16-2004, IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems (October 2004)
4. IEEE Std. 802.16e-2005, IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems - Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in License Bands (February 2006)
5. IEEE Std. 802.16j, Air Interface for Fixed Broadband Wireless Access Systems: Multihop relay specifications (July 2008)
6. DEMOCLES® - Simulation-based testbed for dynamic RRM, `http://www.cttc.es/en/projects/testbeds/project/DEMOCLES.jsp`
7. ns-MIRACLE: Multi InteRfAce Cross Layer Extension for ns-2, `http://telecom.dei.unipd.it/download`
8. Baldo, N., Miozzo, M., Guerra, F., Rossi, M., Zorzi, M.: Miracle: The Multi-Interface Cross-Layer Extension of ns2. EURASIP Journal on Wireless Communications and Networking, Article ID 761792, 16 pages (2010)
9. The Network Simulator - ns2, `http://www.isi.edu/nsnam/ns/`
10. The Network Simulator ns-2 NIST add-on - IEEE 802.16 model (MAC+PHY), Technical report, National Institute of Standards and Technology (June 2007)
11. Chen, J., Wang, C.-C., Tsai, F.C.-D., Chang, C.-W., Liu, S.-S., Guo, J., Lien, W.-J., Sum, J.-H., Hung, C.-H.: The design and implementation of WiMAX module for ns-2 simulator. In: WNS2, Pisa, Italy (October 2006)
12. Sayenko, A., Alanen, O., Martikainen, H., Tykhomyrov, V., Puchko, O.: WINSE: WiMAX NS-2 Extension. In: SIMUTools, Rome, Italy (March 2009)
13. Farooq, J., Turletti, T.: An IEEE 802.16 WiMAX Module for the NS-3 Simulator. In: SIMUTools, Rome, Italy (March 2009)
14. NS2 Miracle Wimax, `http://sourceforge.net/projects/ns2miraclewimax/`
15. Pfletschinger, S., et al.: D2.2.3 Modulation and Coding schemes for the WINNER I System. IST-4-027756 WINNER II
16. WiMAX Forum White Paper, WiMAX System Evaluation Methodology (July 2008)
17. Stüber, G.L.: Principles of Mobile Communication. Kluwer Academic Publishers, Dordrecht (2003)
18. Goldsmith, A.: Wireless Communications. Cambridge University Press, Cambridge (2005)
19. Mumtaz, S., Gamerio, A., Rodriguez, J.: EESM for IEEE 802.16e: WiMaX. In: ICIS, Portland, Oregon, USA (May 2008)
20. Brueninghaus, K., Astely, D., Salzer, T., Visuri, S., Alexiou, A., Karger, S., Seraji, G.A.: Link performance models for system level simulations of broadband radio access systems. In: IEEE PIMRC 2005, Berlin, Germany (September 2005)
21. EvalVid - A Video Quality Evaluation Tool-set, `http://www.tkn.tu-berlin.de/research/evalvid/`
22. Akyildiz, I.F., Wang, X.: Cross-Layer Design in Wireless Mesh Networks. IEEE Transactions on Vehicular Technology 57(2), 1061–1076 (2008)
23. Debbah, M.: Mobile Flexible Networks: The challenges ahead. In: IEEE ATC (October 2008)

A Simulation Implementation of the LTE-Uu Interface Datalink Layer in OMNeT++

Mohammad Arouri[1], Ziyad Atiyyeh[1], Anas Mousa[1],
Amna Eleyan[1], and Hussein Badr[2]

[1] Computer Systems Engineering Department
Birzeit University, Birzeit, Palestine
{mohammadarouri,ziyad.a.2010,anas.n.mousa}@gmail.com,
aeleyan@birzeit.edu
[2] Department of Computer Science
University at Stony Brook, Stony Brook, New York, U.S.A.
badr@cs.stonybrook.edu

Abstract. The 3rd Generation Partnership Project (3GPP)'s Long Term Evolution (LTE) standards define the next major step in the evolution of cellular systems towards higher data rates, low latency, and greater spectral efficiency. This occurs in the context of a System Architecture Evolution (SAE) that specifies a packet-switched IP architecture for both voice and data transmission. We present a simulation implementation of a key component of the overall LTE-SAE: the L2 (Datalink) layer of the LTE-Uu interface between mobile User Equipment (UE) and a base station eNodeB. The simulation is developed for the INET Framework of OMNeT++4.0, an open-source computer-network simulation environment, and implements the Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), and Medium Access Control (MAC) layers of the LTE-Uu. The implementation is extensible, and is intended to serve as a publicly-available, open-source platform for further simulation development of various aspects of LTE-SAE.

Keywords: LTE (Long Term Evolution), SAE (System Architecture Evolution), PDCP (Packet Data Convergence Protocol) Layer, RLC (Radio Link Control) Layer, MAC (Medium Access Control) Layer, Simulation, OMNeT++, INET Framework.

1 Introduction

The Long Term Evolution (LTE) standards of the 3rd Generation Partnership Project represent a major development in the evolution of UMTS (Universal Mobile Telecommunication System) beyond 3G (3rd Generation) mobile cellular technology, and aims at providing high data rates, lower latencies, and greater spectral efficiencies. LTE development goes hand-in-hand with SAE (System Architecture Evolution) which defines an AIPN (All IP Network) core network architecture, the EPC (Evolved Packet Core). The major elements of the combined LTE-SAE system, collectively referred to as the EPS (Evolved Packet System), are shown is Fig. 1. The reader is referred to [1] for further details.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 270–284, 2011.

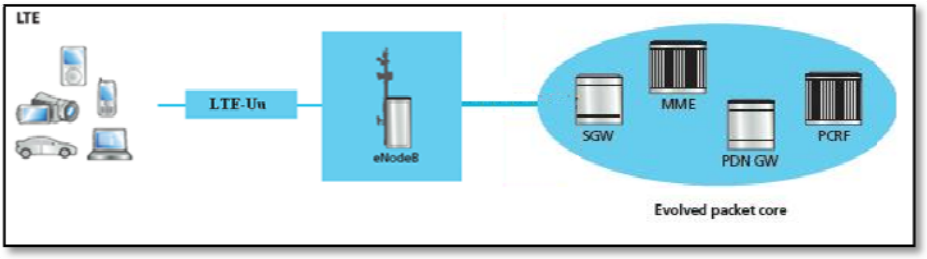


Fig. 1. Elements of the LTE-SAE architecture. The node *PDN GW* is the gateway to a Public Data Network.

In this paper, we present the development of a simulation model for a key component of the LTE-SAE architecture: the L2 (Datalink) layer of the LTE-Uu wireless interface between mobile User Equipment (UE) and a base station eNodeB. The model is implemented as a component that extends the INET Framework [2] of the open-source, extensible, discrete-event network simulator OMNeT++4.0 [3]. Our model design and implementation are themselves also extensible, and are intended to serve as a publicly-available, open-source platform for further simulation development of various aspects of LTE-SAE.

So far as we can tell from the published literature, the only previous work that is similar to ours in breadth and scope is that of Qiu *et al.* [4], in which the authors report on the implementation of an LTE-SAE simulation model using the well-known, open-source network simulator ns-2 [5]. Their work, however, differs from ours in several important respects. Our work focuses on a highly-detailed implementation of the L2 (Datalink) layer protocols (see §2.1 below), strictly in accordance with the relevant defining specifications. The work in [4] models these protocols' dynamics in a more abstract manner using queues. On the other hand, [4] uses a richer and more varied set of traffic classes to report their results than we do. It is probably fair to say that the primary aim of [4] is to provide – quoting the authors – an "accurate enough" simulation model for the study of various "optimization features" for the performance improvement of the LTE-SAE network. Our primary aim, on the other hand, is to provide the research community with an extensible development platform in the form of a specification-compliant simulation implementation of the protocols at a fine granularity of detail.

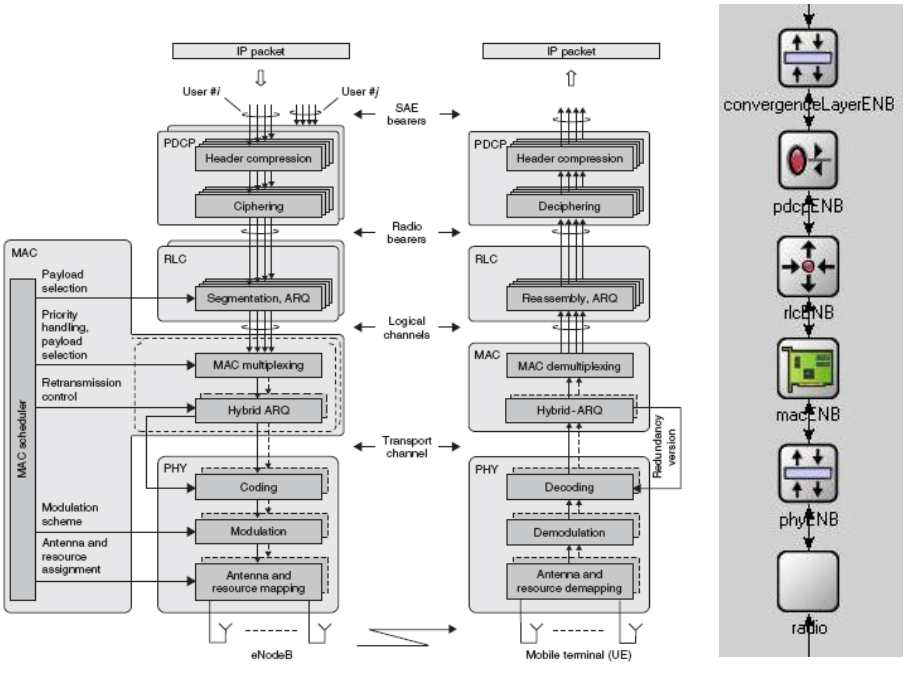
The rest of the paper is structured are follows. Section 2 describes the model design and implementation. Section 3 presents some results from the model. In Section 4 we close with some brief concluding remarks.

2 The Simulation Model

2.1 LTE-Uu Interface

The (user-plane) protocol stack of the LTE-Uu interface is composed of the L2 (Datalink) and L1 (Physical) layers as shown in Fig. 2(a). Fig. 2(b) shows our OMNeT++ implementation. From Fig. 2(a), it can be seen that the L2 layer is composed of three sublayers [1, 6, 7]. From top down:

- **Packet Data Convergence Protocol (PDCP)** [1, 6, 8]. Performs IP header compression/decompression to reduce the number of bits transmitted over the radio link; and ciphering/deciphering where required.



(a) (b)

Fig. 2. (a) The L2/L1 LTE protocol architecture, shown here operating on the Downlink Shared Channel from *eNodeB* to *UE* (reproduced from [6]). Note the *PDCP*, *RLC* and *MAC* sublayers of the L2 layer. *PHY* is the L1 (Physical) layer. (b) INET Framework simulation implementation of the protocol architecture, shown here for the *eNodeB*. The architecture in the *UE* is structurally the same.

- **Radio Link Control (RLC)** [1, 6, 9]. On the outgoing side, it performs segmentation and concatenation of the incoming PDCP protocol data units (PDUs – called RLC Service Data Units (SDUs) from the perspective of the RLC) to produce dynamically-resized, rate-adapted RLC Protocol Data Units (PDUs) for the MAC sublayer, in line with decisions made by the latter's scheduler. On the incoming side, it performs reassembly of incoming MAC PDUs (RLC SDUs).
  The RLC has two modes of operation, Unacknowledged Mode (UM) and Acknowledged Mode (AM). In AM, the RLC implements an ARQ mechanism to guarantee reliable, in-order data transmission on the Downlink Shared Channel (DL-SCH) from eNodeB to UE.
- **Medium Access Control (MAC)** [1, 6, 10]. The MAC sublayer in the eNodeB performs scheduling on both the Downlink (eNodeB to UE) and Uplink (UE to eNodeB) Shared Channels (DL-SCH and UL-SCH, respectively). These are the

main channels for data transmission. The MAC sublayers in the UE and the eNodeB produce dynamically-sized Transport Blocks (TBs) for transmission on these channels, with one TB, if available, being transmitted during a Transmission Time Interval (TTI; typically 1 msec.). In the case of Multiple-Input Multiple-Output (MIMO) antenna spatial multiplexing, more than one TB may be transmitted per TTI, but this is not currently implemented in our simulation. The MAC sublayer may also perform multiplexing/demultiplexing between multiple logical channels (data, control, *etc*.) for transmission on the DL-SCH and UL-SCH, but this feature is also not currently implemented in our simulation.

The MAC sublayer implements a hybrid-ARQ (HARQ) mechanism on the DL-SCH and UL-SCH, in the form of up to eight (the exact number is a parameter of our simulation model) parallel Stop-and-Wait ARQ processes with ACK/NACK signalling in the reverse direction.

Fig. 2(b) shows the protocol architecture implemented in our simulation model, which parallels that of Fig. 2(a). The figure shows the implementation in the eNodeB; the implementation in the UE is structurally exactly the same. At the top of the protocol stack, we introduced a 'convergence sublayer' whose purpose is to 'stitch' the pre-existing INET Framework structure to the LTE sublayers. Its sole function is to implement conversion between pre-existing OMNNeT++ message object types that are passed to and from the next layer up (not shown in the figure; typically, in the UE this would be the Network Layer, and in the eNodeB it would be an INET Framework *relayUnit* module) and those we implement for PDCP SDUs. At the bottom of our stack we have a physical and a radio sublayer. These will be discussed in §2.3 below.

2.2 Connection Establishment

Before data transfer can occur, connection establishment between UE and eNodeB has to take place [1, 6]. In the simulation model, connection establishment consists of a sequence of four phases, as shown in Fig. 3. Some of the activity of these phases pertains to the Radio Resource Control (RRC) layer [1, 6, 7, 11] which lies in the control-plane protocol stack, immediately above the PDCP sublayers shown in Fig. 2, but is not otherwise explicitly represented in our model.

1. **Cell search** [1, 6, 7]. The UE acquires time and frequency synchronization with a cell and detects the Cell ID, based on the Primary and Secondary Synchronization Signals (PSS and SSS, respectively) transmitted in the downlink by the eNodeB.
2. **Cell acquisition and system information** [1, 6, 7, 11]. The UE acquires cell system information by means of a Master Information Block (MIB) and System Information Blocks (SIBs) of multiple types, transmitted with regular periodicities by the eNodeB. In our implementation, the number of SIB types that a UE must receive during connection establishment is sampled from a uniform, integer distribution that is defined as a model parameter.
3. **Random access procedure** [1, 6, 7, 10]. The UE requests and achieves connection setup. The simulation model implements the contention-based form of the random access procedure because we wanted to simulate a UE initiating connection establishment as it comes to the E-UTRAN (Evolved Universal Terrestrial Radio

Access Network; *i.e.*, the LTE network) from the outside. In the contention-free form of the procedure, it is the E-UTRAN that initiates connection establishment.

4. **Initial security activation and radio bearer establishment** [1, 6, 11]. Activates integrity protection and ciphering, and establishes the radio bearer(s). Our model simplifies this somewhat by implementing only the first and last message exchange in what is actually a two-step process. Each step has two handshakes and the messages exchanged in the first step may be interleaved with those of the second.

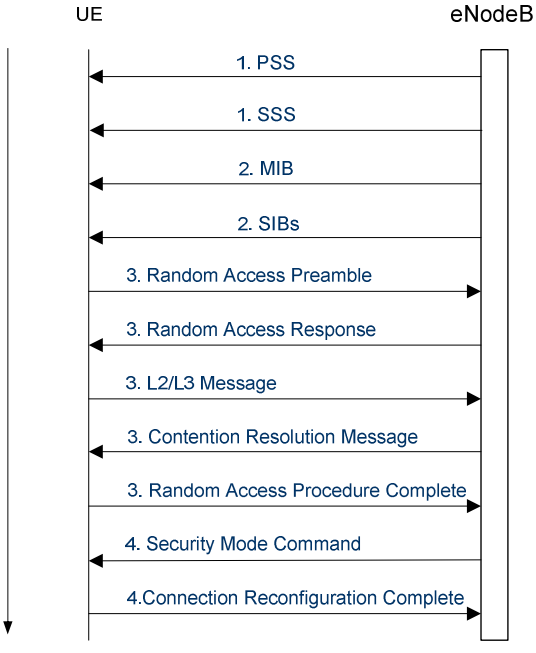


Fig. 3. The four phases of connection establishment

2.3 Some Salient Details of the Model Implementation

As noted in §2.1, our INET Framework implementation for the LTE-Uu protocol stack includes a physical sublayer and a radio sublayer. It is important to note, however, that these sublayers are, at present, basically just stubs. Our implementation models the LTE L1 (Physical) layer in a highly abstract, reductive manner that is nevertheless flexible. The wireless channel is modelled by three distributions which define, respectively: the channel transmission rate; the Block Error Rate (BER) for TBs; and the signalling error rate in the reverse direction for the ACKs/NACKs of the MAC sublayer HARQ. All three distributions are implemented as model parameters and may thus easily be changed from run to run; each may also be dynamically resampled during a simulation run. Our radio sublayer is based on the *AbstractRadio* module that exists in the INET Framework (where it is used to implement the IEEE 802.11 radio link). We have stripped this down so that it just simply transmits available TBs over the air link at 1-msec TTIs, with appropriate propagation delays

that are a function of the relative positions of UE and eNodeB. The fate of these transmitted TBs is determined in our physical sublayer module on the receiving side, in accordance with the values sampled from the BER distribution. Segmentation and concatenation in the RLC sublayer, on the other hand, is determined by the values sampled from the distribution that defines the current channel transmission rate. Future, substantive OMNeT++ simulation models for the LTE L1 layer may, of course, easily be incorporated into our framework. We note that such simulation models already exist, but not for OMNeT++: [12], for example, written in MATLAB and publicly available, is one such. We have not investigated the possibility of combining such simulators with the OMNeT++ framework; though possible in principle, this would probably be a far from trivial task.

Implementation of the HARQ mechanism[1] in the MAC sublayer, and of the interaction between HARQ and the ARQ of the RLC sublayer in AM, is based on the simulation model developed by Chuan & Lin [13]. Appropriate modification and extension to this model were made to take account of the different framework structure for our implementation; of the enhanced functionality we introduced in the form of RLC segmentation and concatenation; and of RLC UM operations.

The following were implemented as parameters of the model, allowing flexible configuration for simulation runs (default values for these parameters are given in parentheses):

- number of parallel HARQ processes (8);
- maximum number of retransmissions per HARQ process (5);
- RLC operates in UM or in AM (AM);
- RLC *t-PollRetransmit* timer [9] (5 msec.);
- RLC *t-Reordering* timer [9] (5 msec.);
- RLC *t-StatusProhibit* timer [9] (5 msec.);
- RLC reception buffer (512 bytes).

Further details on this and other aspects of the simulation model design and implementation are made available in [14].

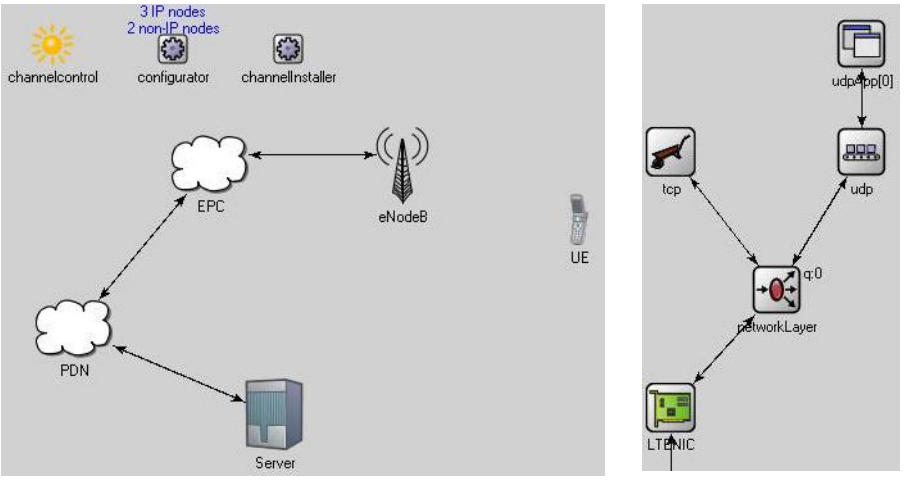
3 Results

To demonstrate the model's capabilities, we present a sample of performance results on various aspects of the RLC sublayer's UM and AM operating modes, and the MAC sublayer's HARQ mechanism. We note that complementary results on related aspects of the MAC HARQ and/or RLC AM have been published elsewhere – *e.g.*, [13, 15], amongst others.

We simulated the network shown in Fig. 4(a), in which a UE downloads a UDP stream from the node *Server*. Node *EPC* represents the Evolved Packet Core, and *PDN* the Public Data Network (in actuality, these are INET Framework switch and router nodes, respectively). Delays on the links *eNodeB* ↔ *EPC*, *EPC* ↔ *PDN* and

[1] The highly simplified, abstract model for the wireless link in our implementation makes the distinction between synchronous/asynchronous HARQ schemes less significant than would be the case in a real system. It also renders the difference between adaptive/non-adaptive HARQ schemes essentially meaningless. The HARQ mechanism we implemented comes closest in flavour to an asynchronous, 'adaptive' scheme.

PDN ↔ *Server* can be set to capture tunnelling latency through the EPC between *eNodeB* and the PDN GW, and latency in the PDN itself. Fig. 4(b) shows the internal structure of node *UE*, which is composed of the typical TCP/IP protocol stack, including a UDP video-streaming client application-layer process. *LTENIC* on the lower left of Fig 4(b) is a compound module, composed of the L2/L1 protocol stack shown in Fig. 2(b). Fig. 2(b) actually shows the equivalent compound module located in *eNodeB*, but the one in *UE* is exactly the same.



(a) (b)

Fig. 4. (a) The network simulated in INET Framework. (b) Internal structure of the node *UE* of Fig. 4(a). The compound module *LTENIC* on the lower left is composed of the L2/L1 protocol stack, which is structurally the same as that shown for *eNodeB* in Fig. 2(b).

For the results reported, the UDP download was a constant bit rate (CBR) stream composed of 10,487 UDP datagrams with a payload of 1,000 bytes each, for a total of 10MB of data. The application process at node *Server* emitted the traffic at a rate of one datagram every 0.1 msec. Note that this implies that the stream is injected into the network at a rate ≥ 80Mbps. The bandwidth of the links on the path between nodes *Server* and *eNodeB* was set to 100Mbps.

In the LTE configuration, and in order to enable straightforward comparison between results for the RLC sublayer's UM and AM[2] mechanisms, RLC segmentation in *eNodeB* was disabled so that each IP packet (PDCP SDU) corresponded to exactly one MAC sublayer TB. Also, the RLC *t-StatusProhibit* timer was disabled in *eNodeB*. The *eNodeB* → *UE* downlink channel transmission rate was fixed such that a 1-msec TTI could accommodate exactly one TB (implying a transmission rate of a little over 8 Mbps). The signalling error rate on the *UE* → *eNodeB* uplink channel was set to 2%. Finally, we ensured that sufficient buffer space was provided in the PDCP

[2] UDP traffic would, of course, normally be handled with the RLC operating in UM not AM, but this does not effect the substance of the results presented.

sublayer at *eNodeB*, with no discard timer in effect [1], so that none of the arriving packet stream is lost prior to transmission over the downlink channel.

Tables 1 and 2 below present performance results for some aspects of the RLC sublayer UM and AM mechanisms, respectively (based on averages from five independent runs of the simulation):

- Columns (a) of the tables give the Block Error Rate (BER), *i.e.*, the loss rate for TBs on the *eNodeB* → *UE* downlink channel.
- Columns (b) give the number of successful IP packets delivered by the RLC sublayer to the PDCP and higher layers at the receiving node *UE*. This is out of a total of 10,487 packets received off the network by *eNodeB* for onward transmission to node *UE*.
- Columns (c) give the total time to download the 10MB stream, in msec.

Table 1. Performance of the RLC sublayer UM mechanism in node *UE*
(with *t-Reordering* timer = 15 msec, *t-PollRetransmit* timer = 15 msec ;
and 8 parallel MAC sublayer HARQ processes, each with maximum retransmission limit = 3)

(a)	(b)	(c)	(d)	(e)	(f)
BER	Number of PDUs (IP packets) successfully received by upper layers in node *UE* (out of 10487 sent)	Total download time (msec)	Number of PDUs dropped from RLC reception buffer	Number of PDUs needing HARQ retransmission (out of 10487)	Number of TBs suffering residual HARQ errors
5%	10461	11404	25	539	1
10%	10385	12008	91	1061	11
25%	9890	14145	430	2645	167
50%	8232	18711	935	5270	1320

- Columns (d) differ between the two tables. For UM in Table 1, it gives the number of PDUs dropped from the RLC sublayer's reception buffer at node *UE*. These are PDUs that were delayed due to retransmission by the MAC sublayer's HARQ mechanism, and subsequently rejected by the reception buffer due to the way the lower boundary of this "sliding window" buffer is updated and the effects of the *t-Reordering* timer [9].
  There was no occurrence of reception buffer drops in AM. This was presumably due to the fact that out-of-order PDUs arriving at the RLC sublayer of *UE* caused its AM ARQ mechanism to trigger off requests for retransmission of the missing PDUs, and these happened to arrive in time before the lower end of the reception buffer was updated past them. So instead, we chose to report on the number of PDUs retransmitted by the AM ARQ mechanism in column (d) of Table 2.
- Columns (e) give the number of IP packets that required HARQ retransmission at the MAC sublayer. (Recall that in our case, each packet forms a single RLC SDU, which in turn forms a single MAC sublayer TB.)

- Columns (f) give the number of packets from columns (e) that experienced residual errors even after the maximum number of 3 retransmissions by the MAC sublayer HARQ mechanism, and hence could not be salvaged.

Table 2. Performance of the RLC sublayer AM mechanism in node *UE*
(with *t-Reordering* timer = 15 msec, *t-PollRetransmit* timer = 15 msec ;
and 8 parallel MAC sublayer HARQ processes, each with maximum retransmission limit = 3)

(a)	(b)	(c)	(d)	(e)	(f)
BER	Number of PDUs (IP packets) successfully received by upper layers in node *UE* (out of 10487 sent)	Total download time (msec)	Number of PDUs retransmitted by RLC AM ARQ	Number of PDUs needing HARQ retransmission (out of 10487 + column (d))	Number of TBs suffering residual HARQ errors
5%	10487	11719	299	555	1
10%	10487	12751	1026	1158	12
25%	10487	19929	4309	3736	228
50%	10487	40044	11277	11363	2829

Figs. 5 and 6 below plot some of the values in Tables 1 and 2 in order to highlight and compare some of the performance differences between UM and AM. Fig. 5 plots the data in columns (b) and Fig. 6 the data in columns (c). For both figures we calculated 99.9% level confidence intervals (in order to make ample allowance for the Bonferroni Inequality). All the confidence intervals turned out to be too small to show in the scale of the figures, except for the AM curve in Fig. 6.

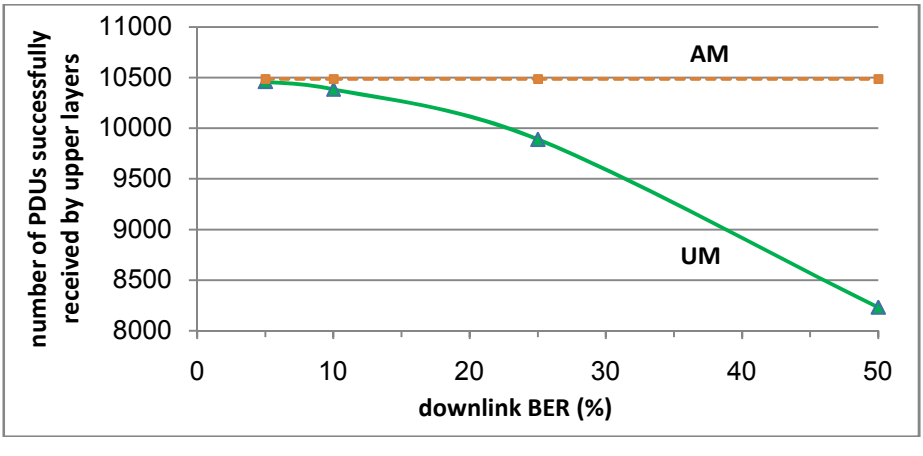


Fig. 5. Number of IP packets successfully received by higher layers at the receiving node *UE*, out of the 10,487 in the 10MB data stream (columns (b) of Tables 1 & 2)

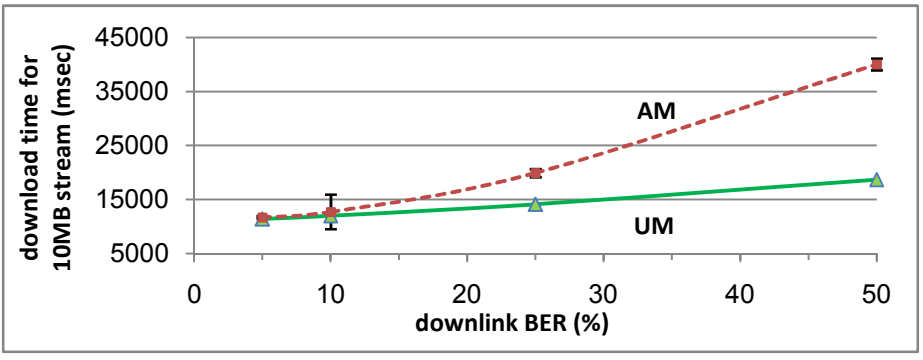


Fig. 6. Total time to download the 10MB data stream (columns (c) of Tables 1 & 2)

Fig. 5 serves to highlight how UM performance degrades in the presence of increasing BER. But it also indirectly demonstrates the effectiveness of the MAC sublayer HARQ mechanism. For example, at a somewhat extreme BER of 50%, an average of 8,232 packets out of 10,487 in the data stream (*i.e.*, 78.5%) arrive successfully. One would expect that, on average, only 50% should do so. The difference is accounted for by the MAC HARQ, which salvages only just short of 25% of the number of packets in the stream.

Fig. 5 also shows that the RLC's AM ARQ is effective in ensuring that all packets are successfully received. But as can be seen from Fig. 6, this comes at a potentially significant cost in download time as the BER increases. Note that with a 1-msec TTI during which one TB constituting one packet may be transmitted from *eNodeB* to *UE*, the theoretical lower bound for the data stream download time is of the order of 10,500 msec. Our simulations can probably do with tighter calibration of the protocol processing delays in the L2 layer stack, so it is probably advisable to approach absolute delay results from simulations with a little caution. But we can use these values on a relative, comparative basis with more confidence. From Fig. 6 we can see that the download time in UM increases by two-thirds (from an average of 11,404 msec. to 18,711 msec.) as the BER increases from 5% to 50%, due to increasing MAC HARQ recovery activity. Now, comparing UM with AM, it can be seen that the RLC AM ARQ mechanism imposes a further, ever-increasing overhead, causing the download time to more than double at a BER of 50% (18,711 msec for UM *vs*. 40044 msec for AM). One further point worth noting in this context is that, returning to Table 2 and comparing column (d) with column (f), it is clear that the amount of retransmission by the RLC AM ARQ mechanism is overwhelmingly disproportionate to the number of packets that the MAC HARQ was unable to salvage.

We now turn our attention to the effects of the RLC sublayer's *t-Reordering* and *t-PollRetranmsit timers*. Results (based on single runs of the simulation) are presented in Table 3 below:

- Column (a) gives various values for the *t-Reordering* timer, in msec.
- Column (c) gives the number of PDUs dropped by the RLC sublayer's reception buffer at node *UE* (*cf.* Table 1, column (d)).

- Column (d) gives the number of PDUs retransmitted by the AM ARQ mechanism (*cf.* Table 2, column(d)). This column does not apply to RLC UM.
- Column (e) give the total time to download the 10MB stream, in msec (*cf.* Tables 1 and 2, columns(c)).
- The *t-PollRetransmit* timer is only used in AM. It is used by the sender's RLC ARQ mechanism to control the solicitation of status reports from the receiver.
  Each table entry for AM has two values for a given setting of *t-Reordering*: one for *t-PollRetransmit* = 10 msec and the other for *t-PollRetransmit* = 20 msec. Thus, each AM entry in columns (c) - (e) is split into two subcells, one for each of *t-PollRetransmit* = 10 and 20 msec, respectively.
  As an example of how to read the table, consider when the *t-Reordering* timer is set at 5 msec (column (a)). In UM, the download time (column (e)) is 12,046 msec. In AM, on the other hand, and with *t-PollRetransmit* = 10 msec (column (e), left subcell), the download time is 17,549 msec; with *t-PollRetransmit* = 20 msec (column (e), right subcell), it is 17,847 msec.

As before, it is probably easier to absorb the data in the table through plots, which we present in Figs. 7, 8 and 9 below.

Table 3. Effect of RLC sublayer *t-Reordering* and *t-PollRetransmit* timers (with 8 parallel MAC sublayer HARQ processes, each with maximum retransmission limit = 3 ; and BER = 10%)

(a)	(b)	(c)		(d)		(e)	
t--Reordering **(msec)**		**Number of PDUs dropped from RLC reception buffer**		**Number of PDUs retransmitted by RLC AM ARQ**		**Total download time (msec)**	
	UM	**1065**		**Not applicable**		**12046**	
5	***t-PollRetransmit*** **(msec)**	**10**	**20**	**10**	**20**	**10**	**20**
	AM	**0**	**0**	**5280**	**5007**	**17549**	**17847**
	UM	**100**		**Not applicable**		**12046**	
10	***t-PollRetransmit*** **(msec)**	**10**	**20**	**10**	**20**	**10**	**20**
	AM	**0**	**0**	**1058**	**944**	**13209**	**13086**
	UM	**0**		**Not applicable**		**12046**	
20	***t-Poll-Retransmit*** **(msec)**	**10**	**20**	**10**	**20**	**10**	**20**
	AM	**0**	**0**	**57**	**69**	**12118**	**12132**

Fig. 7 below highlights the negative impact that too small a value for *t-Reordering* has on the number of PDUs dropped from the RLC sublayer's reception buffer in UM. Each PDU not dropped from the buffer translates, of course, to one more PDU successfully delivered to the upper layers at the receiving node. Note from Fig. 9 below that increasing the value of *t-Reordering* in UM has no effect on the download time of the data stream. In AM, no PDUs are dropped from the buffer, as was has already been discussed in the context of the Tables 1 & 2 results.

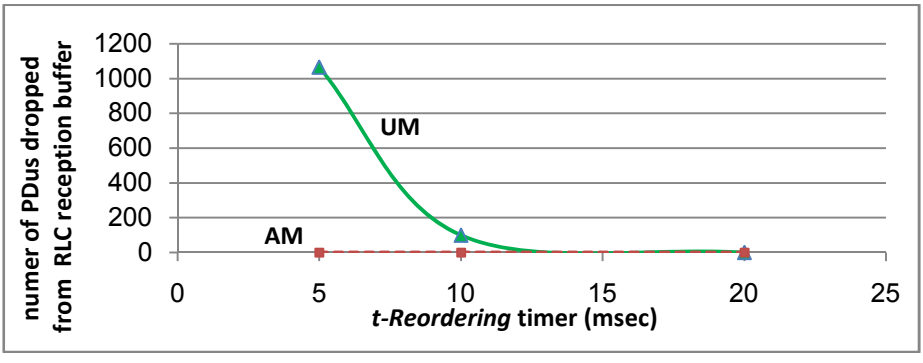


Fig. 7. Number of PDUs dropped from the RLC reception buffer (column (c) of Table 3)

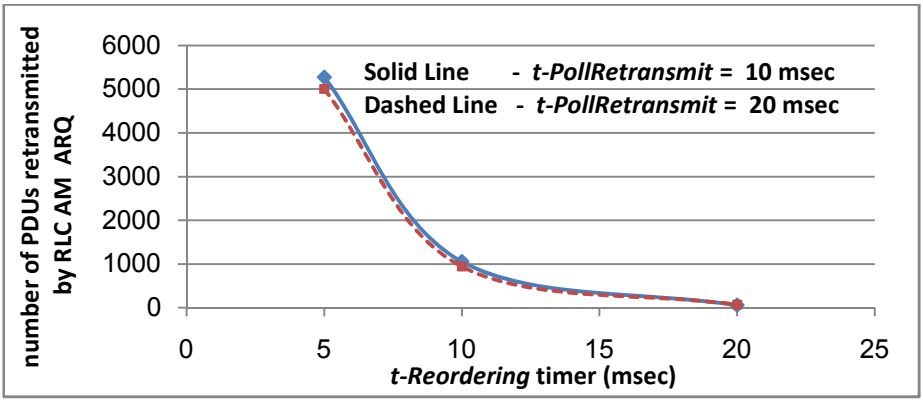


Fig. 8. Number of PDUs retransmitted by the RLC sublayer AM ARQ (column (d) of Table 3)

Fig. 8 above shows the number of PDUs retransmitted by the RLC ARQ mechanism in AM. While the *t-PollRetransmit* timer values we used had no significant effect on performance, the figure again highlights – as does Fig. 7 for UM – the critical need to configure adequate values for *t-Reordering* in order to achieve protocol efficiency. We have already previously noted that an overwhelming proportion of the AM ARQ retransmissions seem to be unnecessary, and this is

further confirmed by the decrease in the download time for AM, as seen in Fig. 9, that goes hand-in-hand with the Fig.8 decrease in the number of ARQ-retransmitted PDUs (AM will always, of course, successfully deliver the entire 10,487 packets of the data stream in whatever time – be it short or be it long – the download takes).

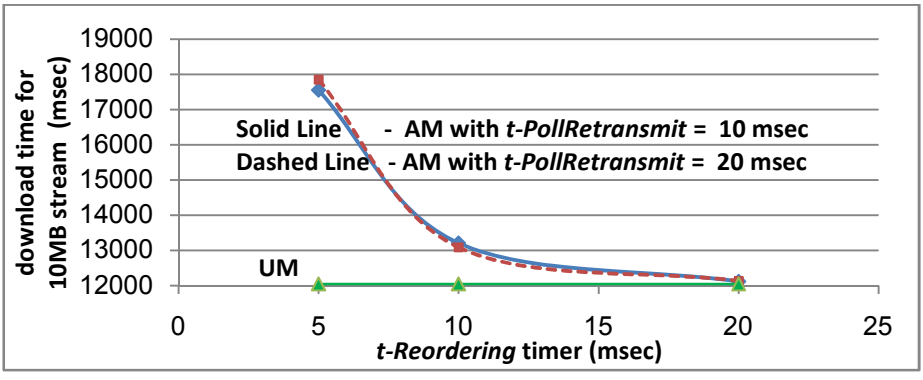


Fig. 9. Total time to download the 10MB data stream (column (e) of Table 3)

Finally, we present some results on the performance of the MAC sublayer's HARQ mechanism in Tables 4 and 5 below (based on averages from three independent runs of the simulation). Table 4 takes a look at the effect of the maximum number of HARQ retransmissions per HARQ process for a fixed number of 8 parallel such processes and a BER of 10%. Table 5, on the other hand, fixes that maximum number of retransmissions at 3, but varies the number of parallel HARQ processes. Constraints of space do not permit us to do more than simply offer these results for the reader's consideration without further ado.

Table 4. Effect of the MAC sublayer maximum retransmission limit for the HARQ processes (with 8 parallel HARQ processes ; BER = 10% ; and RLC operating in UM with *t-Reordering* timer = 15 msec)

(a)	(b)	(c)	(d)
Maximum retransmission limit per HARQ process	Number of TBs needing HARQ retransmission (out of 10487 sent)	Total number of TB transmissions by the HARQ processes (also counting multiple retransmissions for the same TB)	Number of TBs suffering residual HARQ errors
3	1070	1185	8
5	1070	1193	0
7	1070	1193	0
10	1070	1193	0

Table 5. Effect of the number of MAC sublayer parallel HARQ processes (with maximum retransmission limit = 3 per process ; BER = 10% ; and RLC operating in UM with *t-Reordering* timer = 15 msec)

(a)	(b)	(c)	(d)
Number of parallel HARQ processes	**Number of TBs needing HARQ retransmission (out of 10487 sent)**	**Total number of TB transmissions by the HARQ processes (also counting multiple retransmissions for the same TB)**	**Number of TBs suffering residual HARQ errors**
1	**1070**	**1186**	**11**
3	**1056**	**1183**	**13**
8	**1073**	**1188**	**9**

4 Conclusion

In this paper we have described the development of a 3GPP-specification-compliant simulation model, with a fine granularity of detail, for the operations of the L2 layer protocols of the LTE-Uu interface. The model is implemented in the INET Framework of OMNeT++4.0. Performance results from the model were presented. The model is intended to provide the research community with an extensible, open source simulation development platform for research in LTE-SAE.

Acknowledgments

We gratefully acknowledge the help extended by Ching-Hsiang Chuan and Professor Phone Lin of the Department of Computer Science & Information Engineering, National Taiwan University, in making available to us the details of their simulation model for the HARQ-ARQ Interaction for LTE [13].

References

1. Sesia, S., Toufik, I., Baker, M. (eds.): LTE – The UMTS Long Term Evolution: From Theory to Practice. John Wiley & Sons, Chichester (2009)
2. OMNeT++, http://www.omnetpp.org
3. INET Framework for OMNeT++4.0, http://inet.omnetpp.org
4. Qiu, Q., Chen, J., Ping, L., Zhang, Q., Pan, X.: LTE/SAE Model and its Implementation in NS 2. In: 2009 Fifth International Conference on Mobile Ad-Hoc and Sensor Networks, pp. 299–303. IEEE, Los Alamitos (2009), doi:10.1109/MSN.2009.58
5. Ns-2 Wiki, http://nsnam.isi.edu/nsnam
6. Dahlman, E., Parkvall, S., Sköld, J., Beming, P.: 3G Evolution – HSPA and LTE for Mobile Broadband. Academic Press, Elsevier, Oxford (2008)
7. 3rd Generation Partnership Project, 3GPP TS 36.300: E-UTRA and E-UTRAN; Overall Description; Stage 2, Release 9 (2010), http://www.3gpp.org/ftp/Specs/html-info/36300.htm

8. 3rd Generation Partnership Project, 3GPP TS 36.323: E-UTRA; PDCP specification, Release 9 (2010), http://www.3gpp.org/ftp/Specs/html-info/36323.htm
9. 3rd Generation Partnership Project, 3GPP TS 36.322: E-UTRA; RLC protocol specification, Release 9 (2010), http://www.3gpp.org/ftp/Specs/html-info/36322.htm
10. 3rd Generation Partnership Project, 3GPP TS 36.321: E-UTRA; MAC protocol specification, Release 9 (2010), http://www.3gpp.org/ftp/Specs/html-info/36321.htm
11. 3rd Generation Partnership Project, 3GPP TS 36.331: E-UTRA; RRC; Protocol specification, Release 9 (2010), http://www.3gpp.org/ftp/Specs/html-info/36331.htm
12. Mehlführer, C., Wrulich, M., Ikuno, J.C., Bosanska, D., Rupp, M.: Simulating the Long Term Evolution Physical Layer. In: 17th European Signal Processing Conference (EUSIPCO 2009), Glasgow, pp. 1471–1478 (2009), http://www.eurasip.org/Proceedings/Eusipco/Eusipco2009/contents/papers/1569184698.pdf
13. Chuang, C.-H., Lin, P.: Performance Study for HARQ-ARQ Interaction of LTE. J. Wirel. Commun. Mob. Comput. (2009), doi:10.1002/wcm.834
14. Public release of our code and documentation: url will be made available in time for the conference. In the meantime, this material is available on demand from the authors
15. Ikuno, J.C., Wrulich, M., Rupp, M.: Performance and Modeling of LTE H-ARQ. In: Proceedings of WSA 2009, Berlin, pp. 130–135 (2009), http://www.eurasip.org/Proceedings/Ext/WSA2009/WSA2009proceedings.pdf

Scenarios, Research Issues, and Architecture for Ubiquitous Sensing

Theo Kanter, Victor Kardeby, Stefan Forsström, and Jamie Walters

Mid Sweden University, Sundsvall 851 70, Sweden
{theo.kanter,victor.kardeby,stefan.forsstrom,jamie.walters}@miun.se

Abstract. This paper describes research issues and work-in-progress concerning ubiquitous sensing. We present scenarios where the current approaches are deficient in addressing the needs for ubiquitous sensing in services and applications on the Future Internet, involving the massive sharing of information from sensors via heterogeneous networks. We propose an information-centric architecture for real-time ubiquitous sensing which capitalizes on the proposed locator/identifier split, thus extending the Network of Information (NetInf) approach. From this we identify the challenges for which we present work-in-progress within the framework of the EU-funded MediaSense project. Firstly, we integrate sensors as addressable objects, exposed by means of sensor gateways and relocatable abstract interfaces. Sensor information is thus made available to applications solely based on identity. Secondly, sensor information is made available in a distributed data model towards searching and browsing. Finally, we evaluate the effectiveness of the architecture in proof-of-concept applications for intelligent commuting, environmental monitoring and seamless media transfer, utilizing two different sensor platforms.

Keywords: Ubiquitous computing, sensors, Future Internet.

1 Introduction

The ability for applications and services to have access to sensor information and be able to act upon it via actuators is becoming increasingly more important; even urgent. This is particularly the case where it concerns our ability to manage energy and to bring about a sustainable environment. As a result of communications becoming more pervasive, in urban areas and rural areas alike, we have become citizens in an electronic world however still needing to stay in touch with artifacts, people and places in the real world. With sensors and actuators massively connected to the Future Internet environment, accounting for the majority of connected nodes, increasing in numbers by orders in magnitude, impacts on the required mechanisms and architecture. Further, it increases the amount of information that should be searchable and accessible to services and applications via the Internet. Thus reachable via heterogeneous networks, involving wireless and mobile communication, networking, and information brokering.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 285–297, 2011.

Several initiatives and projects have addressed this area and recognized the need for an information-centric approach in contrast to earlier network-centric approaches. In particular, the EU-funded 4WARD-project proposed a Network of Information (NetInf) approach in which end-devices are integrated as sources and consumers of network information, thus enabling end-devices as parts of a whole [1]. The NetInf approach, capitalizing on a locator/identifier split, is proposed to enable content-centric networking (CCN) [2], which retrieves content by name instead of network location. The NetInf approach envisages multimedia associated with real-world information from entities (e.g., people, places, artifacts, etc.). Extension though integration of real-world information is a possibility, which is left unexplained.

Scenarios for a sustainable planet involving ubiquitous sensing regarding the environment, transport and social mobile media, which will be further discussed below, constitute requirements on real-world integration. Firstly, sensors and actuators should be integrated as addressable objects. Sensor information is thus associated with entities via interfaces in end-devices and made available to applications in an extended network of information with ubiquitous sensing. Secondly, sensor information must be incorporated and made available in an extendible, searchable and browseable (distributed) datamodel. Thirdly, a pre-condition is that the provision of sensor information is scalable to accommodate the billions of nodes that will populate the Future Internet[1]. Fourthly, sensor information should be available in real-time (i.e., within predictable and reasonable finite amounts of time, relative to the application domain). Finally, as a consequence of extending the NetInf approach, sensors must be reachable by identities and be associated with entities.

Previous efforts have focused on the brokering of sensor information via web service infrastructures and 3G mobile systems, such Mobilife [3], or via clients connected with web services to servers on the Internet [4,5,6,7]. Brokering sensor information centralized via IMS in 3G Mobile Systems is not suitable for our purposes. Web technologies using DNS are not a suitable architecture.

Other related work [8] has focused on network aspects, and the granularity of the actual exchanged context information. While in other cases, on entire systems and exploring what is possible within the constraints of current research[9,10].

The EU-funded project SENSEI [11] proposes a logical architecture which is compatible with the NetInf approach but has as yet not provided answers about how sensors are integrated, how such information be available and searchable in real-time or in a scalable fashion, alluding in its architecture to centralized mechanisms such as LDAP.

In order to address these deficiencies we propose an architecture and middleware for the scalable integration of actuators and sensors in a network of information for ubiquitous sensing.

The remainder of this paper is organized as follows: Section 2 discusses the urgent need for ubiquitous sensing in three key areas from which we derive our

[1] http://gigaom.com/2010/04/14/ericsson-sees-the-internet-of-things-by-2020/

requirements for an extension of a network of information. Section 3 discusses the principle operation of the architecture and its components for the scalable integration of sensors and actuators in a browseable network of real-time sensing information available to applications and services via heterogeneous networks. Section 4 examines the effectiveness of the proposed architecture and middleware components applied to the key scenarios in relation to the identified goals. Finally, in section 5 we summarize our findings and section 6 concerning further work beyond the work-in-progress discussed in this paper.

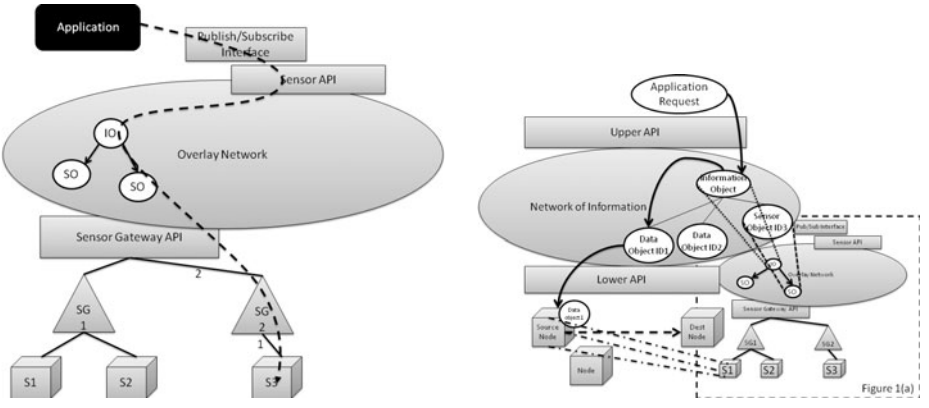
2 Scenarios

In this section, we discuss the urgent need for ubiquitous sensing along with its consequences. The requirements summarized at the end of the introduction above are pivotal to ubiquitous sensing. These general requirements are further elaborated below through an examination of certain key scenarios.

There are huge challenges ahead in the management our environment on a global level. Our knowledge that our way of living is not sustainable has improved due to the proliferation of satellite and other climate monitoring systems. Equally has our understanding that course-grained measures as restrictions for industry or private consumption are very blunt tools. The situation requires more fine-grained tools in interacting with the real world in terms of applications and services that have access to sensor information and actuators on a massive scale. Such tools would enable authorities, industry and consumers to help minimize waste and the misuse of our planet's resources such as energy. This however requires real-world information to be available as distributed data that is searchable and browseable.

Transport infrastructure has to be more responsive to our needs. Thus, the status and actions of the parts and components in a transport system must be monitored and controlled on a much more fine-grained level, with the information being accessible to applications and services that are available to users. The need and even necessity for applications and services to interact with transport infrastructures via communicating sensors and actuators applies to both public transportation (e.g., bus, train, air, etc.) and private transportation (e.g., cars, roads, etc.). When entities participate in such a system, applications would benefit greatly from having timely access to real-world information provided by other entities.

Further, on a global level, individuals have become citizens in an on-line world with an ever-increasing range applications and services. Our communications infrastructure is increasingly populated with novel devices and connecting to appliances in our personal environment for entertainment, utility services, etc. The number of connected devices that we as individuals or as members of families or communities alone wish to interact with, even using a cautious prediction will exceed 50 billion. Entities, sensors and actuators should be able to join spontaneously and be able move in the infrastructure. Thus, individuals will be



(a) The Mediasense Architecture

(b) NetInf framework with the Mediasense architecture

Fig. 1. Extending the NetInf architecture with the Mediasense architecture to support ubiquitous sensing

enabled to stay in touch and perform tasks using content, devices and real-world information that is accessible via heterogeneous networks and reachable based on identity.

We summarize the discussion of the consequences of ubiquitous sensing in relation to a network of information in an elaborated list of requirements:

a) applications can reach sensors using an identity independent of location *b*) sensors are attached to the network of information interfaces which abstract from the implementation of communicating with the sensors *c*) abstract sensor interfaces allow physical sensor hand-off and roaming *d*) sensor information is organized in an extendible datamodel, which allows spontaneous additions *e*) sensor information dissemination support is scalable to accommodate a massive number of nodes exceeding current predictions of 50 billion *f*) a sensor information database must be distributed *g*) a sensor information database must be searchable and browseable *h*) the architecture must support sensor information discovery *i*) a sensor information database must be updated in real-time

3 Concepts and Architecture

The NetInf architecture has identified several base concepts which are influencing the overall architecture, most prominent is the Identifier/Locator split which is a necessary part to enable the addressing of individual ubiquitous sensors (requirement a). On the future Internet when sensors are attached to not only stationary devices, such as weather stations, but also on people via their mobile phones, as well as vehicles and animals, sensors must be able to move and attach to the Internet from the different locations (req. c). Due to an increasing number of sensors, we cannot assign an individual IPv4 address to each sensor. To cope

with the increasing amount of sensors and to keep the cost of such sensors to the minimum, we introduce the concept of a Sensor Gateway represented as a triangle in figure 1(a).

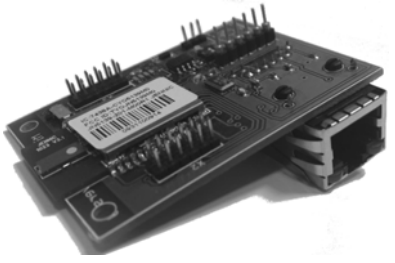
Sensor gateways constitute a point of network attachment for sensors in our architecture. Thus sensor gateways communicate with sensors (see boxes in fig. 1(a)) and enable sensor access through a Sensor Gateway API to and from even the simplest sensors (req. b). All sensors and sensor gateways are given a globally unique identity which is stored in a distributed overlay (req. f) and are divided into two types of objects akin to objects in NetInf. The first, a sensor object (SO in fig. 1(a)), contains the identity of the sensor and the location of the sensor gateway that is currently hosting the sensor. The second object is an information object (IOin fig. 1(a)) which similarly to NetInf contains semantically meaningful information. The distributed overlay provides an upper API (see fig. 1(a)) which mediates access to sensors without needing to know the location of an object. The distributed overlay is capable of exchanging context information in real-time by utilizing a distributed hash table (DHT) based on Chord [12] (req. i).

A publish/subscribe API and a socket-like interface are located on top of the overlay API. The application may use this to combine sensor information into new information that is logically stored as a new sensor inside the overlay. For instance, an application could connect to, retrieve all temperature sensors from a municipality, and then provide a new sensor that represents the average temperature of the whole municipality. This publish/subscribe API allows applications to access the raw sensor information, which is illustrated using a dashed line in figure 1(a). In addition to this, the socket interface can ubiquitously find the information, regardless of location, which enables mobile context-aware applications to utilize the architecture. In contrast to previous publish/subscribe solutions, our architecture and API's will be extended to support searching and dissemination in real-time.

3.1 Extending NetInf for Ubiquitous Sensing

By extending the NetInf architecture with the Mediasense architecture we introduce a new kind of data object, a sensor object. For example, an information object that represents a certain song (e.g., Mozart's Eine Kleine Nachtmusik) is associated with a data object. The data object contains a payload (e.g., an mp3 file with a certain encoding). In this example, a sensor object's payload contains the current state or value that is retrieved in realtime (e.g., of the user's proximity, temperature, light, mood, etc.). The information object that represents the sensor contains the sensor's semantical information since sensor values need context to be usable for applications (e.g. it's geographical location, technical specification or other important information), see further [13]. Both sensor objects and their corresponding information objects are duplicated by their realtime updating counterparts in the mediasense architecture as the dashed lines in figure 1(b). The lower API in the NetInf architecture may thus access sensor information through the mediasense upper API in the same way as it accesses a source node. Hence, the upper API of NetInf accesses information regarding

(a) Wireless Sensor Platform



(b) Ethernet Enabled Sensor Platform

Fig. 2. Wireless Sensor Network Gateway

sensors in the same way as other objects. Figure 1(b) shows how the mediasense framework from figure 1(a) is integrated into NetInf.

Mobility in the architecture can happen, as sensor and data object move around in the world. The system is however agnostic of communication medium, and therefore most communication which can move, will utilize ubiquitous mobile Internet-access such as the 3G or 4G networks. However if a sensor or data object has moved between connectivity mediums or devices, and therefore also changed the network connection, an update is required in the architecture. Therefore, when a sensor object appear with a new connection, it will update it's corresponding information object, with current information on the physical location of the sensor object. Which will result in a successful movement of all related information between two physical locations.

3.2 Sensor Network Platform

A Wireless Sensor Networks (WSN) consist of cheap, small and energy constrained sensors interconnecting to create a network, enabling reliable and automated sensor data acquisition with minimal human intervention. These WSNs are an important source of context information and by utilizing the support proposed in section 3 the sensor information can be efficiently delivered to other entities. By extending and improving the wearable bridge described in [14] we have built a prototype sensor-actuator combination, figure 2(a). The prototype consists of two individual devices.

The first device is the large box, which is capable of measuring temperature and humidity; by connecting the box as a power strip, we can also measure the power consumption of up to two household appliances simultaneously. Apart from measuring the power consumption, the prototype provides actuators that can control the power sockets individually. This sensor is aimed towards home

usage, and utilizes the Zigbee radio protocol instead of bluetooth to communicate with the second part. The line in figure 1(a) denoted with number 1 indicates communication using wired or wireless sensor-specific hardware and protocols.

The second device is shown in the embedded picture in figure 2(a) and is the hardware part of a sensor gateway. Several sensors can communicate with this gateway and several gateways can communicate with an entry point into the mediasense overlay in figure 1(a) denoted by the number 2. This protocol is based on HTTP and uses GET and POST messages for communication with a subset of response messages.

Important to note is that when a new sensor registers with a sensor gateway and the sensor's type is not recognized, then the sensor gateway will retrieve a module from the overlay, which enables communication with the new sensor without user intervention. This enables sensors to be truly mobile by only requiring a nearby sensor gateway, which possesses the required hardware protocol and a basic sensor identification protocol. The sensor gateway detects a node that has the sensor gateway API during the initialization phase using IP broadcasts. This API may be located on the local host or the local network.

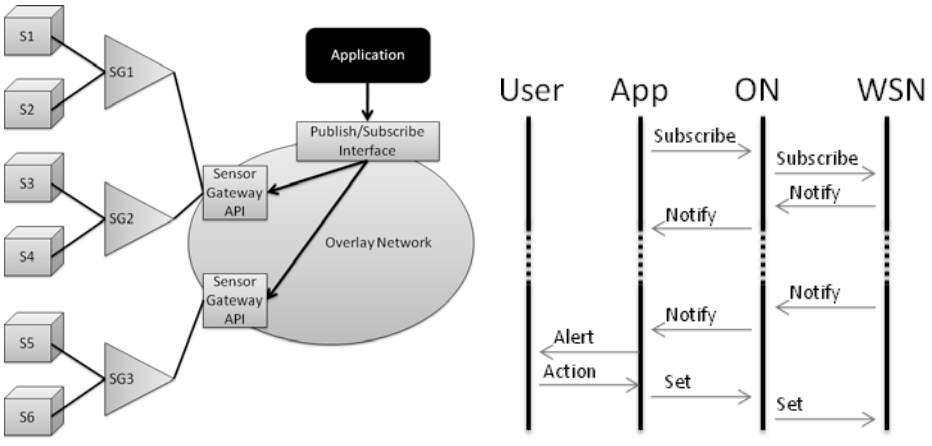
An extension to this sensor network platform has been developed for use in areas where the use of a computer as a gateway between the Zigbee network and IP network is not possible. This sensor platform employs simpler radio protocol but has an Ethernet interface, see figure 2(b), and offers both sensor information as IP-addressable web pages and sensor information in the mediasense overlay. Currently, only environmental monitoring(temperature, humidity) is enabled with this platform but more sensors and eventually actuators will further complement the platform later.

4 Functional Verification

This section evaluates the effectiveness of the architecture as it is applied to three different scenarios that are presented as three proof-of-concept prototypes. In different places we reference back to the requirements list presented in section 2.

4.1 Environmental Awareness

Due to advancements in technology, sensor solutions are becoming inexpensive; this enables more pervasive use of sensors in environmental awareness applications. One scenario under current investigation is the use of sensors to monitor the environment in cross-connection rooms and other broadband installations of ISPs. There exists an urgent need to monitor the local environment of such infrastructure with regard to temperature, humidity and power consumption, etc. ISPs need to react pro-actively to changes in the environment that currently remain undetected. By having sensors able to detect events, from temperature and humidity to events such as flooding or fire, a more rapid and correct response can be engaged.



(a) Architecture Overview (b) Signalling Overview

Fig. 3. Environmental Sensing Architecture

Thus, sensors must be able to provide environmental information in real-time (req. i) and be accessible to an ISP's monitoring center (req. g). The information must also be available to maintenance staff who are mobile (req. a). The architecture should scale and use a low amount of bandwidth to allow for extending the ISP's infrastructure globally (req. e).

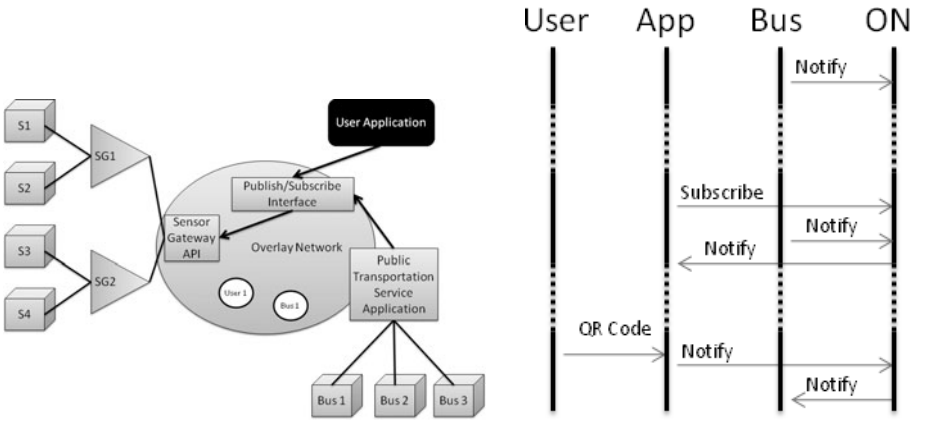
Our architecture supports this scenario by enabling sensors to be accessed globally with the aid of a sensor gateway at each location(req. a). Our architecture allows access to sensors from a remote monitoring application, be it a central facility or a mobile device.

Figure 3(a) presents the proof of concept. On the left side there are multiple sensors that may be located at several installations of the operator. The sensors are attached to local sensor gateways, which communicate with the overlay through a sensor gateway API. Management software, be it a stationary computer or a mobile terminal, is connected to the infrastructure from the other end. Compare with figure 1(a). The system scales by adding more nodes associated with sensor gateways and APIs (req. e).

Figure 3(b) shows the signaling in a scenario when a monitor application subscribes to sensors at each location. The sensor gateways join the overlay, which in turn shares information from associated sensors. Later the user is notified of some event that in response can access actuators via the overlay.

4.2 Context Aware Commuting

With the introduction of smart phones for the mainstream market, there are an increasing number of sensors available to be exploited for the benefit of the user. Mobile applications and services may benefit from knowledge about the user's context through gaining access to sensor information (e.g. temperature, GPS coordinates, nearby WLAN). Sensor information may be further combined



(a) Architecture Overview (b) Signalling Overview

Fig. 4. Commuting Support Architecture

with the user's digital information (e.g., schedule and contacts) in order to determine other relevant information such as the user's current destination. Mobile applications and services may combine such knowledge with public transportation system information in order to provide suggestions concerning travelling preferences, travel time, cost, or information concerning the destination. Such a system require both access and storage of a digital profile of a user containing personal preferences and automatically acquired sensor information as well as access to information from public transportation systems thus creating a NetInf containing real-world continuously updating data.

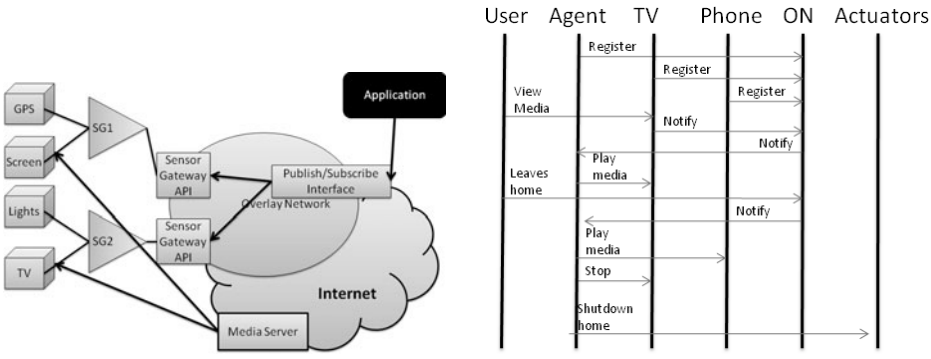
Our architecture supports this by providing both the ability of the public transportation system to input timetables and location of their vehicles as well as the user's ability to connect when desired to synchronize their personal profile with fresh data and acquire travel suggestions from the system.

The initial prototype uses QR codes[2] to detect the proximity of a bus stop, but the context service can also find the bus stop by other means, such as GPS and nearby bluetooth devices. After positioning the user at a bus stop, the service can provide information about the bus route; the information could include a timetable, the time until the next bus arrives and its route displayed on a map. Figure 4(a) shows how several busses are connected into the architecture from one side and the users' connection from the other.

The architecture enables users to retrieve information from remote buses and maintain an online profile allowing bus drivers to receive notifications when someone is waiting at a bus stop.

The signalling for this scenario is detailed in 4(b), where a bus is continuously submitting context information to the service. Later a user connects and registers in order to obtain information regarding the bus.

[2] http://www.qrcode.com/index-e.html



(a) Architecture Overview (b) Signalling Overview

Fig. 5. Seamless Media Transfer Architecture

4.3 Seamless Media Transfer

By incorporating sensors providing real-time data about users and media into a NetInf, we enable the architecture to decide where it should deliver content based on the recipient's current and future whereabouts. When leaving our residence, an ongoing media session may follow us. Seamless media support transfers the media session from the media center in our residence to a mobile device. In the process the seamless media support shuts off devices that are not required using the actuator power strip presented in section 3.2. Figure 5(b) presents the signalling used when the user starts playback of some media and then leaves the home. In figure 5(a) the first sensor gateway (SG1) is connected via Bluetooth to a mobile phone and the sensor gateway API runs in user's mobile phone. The second sensor gateway (SG2) is located in the residence. Mobile phone, TV, and media server are connected to the MediaSense overlay, which enables the user's agent to respond to sensor information, move the media session to mobile and minimize energy consumption at home.

5 Conclusions

Above we described research issues and work-in-progress concerning ubiquitous sensing. In order to derive requirements and key research issues, we presented three key scenarios where current approaches are deficient in addressing the needs for ubiquitous sensing in services and applications on the Future Internet. In particular, services and applications require the massive sharing of information from sensors via heterogeneous networks. Further, services and applications must be able to connect to sensors using an identity and not a location, adhering to the observations of the NetInf approach. Sensor information must be current (in real-time) and available in a searchable and browseable distributed data model.

In response to these challenges, we propose an information-centric architecture for real-time ubiquitous sensing which capitalizes on the proposed locator/identifier split, thus extending the Network of Information (NetInf) approach. Sensors are exposed as first-class objects in a distributed information base using relocatable abstract interfaces. Sensor information is thus made available to applications solely based on identity through sensor sockets. Further, we presented work-in-progress within the framework of the EU-funded MediaSense project, involving proof-of-concept prototypes including two sensor platforms and gateways for the sharing of sensor information in a peer-to-peer context information network via abstract interfaces. We evaluate the effectiveness of the architecture in prototypes pertaining to the three key scenarios (intelligent commuting, environmental monitoring, and seamless media transfer) in terms of the list of requirements and key research issues. We demonstrate the integration of sensors and actuators as information objects that may be reached via their identities, by means of sensor gateways and relocatable abstract interfaces.

6 Future Work

Further research in the area of ubiquitous sensing on the Future Internet should focus on extending the proposed support. In particular, the massive and seamless sharing of sensor information requires new mechanisms to enable seamless and ubiquitous sensor connectivity. We are working on extensions to the presented support to include the ability to establish and maintain context sockets, which can deliver seamless connections to heterogeneous sources. Further, the dynamic relation between presence entities and sensors mandates the search for effective extensions to presented middleware for the massive sharing of sensor information. The sharing of sensor information in real-time, search and browse, as well the discovery of sensor information require scoping mechanisms to be effective, where current approaches offer insufficient answers. Real-time properties and other properties involved in ubiquitous sensing which were discussed earlier require evaluations of not only of the architecture, as presented above, but other aspects as well. Such an evaluation may require a method to classify the constraints in real-time context aware applications. We envision that the main aspects would be responsiveness of the system, i.e. the delays, together with scalability since the system will contain the billion sensors of tomorrow. The goals for the system is to enable a perceived real-time delivery of it's content which also allows for the inclusion of millions of sensors. The tools used for the evaluation could initially comprise of simulation software, but as methods are developed subsequently entail monitoring of infrastructure, including analysis support. Further research should also entail efforts towards abstracting physical sensors into logical objects creating higher-level interaction for applications and services attempting to navigate the heterogeneous properties inherent in reasoning over ubiquitous sources.

Acknowledgments

The work was conducted under the MediaSense project and partially funded by the EU Regional Fund and the County Administrative Board of Vsternorrland. The authors acknowledge Muhammad Amir Yousaf and Prof. Bengt Oelmann for the development of sensor platforms used throughout the project, as well as Roger Norling for implementing the context-aware bus system in cooperation with Swedish Connection.

References

1. Dannewitz, C., Pentikousis, K., Rembarz, R., Renault, E., Strandberg, O., Ubillos, J.: Scenarios and Research Issues for a Network of Information. In: Proceedings of the 4th International Mobile Multimedia Communications Conference (2008), http://eudl.eu/?id=3980
2. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M.F., Briggs, N.H., Braynard, R.L.: Networking named content. In: CoNEXT 2009: Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies, pp. 1–12. ACM, New York (2009)
3. Klemettinen, M.: Enabling Technologies for Mobile Services: The MobiLife Book. John Wiley and Sons Ltd., Chichester (2007)
4. Abdelzaher, T., Anokwa, Y., Boda, P., Burke, J., Estrin, D., Guibas, L., Kansal, A., Madden, S., Reich, J.: Mobiscopes for human spaces. Pervasive Computing 6, 20–29 (2007)
5. Hull, B., Bychkovsky, V., Zhang, Y., Chen, K., Goraczko, M., Miu, A., Shih, E., Balakrishnan, H., Madden, S.: Cartel: A distributed mobile sensor computing system. In: Proceedings of the 4th International Conference on Embedded Networked Sensor Systems, pp. 125–138. ACM Press, Boulder (2006)
6. Grosky, W., Kansal, A., Nath, S., Liu, J., Zhao, F.: Senseweb: An infrastructure for shared sensing. Multimedia 14, 8–13 (2007)
7. Santanche, A., Nath, S., Liu, J., Priyantha, B., Zhao, F.: Senseweb: Browsing the physical world in real time. In: Demo Abstract, Nashville, TN (April 2006)
8. Tahayori, H., Degli Antoni, G., Pagani, E., Astaneh, S.: Context network. In: Canadian Conference on Electrical and Computer Engineering, CCECE 2007, pp. 353–356 (2007)
9. Raz, D., Juhola, A.T., Serrat-Fernandez, J., Galis, A.: Fast and Efficient Context-Aware Services. Wiley Series on Communications Networking & Distributed Systems. John Wiley & Sons, Chichester (2006)
10. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. International Journal of Ad Hoc and Ubiquitous Computing 2(4), 263–277 (2007)
11. Bauer, M.: Towards a real world internet: Context and actuation based on the sensei system. In: Fischer, S., Maehle, E., Reischuk, R. (eds.) GI. LNI, vol. 154, p. 236 (2009), http://dblp.uni-trier.de/db/conf/gi/gi2009.html#Bauer09
12. Kanter, T., Pettersson, S., Forsstrom, S., Kardeby, V., Norling, R., Walters, J., Osterberg, P.: Distributed context support for ubiquitous mobile awareness services. In: Fourth International Conference on Communications and Networking in China, ChinaCOM 2009, pp. 1–5 (August 2009)

13. Dobslaw, F., Larsson, A., Kanter, T., Walters, J.: An object-oriented model in support of context-aware mobile applications. In: Cai, Y., Magedanz, T., Li, M., Xia, J., Giannelli, C. (eds.) Mobilware 2010. LNICST, vol. 48, pp. 205–220. Springer, Heidelberg (2010)
14. Kanter, T., Pettersson, S., Forsström, S., Kardeby, V., Österberg, P.: Ubiquitous mobile awareness from sensor networks. In: Hesselman, C., Giannelli, C. (eds.) Mobilware 2009 Workshops. LNICST, vol. 12, pp. 147–150. Springer, Heidelberg (2009)

Challenges for Cloud Networking Security

Peter Schoo[1], Volker Fusenig[1], Victor Souza[2], Márcio Melo[3], Paul Murray[4], Hervé Debar[5], Houssem Medhioub[5], and Djamal Zeghlache[5]

[1] Fraunhofer Institute for Secure Information Technology SIT, Garching near Munich, Germany
{peter.schoo,volker.fusenig}@sit.fraunhofer.de
[2] Ericsson Research, Stockholm, Sweden
victor.souza@ericsson.com
[3] Portugal Telecom Inovação, Aveiro, Portugal
marcio-d-melo@ptinovacao.pt
[4] HP Labs, Bristol, United Kingdom
pmurray@hp.com
[5] Institut Telecom, Telecom SudParis, France
{herve.debar,houssem.medhioub,djamal.zeghlache}@it-sudparis.eu

Abstract. Cloud computing is widely considered as an attractive service model since the users commitments for investment and operations are minimised, and costs are in direct relation to usage and demand. However, when networking aspects for distributed clouds are considered, there is little support and the effort is often underestimated. The project SAIL is addressing *cloud networking* as the combination of management for cloud computing and vital networking capabilities between distributed cloud resources involved to improve the management of both. This position paper presents new security challenges as considered in SAIL for ensuring legitimate usage of cloud networking resources and for preventing misuse.

Keywords: Cloud Networking, Cloud Computing, Network Virtualisation, Security.

1 Introduction

Initially driven by the deployment of IT applications leveraging the economy of scale and multi-tenancy, cloud computing is today becoming the platform of choice for many different applications. The advantages of running applications in the cloud are manifold: lower costs through shared computing resources, no upfront infrastructure costs, and on-demand provisioning of computing nodes to fit transient requirements. Thus, applications that show high degree of variable demand for resources fit the cloud computing model well. Virtualisation in the data centres has been a key enabler to allow the dynamic provisioning of computing resources to become reality.

While little focus has been given to the network aspects so far, it is obvious that the perceived performance of some applications running in the cloud

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 298–313, 2011.

depends heavily on the network connecting the different cloud sites and connecting the user to the cloud. Applications with interactive and bandwidth hungry characteristics are a good example of the above. As these applications move to the cloud, more will be demanded from existing networks in terms of, e.g., capacity (likely more data to be sent across network links), quality (low delay for interactive applications), and availability.

Besides, cloud applications will demand a network that is more flexible. Since applications and entire cluster of servers can be moved to (or created in) another data centre, existing networking pipes need to be *re-plumbed*. Existing technology provides the allocation of computing resources in the cloud in a dynamic and quick fashion while network connections to those resources are more or less statically established by network operators. Networks that can swiftly be reconfigured will enable the full benefits of the cloud environment. This is the envisioned concept of cloud networking - it encompasses provisioning of on-demand guaranteed network resources in a time span that is compatible with the allocation of computing resources in a cloud today.

This paper presents the research challenges of providing a *secure* cloud network system. These research challenges will be explored in the course of a 30 months project called SAIL (Scalable Adaptive Internet soLutions) that has started in August 2010. SAIL [1] is an EU funded project (part of the 7th Framework Programme) whose consortium includes 24 partners from industry, academia, and research institutes. SAIL aims at creating technology to address some of the shortcomings of the current Internet. This includes the lack of a content-centric model for large scale content distribution, support for connectivity services providing point-to-multipoint capabilities, insufficient support for deployment of dynamic guaranteed network connections in a cloud computing scenario, and non-technical work that will evaluate, identify, and propose, among others, new business models, address socio-economic questions.

For Cloud Networking, SAIL will develop networking functions for applications with highly variable demands, integrating these functions with computing and storage, along with the necessary tools for management and security. In that way, the allocation of both computing and networking resources will be solved as only one optimisation problem. A prototype of the proposed solutions will be developed and refined under the course of the project. The prototype will be hosted on some partners premises distributed across Europe. An iterative approach to research will be taken, whereby proposed solutions will be assessed through prototyping, providing feedback to the architecture, management, and security tasks of CloNe. Just as important, the workpackage will provide a migration path whereby developed technologies will be deployed in the existing Internet and standardized.

Besides the cloud networking security related requirements and challenges, more fundamental cloud computing security aspects will be considered. Cloud computing environments are likely to suffer from a number of known vulnerabilities, enabling attackers to either obtain computing services for free (attack against cloud providers), steal information from cloud users (attack against cloud

customers data), or penetrate the infrastructure remaining in client premises through cloud connections (attack against cloud customer infrastructures). Typical examples of these attacks today are VoIP free calls, SQL injection, and drive-by downloads [2]. Cloud networking will not change the fact that vulnerabilities will continue to exist and that attackers will continue to exploit them. However, the concentration of massive amounts of computing power and data will make these targets more visible and more attractive.

This paper is organised as follows. Section 2 presents the concepts of cloud networking and cloud computing in more details. Section 3 explores the security issues when implementing the cloud networking vision. Section 4 presents closing remarks and summarises the next steps of this research project.

2 From Cloud Computing to Cloud Networking

Cloud computing has gathered a lot of attention in recent years from parties across the computing and communication industries including vendors, network operators, and service providers. The service utility business model on which cloud computing is based is far from new. In 1961, Prof. John McCarthy was one of the first to introduce it by the claim that computer time-sharing technology might lead to a future in which computing power and even specific applications could be sold through the utility business model, i.e., water or electricity [3].

The existence of the Internet and web technologies, and the introduction of infrastructure virtualisation has enabled the current realisation of that vision. Separation of the service provider from the infrastructure provider, is making it easier to generate new services on-line and to scale those services as demand dictates. For the service provider this reduces capital and operational expenditure, and financial risk, as they pay for access to resources on an as-needed basis, with little or no lead time to change capacity. For the infrastructure provider this gives the opportunity to build large infrastructures that benefit from economies of scale [4] and amortise the costs across the workload of multiple customers.

2.1 Virtualisation Technology Supporting Cloud Computing

Today's infrastructure-as-a-service (IaaS) is built on server virtualisation (virtual machine hypervisors such as Xen [5] or VMWare [6]), network virtualisation (implemented in network equipment or distributed routers such as [7]), and storage virtualisation (including network attached storage arrays or storage services such as Amazon's Elastic Block Store [8]). Data centre management systems deploy and manage virtual machines, networks, and data stores to construct any infrastructure topology required by the customer by dynamically re-configuring the virtualisation layers. These virtualisation techniques are now so common place that hardware support has been introduced to standard server chip sets by vendors such as Intel (VT-x [9]) and AMD (AMD-V [10]).

The IaaS business model drives infrastructure providers towards a centralised architecture, as depicted in Figure 1. Very large data centres located near low

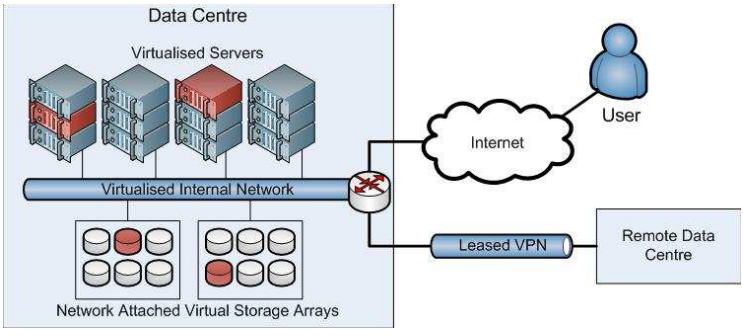


Fig. 1. Cloud Physical Infrastructure Architecture

cost power, land, and labour result in the lowest costs for the provider. However, the global nature of the business introduces opposing factors. From a regulatory perspective, the location of a data centre determines in part the legal jurisdiction that applies to hosted services (e.g., USA Patriot Act [11]). The use of the services can restrict their location or transfer of data (e.g., EU Data Protection Law [12]). From a technical perspective load, data transfer, or disaster-tolerance may require multiple geographical locations. Processing load and data transfer is typically dealt with by parallel implementations of the service, each located geographically near the users. Disaster- tolerance requires replicating services in geographically diverse sites. As a result of these driving factors, today's cloud infrastructure providers operate a few, very large data centres, located in a select number of geographical locations.

Connectivity between data centres owned by a single provider is usually implemented by leased virtual networks providing guaranteed, but static quality of service for the IaaS owner. Connectivity between the data centre and the IaaS user is generally handled by the open Internet. As such, the user's network experience is based on access to a shared medium, which is not under control of the cloud provider.

Although it is possible to scale the virtual infrastructure implemented by a IaaS provider, it is not possible to scale the connectivity to that infrastructure. Recently IaaS providers have added VPN tunnelling connectivity for their customers based on secure connection-oriented protocols such as IPsec (e.g., Amazon Virtual Private Cloud [13]). This allows, for example, the creation of an IT infrastructure in the cloud that is connected to the site network of an enterprise in a way that enables them to use their own address space and network services across both. However, they are still subject to the limitations of bandwidth, jitter, and latency offered by their Internet service provider and lack of support for dynamic provisioning.

The class of applications that are currently deployed in cloud infrastructures are those that are suited to this architecture, for example: batch processing, such as large scale simulations or graphics rendering, on-line web services, and

hosted IT systems. Where sensitivity to network performance is an issue, such as content delivery [14], it is still necessary for the service provider to own the infrastructure or to enter into a long term contractual engagement with the infrastructure provider. The network components and topology of these services are still largely static.

2.2 Virtualisation Technology Supporting Cloud Networking

Network virtualisation brings a missing piece to the cloud computing puzzle. Virtual networks are not at all new in themselves; [15] provides a survey of technologies used at various layers. A number of network virtualisation architectures and frameworks have been proposed in the literature, including VINI [16], CABO [17], 4WARD VNet [18], and FEDERICA [19], to offer customised virtual networks with end-to-end control.

The possibility to specify and instantiate networks on demand and in useful time is one of the great advantages of network virtualisation. Virtual networks can be freshly created according to the different requirements, such as bandwidth, end-to-end delay, security, and protocols. Network virtualisation brings other advantages into stage, such as the ability to reconfigure the network in real time without losing connectivity, to change the physical path, or even to move one or more virtual nodes from one place to another [20].

Cloud networking extends network virtualisation beyond the data centre to bring two new aspects to cloud computing: the ability to connect the user to services in the cloud and the ability to interconnect services that are geographically distributed across cloud infrastructures. These apects transform the cloud architecture of Figure 1 into that shown in Figure 2. Cloud networking users would be able to specify their needed virtual infrastructure and the desired networking properties to access these resources. Users would be able to specify how their infrastructure should be distributed in space and how it should be interconnected. They would be able to do this dynamically, on-demand, and through a single control interface.

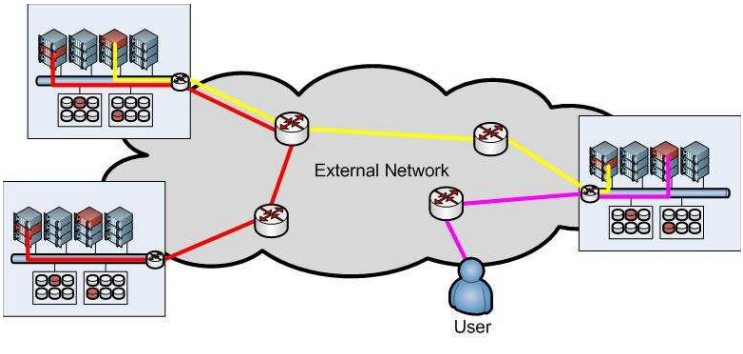


Fig. 2. Dynamic Virtual Networks Connecting a Distributed Service

The cloud paradigm has also encouraged the use of service automation. Applications running in cloud infrastructure can be programmed to monitor their own load and resource usage and dynamically scale themselves according to demand without the intervention of a human operator. Similarly, IaaS management systems optimise the use of physical resources by selectively deploying and migrating virtual machines. By introducing virtual networks to the same control plane the user and provider can make optimisation decisions based on network utilisation as well.

As more classes of applications are introduced to the world of cloud computing new requirements are brought with them. In some cases it may be more appropriate to deploy processing and storage functions across a network, that is, closer to the user, than to centralise processing and storage in a single location. Network conditions, such as latency, may hinder the execution of certain cloud applications in a data centre centralised fashion. Depending on the usage patterns one may need more servers in a certain geographical region. A geographically distributed cloud will enable finer control over the user experience. The previously mentioned content distribution services, as well as virtual desktop services are examples of this class of applications.

We anticipate that a wider range of trade-offs between costs and performance requirements will lead to a wider range of deployment options. To enable these new possibilities, it is critical that we understand the security implications and build appropriate mechanisms into the technologies we develop.

3 Security Problem Space

It is anticipated that security is one of the major factors influencing the acceptance for cloud computing in practical application domains, especially when sensitive information shall be brought into the cloud or IT governance requires an elaborated control regarding the (legal) liability of computing in clouds [21]. From a user's perspective the security topics distinguish infrastructure security, platform and application security, the security of the management processes, and finally compliance and governance [22].

The strength of the solutions that address these topics can be distinguished by the extent to which security objectives are met: who is allowed to do what (authentication & authorisation), how are system components and content protected (availability, confidentiality & integrity), how can the fulfilment of security properties be validated and checked (auditing), and how can the cloud provider prevent others from doing forbidden things (misuse protection).

Cloud networking adds new security challenges to the cloud computing security issues, arising from additional networking capabilities. On the other hand, there are indications that cloud networking can potentially improve control over the cloud computing deployment model, thus solving the security challenges that impact acceptance of this technology. The following is a preliminary threat model that is used in the SAIL project followed by a description of the security problem space as seen by the authors of the SAIL project at project start.

3.1 Threat Model

Information security properties are classically represented as Confidentiality, Integrity and Availability. We deal in this paper with technical threats; clearly, trust issues in cloud computing are also of a contractual or legal nature, but we do not intend to cover these herein. The cloud computing threat model addresses all three, but not necessarily in that order. We believe that the most important threat to information in a cloud computing environment is availability of the information to users whenever they need it. This availability issue manifests itself on the spot, e.g., through denial of service attacks. The likelihood and easiness of these attacks will increase as the volume of information exchanged between user and cloud provider increases. However, this property also needs to be preserved over time, avoiding for example format changes (so keeping legacy viewers, or translators). This also introduces integrity issues, as users must be certain that the information they retrieve is the same they stored. This might become difficult in a world where information is concentrated over such volumes of hardware that the checksum mechanisms currently in use do not allow us to ascertain that the data has not mutated, or that the translation applied preserves the content. Finally, confidentiality issues may arise, for example over (accidental) disclosure of information to third parties or because of aggregation. Most computer compromises result in information leakage, so this is also an important issue, but it clearly includes a regulatory compliance component which is outside the scope of this paper, hence our feeling that this is slightly less important than the two others.

Attackers will of course exploit the vulnerabilities that result of these threats according to their capabilities. In this paper we introduce a preliminary version of the attacker model that will be used in the SAIL project. This following gives a rough overview on roles and capabilities that an attacker might have. A more detailed description of the attackers, adjusted to the given scenarios, will follow in the project.

During the project we plan to base our attacker model on an external attacker that tries to access resources on the cloud infrastructure. To do this he can eavesdrop incoming and outgoing communication of the cloud networking infrastructure and try to get access to the infrastructure itself, e.g., by using vulnerabilities of the system. Additionally the attacker might be a legitimate user of the cloud networking infrastructure and uses this acces to attack other users' data [23].

For some scenarios also an internal attacker might be of interest, e.g., an employee of the cloud networking provider that accesses customers' data. A similar attack might be a supplier that introduces trapdoors in hard- or software in order to access data that is processed on the infrastructure.

External and internal attackers are also often used for analysing cloud computing. In the cloud networking case additionally legal aspects and legal intercepts have to be covered. Due to the fact that virtual components can move to arbitrary physical cloud networking infrastructures they might pass legal borders. Beside the fact that legal intercepts are not classical attacks they might violate

security goals of the cloud networking customers. Therefore, the location (legal space) has to be considered when distributing virtual components.

3.2 Information Security in Clouds

Information security relies on the classical three pillars, confidentiality (information should not be disclosed to unauthorised third parties), integrity (information should not be transformed without evidence of the transformation), and availability (information should not be withheld from rightful access).

The cloud scope adds a significant dimension in the mixing of code and data. Cloud users will need to ship code for execution on their data to cloud providers. Cloud providers will in turn ship code to users to easily manipulate the data. This is exemplified by the current rapid development of AJAX-based web services. Yet, this mixture of code and data is one of the major causes of malware infection, as it becomes extremely challenging to distinguish code from data and qualify the acceptability of both.

3.2.1 Trust in an Adversarial Environment. Cloud environments are by their very nature adversarial. Cloud *providers* balance the needs of their multiple users, and attempt to monetise by-products of their activity. Cloud *users* strive to obtain the cheapest possible services, while requesting services of high quality and respecting their privacy. *Attackers*, who have become very skilled at operating huge botnets (which can be seen as the first large scale clouds), will attempt to either access the information available in clouds or avail themselves to this processing power free of charge. All actors thus have their own trust objectives implemented in their security policies.

This adversarial setting promotes the use of security policy negotiation systems [24]. To maintain their trust relationship while ensuring sufficient flexibility to share resources, users and providers need to dynamically negotiate security policies balancing between operational trade-offs, such as cost and response time. This need will further develop as cloud providers aggregate and weave together complex service infrastructures federating many actors, creating the need for flexible and automated security policies aggregators and negotiators.

3.2.2 Confidentiality of Information and Processes. One of the most effective ways to maintain integrity and confidentiality of information is encryption. While encryption in its current form is sufficient for data storage and transport, it fundamentally prevents data processing. Thus, sending encrypted data to cloud providers for processing is quite useless. This challenge has been met by homomorphic cryptography (HC). Homomorphic cryptography ensures that operations performed on an encrypted text results in an encrypted version of the processed text. Recently, a solution under ideal circumstances has been presented [25,26]. However, practical application is still far away since the computational effort required to retrieve the results of the computation is still too high and thus HC remains of theoretical value only for the coming years.

As a result, users will not have a solution based on cryptography that allows them to rely on information confidentiality and integrity when providing code

and data to an arbitrary cloud. They have no means to ensure that their data is not misused. Until HC provides a formal solution to this issue, we need to rely on audit traces to assert "after the fact" usage control demonstrating that data and code have not been misused by service providers and cloud users. These audit traces can be part of a security policy specification, and can be supported for example by the OrBAC (Organisation-based access control) language. Further, we do not know yet if other solutions, e.g., watermarking will be portable to the cloud computing world and if their properties will be preserved in this world.

3.2.3 Policy Models and Policy Enforcement. The currently available security policy models are not sufficiently flexible. For example, the OrBAC model [27], one of the recent attempts to further develop the classic RBAC model, introduces *organisations* and *contexts* in addition to the classic notion of roles. Both concepts are extremely useful to define security policies that span organisational boundaries (in our context multiple users sharing a cloud provider or a federation of cloud providers uniting for a specific service) or security policies that are flexible according to environmental conditions (for example service load or cost). However, the combination of organisations and contexts with negotiation remains largely an unsolved problem. The complexity of these policies has not been resolved either. Even in simple environments such as network firewall filtering, users have difficulties understanding the impact of filtering rules when the number of rules is large or when multiple firewalls are traversed. We expect that this complexity may become a barrier to the deployment of cloud computing if these policies cannot be simply explained and proven to all parties.

Security policies need to be enforced. Technology for this enforcement is reasonably well established using Policy Enforcement Points (PEP) controlling access to resources. Network firewalls, web application firewalls, identity management systems, file system access controls are well-know entities with clear properties. Policy Decision Points (PDP) are in charge of managing such PEPs and taking over for complex access control requests.

However, there is no such clear picture for the cloud computing world. First, it is not known if the policy enforcement technology can be ported into the cloud world and how. Second, it is unclear if cloud computing enabling technologies, such as virtualisation, will bring new PEPs. Once this setting is clearer, we will need to define techniques that will weave PEPs and PDPs into a cloud service definition and tools for verifying that the resulting "secure service" definition meets the security objectives of all parties. It specifically requires new tools that will enable partial verification of security objectives so that all parties (users and providers) can reliably verify that their security objectives are met, without knowing the security objectives of the other parties.

3.3 Virtualisation Environment Threats

Analysis of security threats in virtualisation environments provides some insight about the challenges raised by virtualisation of computing resources and

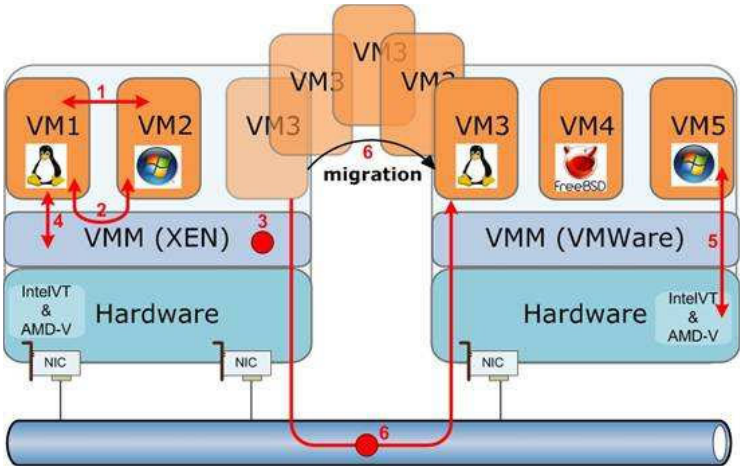


Fig. 3. Security Threats in Virtualised Environments

networking. Figure 3 highlights six different security threats that might emerge when a type II hypervisor is used. They can be classified into software level (1, 2, 3, 4 and 6) and system level (5) concerns.

3.3.1 Isolation between Virtual Machines. In this case, each virtual machine uses and only reads its allocated resources. For example, the memory management is subdivided into multiple levels (Hypervisor level, Host VM level and Guest VM level). The Hypervisor can read all the physical memory space. The Host Virtual machine (dom0 for XEN) can read all the memory except the memory allocated to the hypervisor. Guest Virtual Machines (domU for XEN) can only read their allocated memory. This isolation between different virtual machines is one of the main important roles of a hypervisor. As a solution, selection of the right hypervisor can ensure this isolation between virtual machines.

3.3.2 Information Theft through Malicious Use of Hypervisor. To share physical resources, the hypervisor uses different techniques depending on the physical components to share. For example, to share physical network cards, the hypervisor (see the case of XEN at [28]) can use Bridged, NATed or Routed networking. In Figure 4, there are two bridges (*xenbr0* and *xenbr1*) that virtualize two physical network cards (*peth0* and *peth1*). The bridge *xenbr0* connects physical interface *peth0* to three virtual interfaces (*vif0.0*, *vif1.0* and *vif2.0*). Each virtual interface is connected to a virtual machine. In this configuration, despite the fact that all interfaces use the same bridge, it is necessary to ensure that a virtual machine cannot read the packets of the bridge that are sent to another virtual machine. This can be accomplished by the hypervisor or just by applying existing security solutions. To reduce the burden on the hypervisor in managing network I/O activities, manufacturers have since introduced Virtual Machine Device Queues (VMDq) [29] and Single Root Input Output Virtualisation

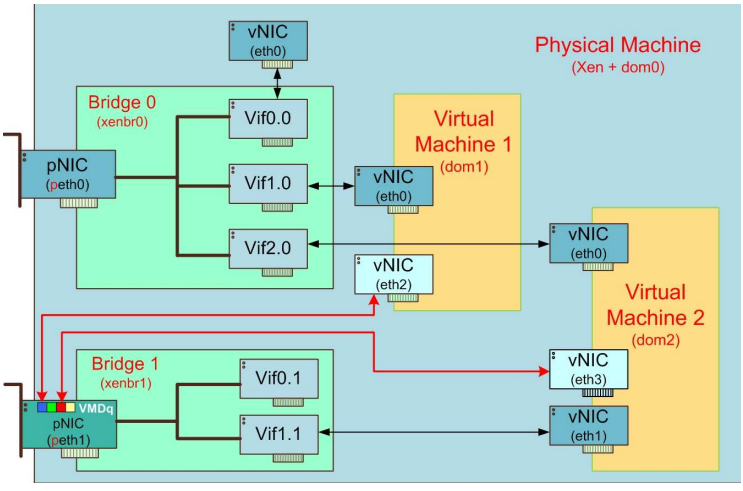


Fig. 4. Bridges Sharing Physical Network Cards

(SR-IOV) [30]. Sorting data packets in the network silicon frees CPU cycles for application processing instead of network I/O processing. These new technologies introduce the additional requirement of securing, protecting, and isolating also the network card virtualisation.

3.3.3 Untrusted Hypervisors. If the owner of the physical machine wants to read and steal the data of virtual machines, she or he can do it using the hypervisor (an untrusted hypervisor). In this case, each user of a virtual machine needs to have a solid contract with the owner of the physical machine. Having a contract is a good and necessary thing, but it is imperative that virtual machines use their own mechanisms to secure themselves. For example, encrypting a virtual machine is a potential solution for this kind of problem.

3.3.4 Untrusted Virtual Machines. It is always possible to have a contract to build some trust between the user of the virtual machine and the owner of the hypervisor. However, there is still the problem of the impossibility to have any idea about the other virtual machines than could be deployed in the same physical machine. A virtual machine can try to get control of the hypervisor using software related security holes without informing the hypervisor owner. Then, this virtual machine can get partial or total control of the physical machine. Technically, this is a similar situation to the untrusted hypervisor scenario. In this situation it is possible to apply the same security solutions as in the untrusted hypervisor case.

3.3.5 Untrusted Virtual Machines Misusing Hardware Virtualisation Functionality. To increase the performance of virtual machines in a virtualised environment, different functionalities (dedicated to virtualisation) have recently

appeared in the architecture of physical components. As an example, new instruction sets (IntelVT-x [9,31] or AMD-V [10]) have been introduced in the most recent processors. With these functionalities, a virtual machine can send instructions directly to the processors, bypassing the hypervisor.

All the previously mentioned security problems can be solved using or adapting existing techniques. However, with these new types of security problems related to the hardware, a new philosophy and family of problems appear. Examples of systems that are sensitive or subject to these security threats are SubVirt and BluePill [32,33].

3.3.6 Unsecure Network Transfer on Inter Device Migrations. In a virtualised environment, a virtual machine can migrate from a physical machine to another. This migration through the network can use traditional or new protocols, which can be exploited to attack the system. It is imperative to protect this migration by using or adapting existing techniques to prevent attacks on migration control mechanisms, transactions, and protocols. This sixth identified threat is central in SAIL that focusses on cloud networking. This key aspect is tightly linked to auto-scaling and elasticity properties of clouds. In addition, there is a need for virtual firewalls for isolating dynamic VPNs and virtual networks allocated on the fly and on demand, to create dedicated flash slices.

3.4 Communication Security

Communication between virtual infrastructure, as well as the distribution of virtual infrastructures, generate traffic in the network, which has to be secured. The following Section 3.4.1 shows the challenges of securing the communication between virtual components, while Section 3.4.2 shows the security challenges of cloud networking, i.e., moving virtual components in space, and its management.

3.4.1 Secure Virtual Networking. In addition to cloud computing, virtual networking introduces new security challenges by enabling communication between different virtual components. From a virtual network user's perspective the network might be private while in reality the communication itself occurs via a public infrastructure. Therefore, mechanisms to secure this communication (e.g., by encryption) have to be established. One option is to do it in each virtual component, which means that the virtual network customer has to care for securing the communication. Another option is to provide secured communication as a service by the virtual network provider, which means that the communication is secured by default and transparent to the customer.

Besides securing the communication itself in virtual networks, the management of the communication also has to be secured. By virtualising networks and network components new attacks arise and need to be handled. Due to the abstraction layer introduced by virtualisation, existing techniques might not be applicable or have to be adjusted or extended to fit this new setting. Especially the integrity of the virtual network topology and components, as well as the security of routing in these networks, need to be addressed.

Additionally, similar challenges as in cloud computing also exist in virtual networks. This includes how the virtual network provider guarantees a certain network capacity to a customer, how the access to this virtual network is controlled, and how the virtual network usage is accounted for.

3.4.2 Secure Management of Cloud Networking. For the management of cloud networking access to the physical infrastructure and to the network properties is needed. This access should be implemented as a single interface, where a user can specify several parameters on-demand.

By the combined access to the physical virtualisation infrastructure and the network infrastructure new attacks arise. One challenge is to define rules for accessing the management interface and how to implement these rules. Also policies for moving virtual infrastructures in space need to be distributed. These policies might define to which location (legal space) a virtual infrastructure is allowed to move, as the location of the physical infrastructure determines the legal restrictions that apply to the virtual infrastructure (e.g., USA Patriot Act [11]).

3.5 Misuse of Cloud Networking Capabilities

The ability of cloud computing and cloud networking to allocate computational resources on demand can also be misused, e.g., for DoS attacks, spamming, and providing illegal content. Attacks that use cloud infrastructures are already known today. One example is Zeus "in-the-cloud" [34] where the command and control of a botnet was located at the Amazon EC2.

Auditing can help to detect these kind of attacks, e.g., by looking for fast fluxing or domain fluxing. The challenge of automated detection of attacks is to distinguish misuse from legitimate use. Trying to find anomalies might be one way to solve this problem. If a misuse can be detected the attack can simply be interrupted by discontinuing the virtual infrastructure, which is involved in the attack.

By introducing cloud networking no new threats are added to those already known from cloud computing. Therefore, countermeasures for misuse of cloud networking can be adapted from cloud computing.

4 Conclusion and Future Work

This position paper introduces the cloud networking specific security challenges that will be addressed in the SAIL project. These challenges can be grouped into protection of cloud content, secure virtualisation technology, distribution transparency control, and secure operations. There are clear benefits that come with cloud networking for cloud users and operators. Also operators have prospect to support effectively cloud operators with their available network and transport capabilities for the benefit of end users. The road may even be open for further scenarios, e.g., connecting multiple clouds or introducing more heterogeneity, which in turn will increase the complexity in multilateral security. Both cloud

computing and virtual networking have each their own security challenges, the ones presented here have to be considered for securing and protecting cloud networking that seeks technical solutions to ensure acceptance of this new concept.

Acknowledgement

This paper describes work undertaken for the project SAIL (Scalable & Adaptive Internet soLutions, Project number 257448), which is part of the EU's IST program. 24 organizations from Europe, Israel and Australia are involved in this Integrated Project, which runs in 2010 – 2013. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the SAIL Project.

References

1. SAIL project website (2010), http://www.sail-project.eu/
2. Provos, N., Rajab, M.A., Mavrommatis, P.: Cybercrime 2.0: When the cloud turns dark. Queue 7(2), 46–47 (2009)
3. McCarthy, J.: MIT Centennial Speech of 1961 cited in Architects of the Information Society. In: Garfinkel, S.L. (ed.) Thirty-five Years of the Laboratory for Computer Science. MIT, Cambridge (1999)
4. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson, D.A., Rabkin, A., Stoica, I., Zaharia, M.: Above the clouds: A berkeley view of cloud computing. Tech. Rep. UCB/EECS-2009-28, EECS Department, University of California, Berkeley (2009)
5. Barham, P., Dragovic, B., Fraser, K., Hand, S., Harris, T., Ho, A., Neugebauer, R., Pratt, I., Warfield, A.: XEN and the art of virtualization. In: SOSP 2003: Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles, pp. 164–177. ACM, New York (2003)
6. VMware (2010), http://www.vmware.com
7. Edwards, A., Fischer, A., Lain, A.: Diverter: A new approach to networking within virtualized infrastructures. Tech. Rep. HPL-2009-231, HP Laboratories (2009)
8. Amazon elastic block store (2010), http://aws.amazon.com/ebs/
9. Intel virtualization (2010), http://www.intel.com/technology/virtualization/
10. AMD Virtualization (AMD-V) Technology (2010), http://sites.amd.com/us/business/it-solutions/virtualization/Pages/amd-v.aspx
11. Fraser, D.: The canadian response to the USA Patriot Act. IEEE Security Privacy 5(5), 66–68 (2007)
12. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (directive on privacy and electronic communications). Official Journal of the European Union, L201, 0037–0047 (2002)
13. Amazon virtual private cloud (2010), http://aws.amazon.com/vpc/
14. Pallis, G., Vakali, A.: Insight and perspectives for content delivery networks. Commun. ACM 49(1), 101–106 (2006)

15. Chowdhury, N.M.K., Boutaba, R.: A survey of network virtualization. Computer Networks 54(5), 862–876 (2010)
16. Bavier, A., Feamster, N., Huang, M., Peterson, L., Rexford, J.: In VINI veritas: realistic and controlled network experimentation. In: SIGCOMM 2006: Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, pp. 3–14. ACM, New York (2006)
17. Feamster, N., Gao, L., Rexford, J.: How to lease the internet in your spare time. SIGCOMM Comput. Commun. Rev. 37(1), 61–64 (2007)
18. Schaffrath, G., Werle, C., Papadimitriou, P., Feldmann, A., Bless, R., Greenhalgh, A., Wundsam, A., Kind, M., Maennel, O., Mathy, L.: Network virtualization architecture: proposal and initial prototype. In: VISA 2009: Proceedings of the 1st ACM Workshop on Virtualized Infrastructure Systems and Architectures, pp. 63–72. ACM, New York (2009)
19. FEDERICA: Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures (2010), http://www.fp7-federica.eu/
20. Wang, Y., Keller, E., Biskeborn, B., van der Merwe, J., Rexford, J.: Virtual routers on the move: live router migration as a network-management primitive. SIGCOMM Comput. Commun. Rev. 38(4), 231–242 (2008)
21. Brunette, G., Mogul, R.: Security guidance for critical areas of focus in cloud computing v2.1. Cloud Security Alliance (2009)
22. Streitberger, W., Ruppel, A.: Cloud computing security - protection goals, taxonomy, market review. Tech. rep., Institute for Secure Information Technology SIT (2010)
23. Ristenpart, T., Tromer, E., Shacham, H., Savage, S.: Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In: CCS 2009: Proceedings of the 16th ACM Conference on Computer and Communications Security, pp. 199–212. ACM, New York (2009)
24. Abi Haidar, D., Cuppens-Boulahia, N., Cuppens, F., Debar, H.: XeNA: an access negotiation framework using XACML. Annals of Telecommunications 64(1), 155–169 (2009)
25. Gentry, C.: Fully homomorphic encryption using ideal lattices. In: STOC 2009: Proceedings of the 41st Annual ACM Symposium on Theory of Computing, pp. 169–178. ACM, New York (2009)
26. Gennaro, R., Gentry, C., Parno, B.: Non-interactive verifiable computing: Outsourcing computation to untrusted workers. In: Rabin, T. (ed.) CRYPTO 2010. LNCS, vol. 6223, pp. 465–482. Springer, Heidelberg (2010)
27. Abou El Kalam, A., Baida, R.E., Balbiani, P., Benferhat, S., Cuppens, F., Deswarte, Y., Miège, A., Saurel, C., Trouessin, G.: Organization Based Access Control. In: 4th IEEE International Workshop on Policies for Distributed Systems and Networks, Policy 2003 (2003)
28. XEN networking blog (2010), http://wiki.xensource.com/xenwiki/XenNetworking
29. Chinni, S., Hiremane, R.: Virtual machine device queues (VMDq) - white paper (2010), http://software.intel.com/file/1919
30. Pci-sig single root i/o virtualization (sr-iov) support in intel virtualization technology for connectivity - white paper (2008), http://www.intel.com/network/connectivity/solutions/SR-IOV-046NTL_Whitepaper_061308.pdf

31. Uhlig, R., Neiger, G., Rodgers, D., Santoni, A., Martins, F., Anderson, A., Bennett, S., Kagi, A., Leung, F., Smith, L.: Intel virtualization technology. Computer 38(5), 48–56 (2005)
32. Price, M., Partners, A.: The Paradox of Security in Virtual Environments. Computer 41(11), 22–28 (2008)
33. King, S.T., Chen, P.M., Wang, Y.M., Verbowski, C., Wang, H.J., Lorch, J.R.: SubVirt: Implementing malware with virtual machines. In: IEEE Symposium on Security and Privacy, pp. 314–327 (2006)
34. CA Community Blog: Zeus "in-the-cloud" (2009), http://community.ca.com/blogs/securityadvisor/archive/2009/12/09/zeus-in-the-cloud.aspx

Video-Enhancing Functional Architecture for the MEDIEVAL Project

Daniel Corujo[1], Albert Banchs[2], Telemaco Melia[3], Michelle Wetterwald[4], Leonardo Badia[5], and Rui L. Aguiar[1]

[1] Universidade de Aveiro, DETI, 3810-193 Aveiro, Portugal
[2] IMDEA Networks, Avenida del Mar Mediterraneo 22, 28918 Leganes (Madrid), Spain
[3] Alcatel-Lucent, Route de Villejust, 91620 Nozay, France
[4] Mobile Communications Dept., EURECOM, 06904 Sophia Antipolis, France
[5] Consorzio Ferrara Ricerche, via Saragat 1, 44122 Ferrara, Italy
dcorujo@ua.pt, banchs@it.uc3m.es,
telemaco.melia@alcatel-lucent.com,
michelle.wetterwald@eurecom.fr,
leonardo.badia@gmail.com, ruilaa@det.ua.pt

Abstract. The MEDIEVAL project aims to leverage today's Internet with the necessary fabric to provide optimized video services in a mobile wireless world. It is expected that video traffic will surpass Peer-to-Peer (P2P) in volume in the coming years, and thus novel mechanisms and techniques need to be provided to better suit its unique requirements. This article describes the key functional elements of the MEDIEVAL architecture, which provides a video-aware networking core coupled with abstracting interfaces which cater to service and access technology specific requirements, aiming to enable efficient video transport and novel video service development.

Keywords: Wireless networks, Mobile communication, Video services, Radio optimization, Multicast/Broadcast.

1 Introduction

The EU project MultimEDia transport for mobIlE Video AppLications (MEDIEVAL) [1] is a collaborative project with a three-year duration starting on 1st July 2010, having as partners Alcatel-Lucent Bell Labs France, Telecom Italia, Portugal Telecom Inovação, Docomo Communications Labs, LiveU Ltd., Instituto de Telecomunicações, Universidad Carlos III de Madrid, Consorzio Ferrara Richerche and Eurecom. It aims to evolve today's mobile Internet architecture to more efficiently support the upcoming growth of video services. According to [2] P2P, as the current dominant source of traffic in the Internet, will be surpassed by video in 2010 achieving volumes close to 90% of consumer traffic by 2012. This increase is motivated by a change in perception and usage of video services such as Internet TV, interactive video, Video on Demand (VoD), among others, which instead of being regarded as simple streaming of content, will become a tool for personal multimedia communication, resembling today's explosive usage of personal messaging (i.e., Short Message Service (SMS) and Twitter.

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 314–325, 2011.

However, the Internet, and the mobile technologies therein, have not been designed to properly sustain such an increase of video, in an optimized way. This is where MEDIEVAL intervenes by providing a more suitable video transport architecture, commercially deployable by network operators. This article is organized as follows. In the next section, we will present the vision of the MEDIEVAL project, focusing on its concept and main objectives. This is followed by section 3 where we discuss the general MEDIEVAL architecture, which provides video-specific enhancements at different layers of the protocol stack, by exploiting cross-layer approaches that aid in better video support. In the subsequent three sections we will describe the major points over which the architecture is impacted and aims to provide solutions, detailing some of the approaches: network requirements for video (Section 4), packetization (Section 5), and multicast mechanisms for video optimization (Section 6). Finally, we conclude in Section 7.

2 Vision

The vision of MEDIEVAL considers the evolution of video as a primary source of content, accessed as well as generated, over the Internet. This is exactly where MEDIEVAL aims to contribute: evolving the mobile Internet architecture for efficient video traffic support.

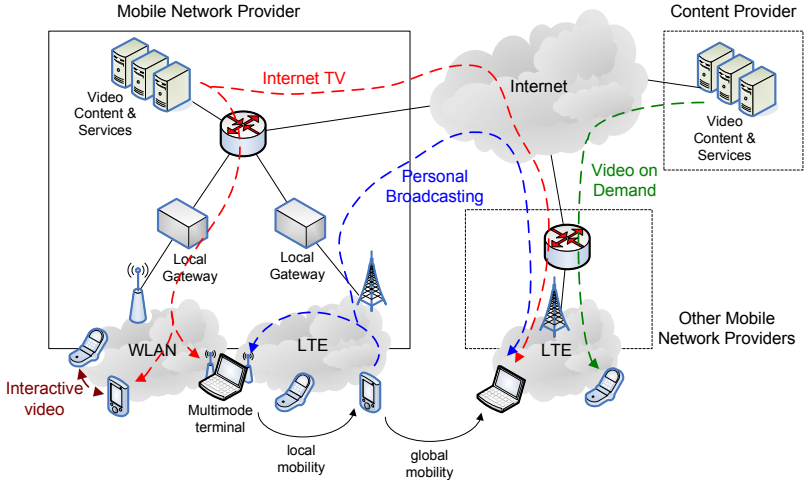


Fig. 1. The MEDIEVAL vision

Fig. 1 presents a visualization of this vision, highlighting what we foresee as the required evolutionary path for a true video-for-all philosophy, providing selected application examples. Four primary video services are considered, comprising VoD, Internet TV, Interactive Video and Personal Broadcasting, offered by the providers in the figure. These services are accessed by terminals supporting different access technologies (Long Term Evolution (LTE) and IEEE802.11 in the figure), while on

the move, through both local as well as global mobility procedures, in intra and inter-domain scenarios. The video services are accessible from content and services providers available at a home operator domain (visible inside the Mobile Operator Provider part of the figure), as well as from content providers or other mobile network providers (visible in the separate Content Provider part of the figure). However, MEDIEVAL also envisages scenarios where users, and thereby their mobile terminals, are the source of generated video content, enabling scenarios of direct interactive video.

The necessary technical solution and problem solving that enable such vision span to all areas of mobile communications, starting from the need to enhance wireless access technologies, requiring efficient mobility management as well as optimized transport, to video distribution mechanisms and network-aware applications and services. We believe that a cross-layer approach will not only provide clear innovations in all the mentioned fields, but will also lead to a realistic evolutionary path for mobile networks, truly providing an environment where users can benefit from the MEDIEVAL vision.

This vision, and the subsequent architecture, will address the following five key issues:

- Design and specification of a set of interfaces between video services and the underlying network mechanisms, allowing the video services to customize the network behavior in an optimal way.
- Enhance the wireless access to provide an optimized video performance experience through the coordination of the features of the wireless technologies and the video services.
- Design and specification of a novel mobility architecture for the next generation of mobile networks, truly adapted to video service requirements.
- Optimize video delivery systems with Quality of Experience (QoE) driven network mechanisms through the combination of Content Delivery Networks (CDN) and P2P techniques for optimized video streaming focusing on the location of caches and peer selection.
- And lastly, support for broadcast and multicast video services, including Internet TV and Personal Broadcasting, through the introduction of multicast-aware mechanisms at the different layers of the protocol stack.

Through the addressing of these key issues, MEDIEVAL intends to perform technological developments based on an operator-driven architecture aiming to have an innovative impact in terms of video services performance improvement over existing solutions, while providing an integrated video solution that can be implemented by an operator. This integrated solution will observe better QoE of video to users by providing a joint view of user and video services requirements, wireless network conditions (such as performance and load) and transport optimization, all of which impact mobility decision taking.

The support of this vision requires not only the development of a video transport architecture, but also a high emphasis on its commercial deployment suitability. To further this, the design of a set of specific mechanisms and enhancements with application on video services will have to be developed and presented to the relevant standardization bodies.

Even though the architecture focuses on video, this kind of traffic is the most demanding in terms of bandwidth thus enabling other applications to work since the network is dimensioned for video, and with the same protocols (albeit using different algorithms) we can treat other kinds of traffic.

3 Architecture

The MEDIEVAL architecture relies on four functional cornerstones, which are depicted in **Fig. 2**, considering Wireless Access, Transport Optimization, Mobility and Video Services and described in the following subsections.

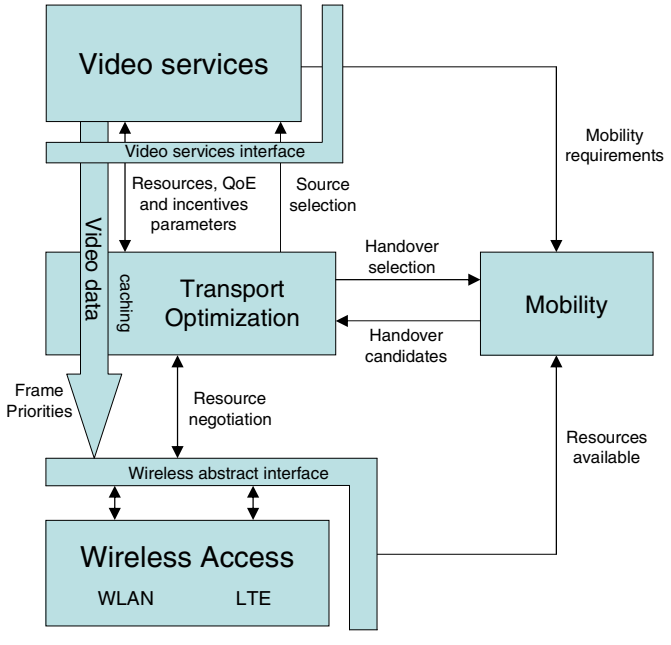


Fig. 2. The MEDIEVAL architecture

3.1 Wireless Access

Wireless access considers coordination mechanisms between video services and the different wireless access technologies in order to optimize overall performance. MEDIEVAL will focus on the LTE of the Universal Mobile Telecommunications System (UMTS) for coordination-based access, and the IEEE802.11 standards for contention-based wireless access, enabling a "video over wireless" concept. On the IEEE802.11 side, the high rates placed by video services when this kind of traffic is prioritized against non-real time traffic allow for good QoE, but only in limited scenarios [16][17]. To counter this, numerous approaches have been presented but are either based on heuristics and do not guarantee optimal performance [18][19] or introduce significant complexity while requiring the interaction between codecs and

the MAC layer, increasing their deployment difficulty [20][21][22][23]. On the LTE side, cross-layer optimizations are key improvements over existing cellular technologies regarding the support of video services, along with data rate increase obtained through intelligent exploitation of radio resources, which enable interesting scenarios complemented with the introduction of point-to-multipoint capability such as Multicast/Broadcast Multimedia Service (MBMS) [5] and evolved MBMS (eMBMS). An interesting project concerning cross-layer mechanisms to improve multimedia performance over wireless links is OPTIMIX [25], but it derives a new wireless architecture from scratch which can seriously hinder deployability with operators and does not consider entirely mobile scenarios.

In MEDIEVAL, interaction with different technologies will be achieved through heterogeneous cross-layer mechanisms for interface abstraction (i.e., IEEE802.21 [6]), facilitating interfacing with the upper layers and their exploitation of specific features for video optimization. Another key innovation will be the introduction of enhanced dynamic configuration procedures based on the current network conditions, while also exploiting advanced terminal features such as multiple-input multiple-output (MIMO) capabilities, cognitive networks and multi-hop transmission in centralized technologies.

3.2 Video Services

Video services control will provide these services with the tools for reliable video delivery over an evolved mobile network. Some approaches already consider content adaptation and the use of RTSP (Real Time Streaming Protocol) but they do not consider, in the first case, the dynamicity of the network, the various network policies, QoS requirements and, in the second case, do not allow session negotiation. Other measures, such as packet prioritization, are not yet investigated in relatively new codecs (i.e., H.264, AAC audio and SVC), and Forward Error Correction (FEC) mechanisms negotiation at application level does not exist. All of these are imperative requirements in mobile environments. PHOENIX [26] proposes a cross-layer optimization of wireless access where video is an application of the proposed optimization framework. However it does not consider entirely mobile scenarios and video is just regarded as another application.

MEDIEVAL will provide video services with an interface enabling them to interact with core network mechanisms, considering requirements and features from both, and properly adapt the service execution. Through this interface, video services are able to provide indications regarding the type of video data in order to request frame prioritization to the wireless layers, as well as providing resources, QoE and incentive parameters to better optimize transport procedures. Other measures such as algorithms to better adjust video streaming to the network conditions will be tackled considering cross-layer dynamic adaptation management techniques, prioritization methodologies and appropriate FEC mechanisms.

3.3 Transport Optimization

To capitalize from the video-aware features of both the wireless access and video services functional components, the architecture also provides video-aware transport

optimization mechanisms for efficient video delivery, offering resilient and mobility-aware QoE to video services. Several techniques already exist such as analysis of interactions of video transport with other layers [29] and CDN. However, they miss a global system view and there is a lack of knowledge about the operation of CDN in the context of mobile networks. The NAPA-WINE [27] project considers P2P systems for high quality TV delivery, managing traffic and quality degradation, albeit not considering mobile networks nor broadcast/multicast solutions as well as operator-controlled mechanisms.

MEDIEVAL aims to develop transport optimization cross-layer mechanisms able to execute resource negotiation with the wireless layers, considering an optimized source selection, while adapting the video service to current network conditions. The usage of CDN architectures in mobile environments will also be pursued as well as providing solutions for dynamic rate-control and caching schemes. Resource reservation procedures are considered to be in place and MEDIEVAL will use them as implemented by each technology (LTE and WIFI). However, here we go one step further addressing QoE in a different manner by enabling the network to understand what traffic is traversing the routers and will be able to take decisions before routing the packets (e.g., dropping specific video frames, using a different path, etc.).

3.4 Mobility

The reference network architecture for MEDIEVAL is the EPS specified by 3GPP (Release 8), and Proxy Mobile IPv6 (PMIPv6) [12] and Dual Stack Hosts and Routers (DSMIPv6) [24] are the two major IP mobility protocols proposed. However, these efforts are still based on anchor points and tunnels, and mobility is offered as a general service which can employ unnecessary overhead when not needed. CARMEN [28] addresses a mesh architecture with mobility support, while considering video at carrier grade. MEDIEVAL also aims at the design of a novel mobility architecture but focuses more on video efficiency rather than just support it.

Considering the mobile environments over which video services will be used, the architecture will provide mobility mechanisms which are also video-aware, and will interact with the core network to ensure optimized connectivity for the terminals. This is achieved through the collection of handover requirements from video services, the identification of available resources and their impact in handover candidate selection taking into consideration an optimized transport execution. These four architectural components will feed handover selection algorithms through the provision of parameter values that consider an overall view of the best choice possible.

These mobility mechanisms will address both local mobility and global mobility, considering, in the first case, the provision of service continuity customized to the different requirements of video services and, in the second case, addressing inter-operator roaming issues without requiring the deployment of global anchor points in the operator's network. Here the focus is also on session continuity. An important innovative point from MEDIEVAL is to consider mobility in terms of specific flows.

Thus, the mobility architectural part will consider three main areas of intervention: i) mobility mechanisms for multi-mode terminals and moving networks supporting mobile and network initiated handovers; ii) video-aware interface for heterogeneous wireless access conveying video relevant information for optimal decision taking by

the mobility function and iii) IP multicast mobility by both sources and receivers, (i.e. considering their issues with tree-based approaches).

4 Cross-layer Mechanisms for Addressing Network Requirements

For an optimized video experience, the full set of MEDIEVAL's functional architecture needs to work cooperatively, providing the bridge that allows video services, and their traffic, to be adapted to current network conditions based on the user terminal's selected technology while on the move. An open problem to be tackled is how to interrelate network dynamics and QoE requirements for video services in wireless environments. Proposals addressing these issues don't consider session negotiation or only allow it before the connection is established [3], which prove unfeasible in dynamic environments. Under these environments, mobility has been thoroughly studied, leading to extensive optimization efforts for handover execution, but never considering video specifically. The transport video traffic under these conditions gains a key importance where increasing requirements for more bandwidth are coupled with stringent delay constraints, while operating in heavily congested networks. Network requirements have to be considered under the general-purpose behavior of the Internet, where other different kinds of traffic coexist, raising the interest of the IEEE802.11 in previous extensions [4].

To address network requirements in such environments, we explore cross-layer interactions, focusing on all layers of the network stack, applying improved management towards reliable and smooth transmission. An important tool to be used for this cross-layer interaction is the IEEE802.21 Media Independent Handovers (MIH) standard. MIH considers the optimization of handovers in multi-technology environments, by providing mechanisms that rely on the abstraction of the different connectivity technologies and provide media independent information and control to deciding entities, regarding the medium status. The introduction of MIH mechanisms work as a layer 2.5 abstraction concept, enabling MEDIEVAL to encompass future communication systems while tackling their inherent heterogeneous characteristics and challenges. Also, the media-independent signaling provided by the IEEE802.21 MIH protocol will provide the interaction between the different cross-layer components on which MEDIEVAL intends to impact.

However, the IEEE802.21 standard was not conceived for any specific kind of traffic and does not attempt to take the optimization perspective of the network for video delivery. Its introduction in the MEDIEVAL framework will leverage and enhance it to support video specific extensions (e.g., link capacity versus packet prioritization), by taking advantage of its intrinsic signaling primitives. Concretely, MEDIEVAL will extend the interfaces with higher layer services considering the interaction between mobility and transport optimization components, enabling IEEE802.21 to provide fine granular IP flow mobility management, while the interaction between the video service and the mobility components will also leverage already existing protocols, such as DIAMETER [7], through the creation of the required extensions. Thus, IEEE802.21 will be extended to convey video-specific information (such as encoding parameters, real-time QoE parameters, among others) enabling the provision of indications to handover decision entities which assist in applying the best procedure possible, depending as well on user credentials.

Although the extension of the MIH protocol to execute new personalized behavior has already been proposed in other contributions (such as European projects), to our knowledge, this is the first time that a similar rationale is applied to video and its inherent services. An important distinguishing point is that the objective is not to use media independent mechanisms to provide the same abstraction for different technologies, since in this case they are based on different principles that operate very differently when it comes to video traffic. As such, our aim is to provide a set of abstract interfaces allowing each medium to report its capabilities to decision video-aware entities which can then exploit them through the same interfaces.

MEDIEVAL will benefit from the inclusion of ODTONE [8], standing for Open Dot Twenty ONE, which is an open-source implementation of the IEEE802.21 standard from the Instituto de Telecomunicações (Aveiro, Portugal). ODTONE is implemented in C++ using Boost and provides an operating system independent Media Independent Handover Function (MIHF), the core entity of the IEEE802.21. To enable its integration with the different link layer technologies, being managed by different operating systems, ODTONE provides a library based on the MIH protocol, which can be used to implement the different link Service Access Points (SAP). This provides an ample platform able to be executed in different environments, featuring different terminals and access technologies and thus not being dependant of a single operating system, which is the case of other initiatives such as [15]. The open-source nature of the project, and the expertise gained by its development, will provide to the MEDIEVAL project with the necessary tools with which to extend the base IEEE802.21 behavior, enabling it to provide optimized execution for video services aware mobility and data transport.

5 Packetization

MEDIEVAL will exploit packet-level mechanisms and techniques available to its functional elements and core network in the various types of technologies. The control plane of the LTE Radio Access Network (RAN) will intervene at user plane entities with the aim of selecting and prioritizing video frames. With respect to the lower layers of the wireless technologies, the project will evaluate mechanisms, such as the ones under standardization in 802.11aa, where dynamic prioritization for frame marking and discarding are supported, and techniques such as graceful performance degradation are employed.

The usage of jumbo frames to aggregate packets while enhancing video delivery mechanisms to achieve higher video throughputs will be evaluated, taking into consideration the necessary extensions for the cross-layer interaction of video services with LTE and IEEE802.11 networks. Jumbo frames allow the usage of larger frames extending them to 9KB which take advantage of reduced MAC overhead, increased throughput and less CPU usage. However, they also introduce new problems such as larger hardware requirements on routers, and more video data is lost when a packet is lost or delayed. This is crucial in wireless environments running interactive video services, requiring the adoption of important measures such as zero-loss mechanisms, needing a feasibility study considering their effect in mobile environments and their impact on real-time services and video buffers. In this feasibility study, the suitability

of current IEEE802.11 mechanisms (i.e., such as the TXOP (Transmission Opportunity) parameter of the MAC protocol) will be analyzed for the case of jumbo frames aiming towards video services optimization. The TXOP parameter has been extensively used as the means for modification of the standard transmission procedure to achieve optimized results [9][13]. However, the MEDIEVAL framework intends to provide a comprehensive cross-layer approach towards the optimization of video services and thus application of jumbo frames at the MAC layer is not enough (or even only at the IP level itself [10]): the other layers, involving the video services, the transport and mobility procedures, must be aware of this factor and to know if the conditions are favorable towards its usage. As such, the interface between layers needs to be extended to convey this cross-layer information, towards the optimization of the usage of larger frame sizes to increase video performance. The studies executed at these two fronts will determine whether jumbo frames will be used in WLAN technologies or not, and, in parallel, the feasibility for the usage of this kind of frames will also be analyzed within LTE, towards the enhancement of video delivery mechanisms. Here the objective is to extend the cross-layer interaction of video services evidenced in the WLAN case, with the LTE architecture [14], enabling the usage of jumbo frames to achieve higher throughputs for video under this wireless technology as well.

6 Multicast Mechanisms

A key development for the proliferation of video traffic is its widespread diffusion within social network, as witnessed today on Facebook or MMS. However, the deployment of such features in today's Internet while considering video traffic being generated by millions of mobile wireless users, emphasizes the lack of interconnection mechanisms supporting this trend. It is just not feasible to send independently video feeds towards users viewing the same content. Although solutions for multicast exist, these approaches do not consider scale service announcement and discovery as well as mapping video service groups into network-based groups while managing different content sources. MEDIEVAL will focus on providing a common interface that allows different applications to efficiently deliver video content to user groups, leveraging multicast and broadcast context solutions (MBMS and eMBMS). The inclusion of these mechanisms at IP level, with special nodes acting as the heads of the multicast distribution trees, will also be enhanced with bearer service preparation to optimize scenarios where terminals change into a new cell not yet in the session topology.

Another key intervention point for MEDIEVAL is the crossing of multicast and mobility mechanisms, particularly network-based localized mobility management solutions. For this, a thorough analysis on optimal multicast support in PMIPv6 [12] will be done. Here, the project will benefit and contribute to standardization via a recently formed IETF working group in the Internet Area: Mobility Multicast (MULTIMOB) [11], aiming to provide guidance and multicast support in a mobile environment. An important consideration to tackle, considering that a mobility management protocol that is network-based such as PMIPv6 does not consider the user terminal as an entity that is involved in the mobility signaling, and thus its

integration into mobility-aware group subscription is problematic. In the context of MEDIEVAL we plan to study not only receiver mobility but also the impact of sender mobility in a network-based localized mobility management architecture. Concretely, this area of the project intends to address:

- Mechanisms for mobility support of listener nodes in a non-relying way to bi-directional tunneling
- Topological correctness and transparency of source addressing
- Mechanisms for optimized multicast distribution tree updating

Additionally, another study item will be the coupling between the handover process, the change between layer-2 point of attachment and the actual group subscription. The strategy for the multicast distribution update depends on the envisioned service in MEDIEVAL, with two possible paths: specific to the source, or related to any source. Regarding this point, and in the context of localized mobility management, the multicast tree creation may interact with route optimization in a mobility point of view.

The use of cross layer information to better synchronize subscription information and actual point of attachment especially in case of predictive handover will be a key study point, particularly in the cases where the change between L2 point of attachment is not synchronized with the L3 change (i.e., homogeneous and heterogeneous handovers). Cross layer information can be used to better synchronize subscription information and actual point of attachment especially in case of predictive handover. Here, the application of IEEE802.21 mechanisms is a possible tool to ensure the feasibility of these processes. Also, IP multicast optimizations will be proposed both from the network mobility perspective (i.e., due to handovers) as well as from the service perspective (i.e., fast change of multicast groups, required by IPTV). MEDIEVAL will also benefit from the on-going efforts for the development of a PMIPv6-compliant protocol stack, performed by several partners of the MEDIEVAL project.

7 Conclusion

In this article we have presented the key points and challenges that the MEDIEVAL project will address aiming to deliver video services in an optimized way over wireless mobile access. The major architectural areas have been highlighted as being wireless access technologies, mobility, transport optimization and video services, which reflect the general work items of the project. We also have detailed key innovation points and research objectives for the areas of jumbo frames, MIH signaling extension and network localized mobility management with multicast support, which are important tools and mechanisms in the overall MEDIEVAL design. The work will start with the exploitation of the individual work items into the development of a cross-layer design that leverages the joint effort of each item, into an evolution of the Internet architecture for efficient video traffic support. The results achieved with the project will fall in a number of research subjects and will provide a set of extensive and measurable outputs, where possible solutions for this architecture will be evaluated and quantitatively assessed, particularly its impact to standardization

and the development of new video services. Lastly, the resulting architecture will be implemented in a demonstrator showcasing the developed functionalities. These results will be further disseminated in scientific fora, including leading conferences and journals in the field, as well as active pursuit of opportunities for standardization bodies influencing.

References

1. FP7 EU project: MultimEDia transport for mobIlE Video AppLications (MEDIEVAL), Grant Agreement no. 258053
2. Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, Cisco White Paper, http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.pdf
3. Lindquist, J., Maenpaa, J., Rajagopal, P., Marjou, X.: SIP/SDP Overlap with RTSP. IETF draft, draft-lindquistmmusic-sip-rtsp-00 (2009) (work in progress)
4. Suzuki, T., Tasaka, S.: Performance evaluation of integrated videoand data transmission with the IEEE 802.11 standard MAC protocol. In: Proc. IEEE GLOBECOM, vol. 1B, pp. 580–586 (1999)
5. 3GPP TR 25.913 V7.3.0: Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN) (2006)
6. IEEE 802.21 Standard, Local and Metropolitan Area Networks – Part 21: Media Independent Handover Services (January 2009)
7. Calhoun, P., Loughney, J., Guttman, E., Zorn, G., Arkko, J.: Diameter Base Protocol. RFC 3588 (September 2003)
8. ODTONE – Open Dot Twenty ONE (June 2010), http://hng.av.it.pt/projects/odtone
9. Majkowski, J., Palacio, F.C.: Dynamic TXOP configuration for QoS enhancement in IEEE 802.11e wireless LAN. In: International Conference on Software in Telecommunications and Computer Networks, pp. 66–70 (2006)
10. Borman, D., Deering, S., Hinden, R.: IPv6 Jumbograms. IETF RFC 2675 (August 1999)
11. IETF Multicast Mobility (MULTIMOB) WG: http://www.ietf.org/dyn/wg/charter/multimob-charter.html (visited in June 2010)
12. Gundavelli, S., Leung, K., Devarapalli, V., Chowdhurry, K., Patil, B.: Proxy Mobile IPv6. IETF RFC 5213 (August 2008)
13. Boggia, G., Camarda, P., Grieco, L., Mascolo, S.: Feedback-based control for providing real-time services with the 802.11e MAC. IEEE/ACM Trans. on Netw. 15(2), 323–333 (2007)
14. 3GPP TR R3.018: Evolved UTRA and UTRAN Radio Access Architecture and Interfaces, Release 7 (2007)
15. Piri, E., Pentikousis, K.: Towards a GNU/Linux IEEE 802.21 Implementation. In: IEEE International Conference on Communications, ICC 2009, June 14-18, pp. 1–5 (2009)
16. Shimakawa, M., Hole, D.P., Tobagi, F.A.: Video-conferencing and data traffic over an IEEE 802.11g WLAN using DCF and EDCA. In: Proceedings of IEEE International Conference on Communications 2005 (ICC 2005), vol. 2, pp. 1324–1330 (2005)
17. Suzuki, T., Tasaka, S.: Performance evaluation of integrated video and data transmission with the IEEE 802.11 standard MAC protocol. In: Proceedings of IEEE Global Telecommunications Conference 1999 (GLOBECOM 1999), vol. 1B, pp. 580–586 (1999)

18. Grieco, L., Boggia, G., Mascolo, S., Camarda, P.: A control theoretic approach for supporting quality of service in IEEE 802.11e WLANs with HCF. In: Proceedings of the 42nd IEEE Conference on Decision and Control, vol. 2, pp. 1586–1591 (December 2003)
19. Xiao, Y., Li, F.H., Li, B.: Bandwidth Sharing Schemes for Multimedia Traffic in the IEEE 802.11e Contention-Based WLANs. IEEE Transactions on Mobile Computing 6(7), 815–831 (2007)
20. Bucciol, P., Davini, G., Masala, E., Filippi, E., De Martin, J.: Cross-layer perceptual ARQ for H.264 video streaming over 802.11 wireless networks. In: Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM 2004), vol. 5, pp. 3027–3031 (November-December 2004)
21. van der Schaar, M., Krishnamachari, S., Choi, S., Xu, X.: Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs. IEEE Journal on Selected Areas in Communications 21(10), 1752–1763 (2003)
22. Zhang, Y., Foh, C.H., Cai, J.: An On-Off Queue Control Mechanism for Scalable Video Streaming over the IEEE 802.11e WLAN. In: IEEE International Conference on Communications (ICC 2008) (May 2008)
23. Haratcherev, L., Taal, J., Langendoen, K., Lagendijk, R., Sips, H.: Optimized video streaming over 802.11 by cross-layer signaling. Communications Magazine 44(1), 115–121 (2006)
24. Soliman, H. (ed.): Mobile IPv6 Support for Dual Stack Hosts and Routers. RFC 5555 (June 2009)
25. ICT-OPTIMIX project: `http://www.ict-optimix.eu/` (visited in June 2010)
26. PHOENIX: Jointly optimising multimedia transmissions in IP based wireless networks. Project, `http://www.ist-phoenix.org/` (visited in June 2010)
27. NAPA-WINE: Network-Aware P2P-TV Application over Wise Networks project, `http://www.napa-wine.eu/` (visited in June 2010)
28. ICT-CARMEN: CARrier MEsh Netorks project, `http://www.ict-carmen.eu/` (visited in June 2010)
29. Cunningham, G., Perry, P., Murphy, L.: Soft, vertical handover of streamed video. In: Proceedings of the IEE International 3G Mobile Communication Technologies, pp. 432–436 (2004)

EARTH: Paving the Way for Future Energy Efficient Broadband Wireless Networks

Luis Sanchez[1], Oliver Blume[2], Manuel Gonzalez[1],
Gergely Biczók[3], Dieter Ferling[2], and István Gódor[4]

[1] TTI (Technologies of Telecommunication and Information).
Albert Einstein 14, Santander, Spain
[2] Alcatel-Lucent Deutschland. Lorenzstraße 10, Stuttgart, Germany
[3] Budapest University of Technology and Economics. Műegyetm 3. Budapest, Hungary
[4] Ericsson Magyarország Kft., H-1097 Budapest, Könyves Kálmán krt. 11., Hungary
lsanchez@tlmat.unican.es, mjgonzalez@ttinorte.es,
{oliver.blume,dieter.ferling}@alcatel-lucent.com,
biczok@tmit.bme.hu,istvan.godor@ericsson.com

Abstract. Currently, the vast majority of mobile subscribers rely on second-generation mobile technologies, but service providers are investing into aggressive rollouts of mobile broadband networks to deliver a fully-featured wireless Internet. While the main focus in research has been put on enhancing the capacity of this kind of networks, very little has been done regarding their energy efficiency. On the other hand, rising energy cost and growing awareness of climate issues require a shift of focus. The EARTH (Energy Aware Radio and neTworking tecHnologies) project addresses this by investigating and proposing effective mechanisms to drastically reduce energy wastage and improve energy efficiency of mobile broadband communication systems, without compromising system capacity and users' perceived quality of service. In this paper we sketch the main research approaches taken within the project. First, the methodologies to evaluate the energy efficiency of cellular networks, as well as the respective energy efficiency metrics are presented. Afterwards, the proposed solutions are described; within EARTH a holistic approach is being used so that advances in radio components, radio network technologies and advanced network management protocols are exploited jointly, resulting in combined gains that enable an expected power consumption reduction by 50%.

1 Introduction

The planet Earth is experiencing global warming mainly due to the rising emission of greenhouse gases requiring immediate action to reduce carbon dioxide (CO_2) emission of all human activities. Even though the United Nations Climate Change Conference COP15 in Dec 2009 has failed to yield legally binding reduction plans many countries have committed to reducing their CO_2 footprint by 20% by 2020, compared to 1990 [1]. It is believed that the Information and Communication Technology (ICT) industry can play an important role in such reduction plans. Scientific findings have indicated that the CO_2 emission of the ICT industry is

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 326–339, 2011.

contributing a considerable percentage to the world energy consumption budget [2]. Within the ICT sector mobile communication is one of the fastest growing contributors with a CO_2 emission of 150 GTo in 2007 [3]. So, today and in the next 5-10 years there is a high demand for energy efficient improvements of ICT, driven by the socio-economic requirement of reducing climate change [4].

In the development of mobile systems, such as GSM and UMTS, operators and equipment manufacturers were focused on improving transmission capacity and spectral efficiency to offer higher data rates and better quality of service. Energy efficiency was not on the mainstream of research and it came as a side effect obtained by improved radio techniques and technological evolution. Now after it has become evident that these improvements do not compensate for the exponentially growing traffic and the related additional deployments, energy efficiency of networks becomes a research theme of its own. The cost of energy resources makes up for 20-30% of the OPEX cost, this has forced the major operators, such as Vodafone and Orange, to aim for reducing energy consumption by 20% to 50% [5] in order to counteract on their growing operational costs. Therefore, in future mobile systems, e.g. LTE and LTE-Advanced, energy consumption will also be a major issue from the point of cost of operation.

Recently, several research projects began to address energy efficiency improvements. For example, MobileVCE [6] and Opera-Net [7] are targeting aspects of energy efficiency. In order to tackle the associated research problems in an even more comprehensive manner and to achieve significant impact, the multi-disciplinary expertise of Europe's most successful and innovative companies, research and academic institutions is required. In January 2010 the highly ambitious 7th Framework Programme project EARTH (Energy Aware Radio and networking TecHnologies) [8] was launched, applying a holistic approach to investigate the energy efficiency of mobile communication systems. EARTH is committed to the development of a new generation of energy efficient equipment, components, deployment strategies and energy-aware network management solutions. EARTH is investigating the energy efficiency limit that is theoretically and practically achievable whilst providing high capacity and no unwanted QoS impact. The target of EARTH is to reduce the energy consumption of mobile systems by a factor of at least 50%. The project is primarily focused on mobile cellular systems of LTE, its evolution LTE-A, where potential impact on standardization is envisaged, but it will also consider 3G (UMTS/HSPA) technology for immediate impact. This paper sketches the technical approach of EARTH and results achieved in the first six month of the project duration.

As it is shown in Figure 1, the EARTH consortium involves 15 partners from 10 European countries in a European large scale Integrated Project (IP). The partners represent industry, operators, research institutions and universities, complementing the full range of knowledge and experience. The resources mobilised by the 15 participants of the EARTH consortium represent a total budget of 14.8 M€ with a total requested EC contribution of 9.5 M€ for the duration of 30 months. The especially strong industrial commitment in the project is expected to facilitate real-world results and fast commercialisation of the project's efforts. A strong contribution by academia assures a high level of innovation and bases the project on cutting edge research know-how.

Next to the environmental improvements, the substantial reduction of network energy consumption will yield large cost savings for mobile operators, substantially reducing the economical barrier to offer ubiquitous mobile broadband coverage. Hence, EARTH also enables the provisioning of high speed mobile services to citizens in countryside areas which are not yet reached by mobile broadband services.

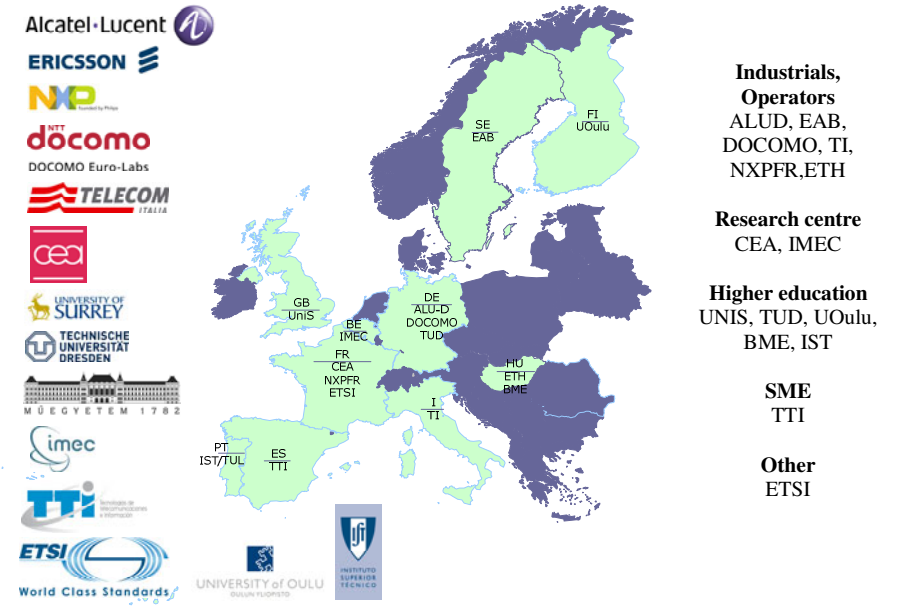


Fig. 1. EARTH Consortium

2 Research Strategy and Expected Impact

Instead of improving single components, the EARTH project [8] addresses the whole system with a holistic approach, from component level to network deployment and network management.

2.1 S&T Methodology

The technical work is organized around a structure with four groups of tracks as can be seen in Figure 2. In the first group of tracks, EARTH studies the global carbon footprint of wireless communications today and provides a forecast for the years 2010 to 2020. Furthermore, reference scenarios and an evaluation framework are being developed. These include the specification of meaningful “green” metrics and optional system extension scenarios in order to enable the proper assessment of energy efficiency for the project’s technical solutions. Based on these metrics and scenarios, the following three groups of tracks investigate theoretical limits, technology potentials and practical targets on different levels of Radio Communication networks.

In a second group of tracks, EARTH identifies and investigates mechanisms and solutions by reconsidering the mobile cellular networks architecture and deployment. The work places special emphasis on cooperative base stations, multi-hop extensions and heterogeneous deployments including relays and small cells.

The third group of tracks studies how network management and radio resource management can be enhanced for energy aware operation and dynamic adaptivity to variable traffic loads. The behaviour of networks and network components in low loaded situation, e.g. during night times, has been neglected in the past. Therefore high potential for large saving is expected in these situations.

The fourth group of tracks is devoted to energy efficient components of radio base stations and to enhancements of transmission techniques.

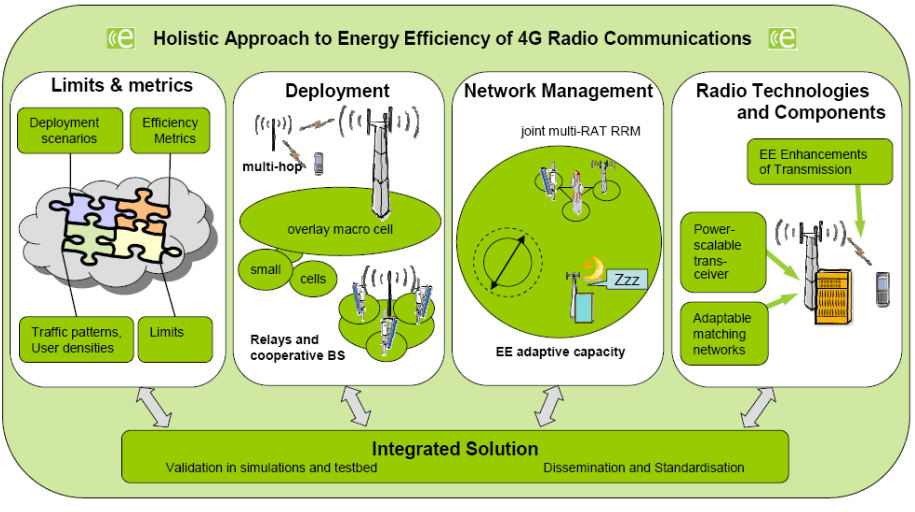


Fig. 2. EARTH holistic approach to Energy Efficiency of 4G Radio Communications

Developing early proof of concepts of sub-systems will be a key factor for the success of the proposed solutions. EARTH validates the investigated concepts by implementing selected solutions in a test platform provided by Telecom Italia's testlab. Other approaches are verified in system simulations. With the experiences gained, the individual solutions of the different tracks will be consolidated and aligned into an integrated solution. The evaluation framework and the defined metrics will be applied to calculate the energy efficiency improvement achieved by EARTH.

The scientific results of the innovative energy efficiency solutions will be widely disseminated. Large parts of the solution will require submission of amendments to standards or of new standards in order to ensure multi-vendor interoperability and wide applicability.

2.2 Estimation of the Ecological Impact of Mobile Communications

Several estimations of the energy consumption and the corresponding CO_2 footprint of ICT have been published [9] [3], the SMART 2020 report probably being the most prominent one. The EARTH project has undertaken to revisit the ecological footprint of mobile communications to predict the impact on the global scale and what mitigation the EARTH project can provide for reducing this footprint.

The study within EARTH is based on comprehensive data of the consortium on deployed base station sites of cellular networks, broadband subscriptions, types of mobile devices and the amount of traffic they generate. The study takes into account the radio access network, core network data centres and mobile devices. For each of these, energy consumed during manufacture and operation is studied separately in a full life cycle analysis. These data are extrapolated to the year 2020, regarding several scenarios: (i) efficiency improvements are leveraged for capacity gains rather than energy savings, (ii) business-as-usual with an observed annual 8% reduction of energy consumption mainly based on Moore's law, and (iii) improvements foreseen from EARTH. Finally (iv), also the usage of renewable energy is studied for a reduction of CO_2 emissions.

The study yields an overall carbon footprint for 2007, that is significantly below the one estimated in SMART 2020 study of 150 Mto CO_2-e. The difference is mostly due to updated electricity consumption data of base stations. The projection to 2020 [10] suggests that the overall carbon footprint of mobile communications will almost triple between 2007 and 2020 if no additional means for reduction are taken. By 2020 the contribution of mobile device manufacturing will catch up with the CO_2 footprint of network operation, driven by the increasing use of smartphones and laptops for mobile access. This prediction is slightly higher than that given by the SMART 2020 study, which did not consider smartphones and laptops. Extrapolations of the Cisco Visual Networking Index suggest that mobile traffic volume might rise by a factor of 100 to 150. In spite of that, the overall RAN energy consumption can be kept flat by realizing both component and system-level energy efficiency improvements that the project EARTH strives to provide [11].

3 Technical Approaches Undertaken

The EARTH project is committed to study the energy saving potential of a broad range of solutions. Following all of them would risk diluting resources and missing the maximum impact on power consumption of future LTE and LTE-advanced systems. Therefore EARTH is devoting the first year of the project to analyse the theoretical gains and limits of all tracks. Leveraging the metrics and the evaluation framework EARTH that are being defined in parallel to that, EARTH will then select the most promising tracks to pursue in more detail in the remainder of the project. At the time of writing this position paper, the project is on half way to the selection process, and this paper can just give a first status of the modelling of the different tracks.

3.1 Reference System, Reference Scenarios and Evaluation Framework

For analysing the potential gains of the different technology tracks studied in EARTH, a baseline system (BLS) and reference scenarios (RSc) are an indispensible prerequisite. In the first months of the project such a common reference system has been defined, building a cornerstone for the EARTH project. To capture the state-of-the-art, the specified parameters and settings are based on 3GPP LTE standards and best-practice of system simulations. Further, the reference system definition will serve for the configuration of the testbed and for the final assessment of achieved gains at the end of the project. The Reference Scenarios (RSc) are technology agnostic use cases chosen to represent relevant cases for energy efficiency evaluation with certain deployment assumptions, channel models and geographical, temporal and service specific traffic models.

The Optional System Extensions (OSE) broadens the baseline system by innovations, spanning the energy efficiency solution space envisaged in the EARTH project, e.g.

- Traffic aware power amplifier concepts
- Remote radio heads
- Relaying concepts
- Heterogeneous networks and small cells
- Coordinated multipoint transmission
- Multiple combined radio access technologies (multi-RAT)

Figure 3 shows a graphical representation of the EARTH reference concept.

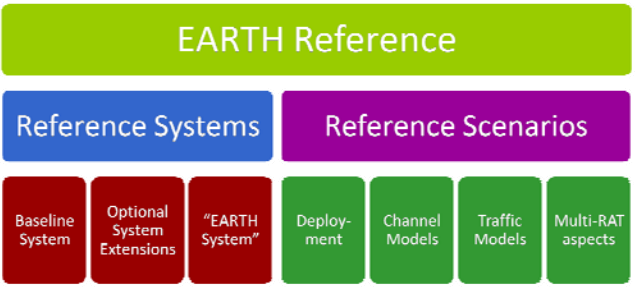


Fig. 3. Graphical representation of reference system and reference scenarios

The specification of the baseline system and of the system with optional extensions are organised in a hierarchical manner, from component level to system level, as shown in Figure 4. This provides a toolbox for studying single tracks or combined effects of the most promising tracks selected in EARTH.

3.2 Energy Efficiency Metrics and Evaluation Framework

Today, there exists no widely agreed methodology to evaluate the energy efficiency of cellular networks. Similarly, meaningful energy efficiency metrics are currently lacking. But both are a prerequisite to determine the energy saving potential of concepts or technologies as developed in EARTH.

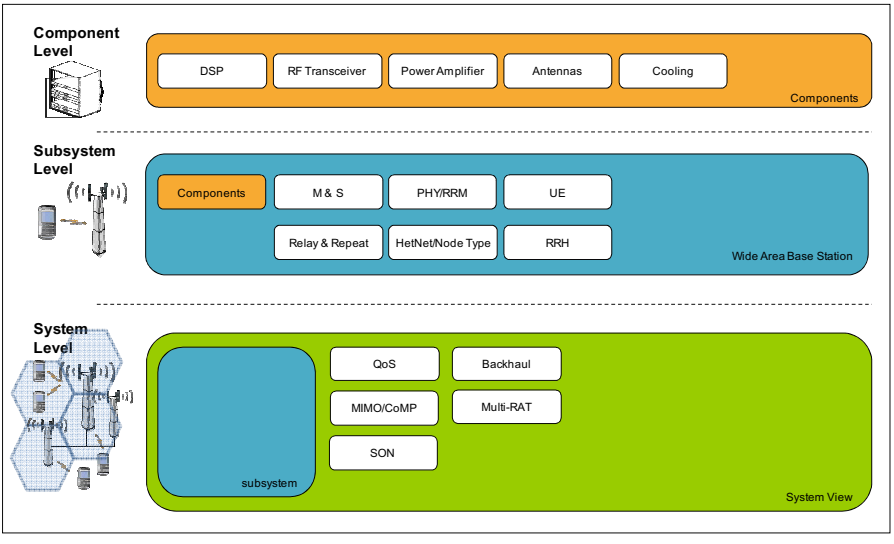


Fig. 4. Hierarchical structure of the Optional System Extensions

The evaluation of energy efficiency on the system level requires combining contributions and interaction of components, deployment, traffic scenarios and system management. This cannot be achieved in a complete system simulation, but requires detailed simulations on radio link level, modelling of energy consumption of radio components, stepwise abstraction of such models into a system level with weighted contributions according to the reference system or its extension.

Such a methodology will be provided by EARTH for a fair comparison of different concepts and technologies. Figure 5 shows a sketch of the evaluation framework that is currently under construction in EARTH. Each step is fed by parameterized models, from component level up to system level. For example, a simplified energy aware scheduler model defines the resource utilisation pattern of the LTE radio frames. Combined with the DC power consumption model of the transceivers (TRX) this yields the TRX power consumption. For a given base station and deployment model the system power can be computed from this for the covered area of the system.

Now, the final step of assessing energy efficiency is the application of a metric to the energy consumption. Simple metrics just sum the power consumption of all components. But obviously the lowest energy consumption is not the best mode of operation, when the quality of service is jeopardized or the coverage is patchy. Proper energy efficiency metric has to capture the amount of service provided by the energy spent, i.e. rating the consumed energy in relation to the transported data, to the covered area or to QoS parameters like call drops or data delay.

Another aspect of the metric is whether it regards the instantaneous power P (i.e. energy consumption on a microsecond scale) or the power averaged over longer times (e.g. a radio frame of 10msec, an hour or a week), and the instantaneous or averaged data rate R served by this power. Note that mathematically it is not the same to compute the average of the ratio of power and data rate compared to the ratio of the average values (Eq. 1). Usually a network operator will not care about instantaneous

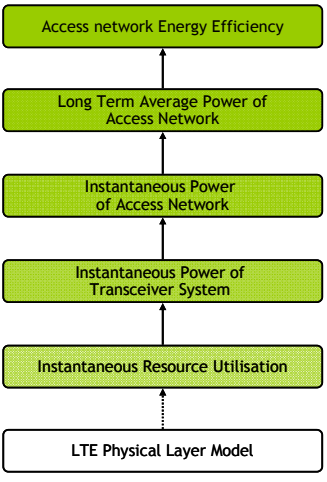


Fig. 5. Stepwise abstraction of system power consumption in the evaluation framework

values, but about the monthly energy bill and the user data throughput per month. Also it may be very difficult to record the instantaneous values in a real-world system.

$$\varepsilon = \text{Average}(R) / \text{Average}(P) <> \text{Average}(R/P) \quad (1)$$

In EARTH the requirements for a suitable metric are being collected and a set of metrics have been identified as candidates for a meaningful system metric. Different key performance indicators (KPI) like data throughput or coverage can be regarded or additional boundary conditions (e.g. delay or acceptable blocking rate of services) are applied. One example metric is given below (Eq. 2)., aggregating the system data rate R(t) as key performance parameter and relates it to the aggregated consumed system power P(t) over a time interval T.

$$\varepsilon_T = \frac{\int_0^T R(t)dt}{\int_0^T P(t)dt} \left[\frac{bit}{J}\right] \quad (2)$$

4 Potentials of EARTH Technical Solutions

The technology potential of some key technical approaches of the EARTH project will be briefly described in the following, grouped according to the work packages named “Green networks” and “Green Radios”. Detailed studies to quantise the achievable energy savings have been started and are ongoing at the time of writing.

4.1 Green Networks Technological Potential

Temporal and spatial traffic variations greatly influence the achievable energy savings. A method which allows extracting traffic variations from cell-specific measurements in the form of spatial distributions will be used to transform operator’s

measurements into a surface map of spatial densities, so that the information can be treated separately from the underlying network layout. Then, network scenarios with different layouts of cells can be compared. An energy-aware network management should match the number and locations of active radio resources to the temporal and spatial variations of traffic demand.

Within the EARTH project several solutions have been studied to enable the reduction of energy consumption at the network level of mobile cellular networks by adapting the resources offered to the users' demands. They are grouped in the following three main areas that encompass the three different time scales used when reasoning over radio networks: network deployment strategies, network management algorithms and radio resource management (RRM). While network deployment is designed for long-term measured in years, network management acts on day-to-hour-basis and radio resource management is almost real-time.

For each of these areas, different solutions have been proposed, and their theoretical potentials have been analysed. An important aspect that will be analysed is the possible combination of solutions from different areas, for example, designing a deployment strategy taking into account the possible management mechanisms that could be implemented on top of it.

4.1.1 Deployment Strategies

Energy efficiency of deployment strategies featuring the following techniques is being analysed:

- Variable cell size and cell mixing: The trade-off between required transmission power and the number of cells to cover an area is studied to derive the optimal strategy for the deployment of macro and micro sites in heterogeneous systems for different load conditions. Results from the studies performed so far indicate that significant gains (up to 70%) can be obtained in terms of area power consumption (W/Km^2) if micro-sites are combined with traditional macro deployment [12]. This solution has also been analysed under the condition of indoor-outdoor scenarios.
- Multi-hop relay: As part of the 3GPP study item LTE-Advanced relays are being studied as a technology to extend coverage and increase capacity [13]. By means of analytical studies, the energy efficiency fundamental limits for Decode and Forward, Amplify and Forward and Compress and Forward relaying schemes have been derived. Additionally, a relaying scenario has been modelled as a non-regenerative cooperative Multiple-Input and Multiple-Output (MIMO) communication system and its energy efficiency in bits-per-joule has been compared. Preliminary results show that for some relevant scenarios relaying strategies do not only enable higher spectral efficiency but also optimize the energy consumption per transmitted bit.
- Base station (BS) cooperation: Cooperative transmission of neighbouring BS like Network-MIMO and Coordinated Multi-Point Communication (CoMP) can enable interference mitigation and thus reduction of transmit power and of retransmissions, at the cost of higher traffic on the backhaul. An energy efficiency analysis of an idealized CoMP system have shown that it can be energy efficient in asymmetric propagation conditions and that the backhauling and cooperative processing power should be kept low for CoMP to provide any gain.

- Multi-RAT deployment: 3GPP networks comprise a range of radio access technologies (RATs), e.g. GSM, WCDMA and LTE. These RATs have different capacities and different ability to provide bearers with QoS profiles of certain services. A multi-RAT strategy can apply load balancing and separation of traffic types to maximise the overall energy efficiency.

4.1.2 Network Management

A first assessment of self-organizing networks (SON) mechanisms for autonomous selection of energy efficient network operation mode has also been carried out:

- Network management mechanisms with ON-OFF schemes are focusing on how to reduce the number of serving base stations according to the current needs of the network. A BSs can be switched off when the load condition is low and other BSs can serve the user demand. The viability of turning off BSs relates to the density of the BS that is required to cope with the capacity demands. In peak-traffic demand the cell sizes are deliberately reduced since the limiting factor is the BS capacity and not its power budget. Results have shown that simple management schemes might achieve, theoretically, significant energy savings (up to 20 %).
- Multi-RAT coordination and cooperation also aims to reduce the number of active BS adaptively to the actual traffic demands of the network. Assuming co-located sites equipped with multiple radio access technologies, the analysis performed in an urban macro environment has shown the real potential of multi-RAT management to enable more energy efficient operation through dynamic adaptation to the actual traffic demands of the network.
- Cooperative BSs can be used, e.g., to multicast traffic from a single source to a set of destinations through a set of BSs. Numerical results show the impact of average channel conditions on how to select the interim BSs between the source and destinations.

4.1.3 Radio Resource Management

Finally, RRM algorithms are being studied from an energy efficiency point of view.

- In order to investigate how the elasticity of the traffic and its characteristics can be utilized, the interaction of call admission control and RRM is being analyzed.
- In multi-RAT environment, it is interesting how to dimension the networks with system wide control of scheduling and call admission control in a more energy efficient manner.
- It is also an open question how neighbouring cells should cooperate with other to reduce their energy consumption. There are ongoing studies on joint scheduling and power control of multiple cells.

4.2 Green Radios Technology Potential

In today's mobile networks, the base stations are responsible for the major part of energy consumption in radio access networks, simply due to the large number of deployed BSs for guaranteeing the QoS and coverage. With the deployment of LTE, in coexistence with GSM and UMTS, the increasing number of BSs will cause an increase in energy consumption [8].

The EARTH project breaks down the consumption of base stations to its components and analyses the technology potential in energy savings, not only optimizing each block, but also providing a holistic approach. Three main fields of techniques and solutions have been identified to improve the energy efficiency of LTE with the low impact on overall performance:

- Energy Efficiency in Antenna Techniques
- Energy Efficiency in Radio Components
- Energy Efficiency in Radio Interface and Protocol Based Techniques

4.2.1 Energy Efficiency in Antenna Techniques

By using multi-antenna techniques such as spatial multiplexing and beamforming it should be possible to reduce the necessary transmit power while still maintaining QoS. The current activities are mainly focused on:

- New antenna topologies and the use of new materials are investigated from the energy efficiency perspective.
- Reconfigurable antennas and active antenna systems are considered for adapting to the traffic demand. On one hand, static parameters of an antenna, like beam width and pointing direction, can be tuned according to current traffic distribution, so that link budget is optimized. On the other hand, active antenna systems implement beamforming with maximum flexibility and energy savings related to space diversity and multiplexing techniques, as well as reduction of feeder losses (with Remote Radio Head, RRH).
- Smart antenna technologies, like MIMO, are also investigated. Multi-antenna operation and MIMO have been adopted by LTE to increase the coverage and capacity. However, the different operations in downlink (transmit diversity, spatial diversity with/without precoding, multiuser MIMO,…) [18] can strongly affect BS energy efficiency. Therefore, the best configuration of MIMO can be selected if an accurate model of performance and the impact on consumption of each MIMO mode is available.

4.2.2 Energy Efficiency in RF Components

The main BS blocks considered in the project are DSPs for baseband (BB) signal processing and control, Small Signal RF transceiver (SSTRX) and Power Amplifiers (PA). The energy consumption of each block presents different breakdowns in function of the cell size. For instance, in a macro cell BS, the power amplifier (including its cooling) consumes around 65% of the total consumption. However, in a pico cell BS, the BB/DSP and SSTRX dominate the energy consumption and use up to 70% of the BS power. Therefore, the investigations of the RF transceiver field puts special emphasis on the search of new solutions and approaches in the PA for macro BSs and in the BB/DSP and SSTRX for pico-BSs, with parallel research in the rest of the blocks.

- For PA, two concepts for Adaptive Energy Efficient Power Amplifiers are proposed to meet the required high energy efficiency performance [15]. Both are based on advanced amplifier design (transistor technologies and architectures) for allowing the adaptability to traffic statistics in LTE. Adaptive PAs leads to reduced

power consumption for low and medium signal levels by adjusting the operating point to the signal level. Fast component deactivation during time slots with no transmitted signals avoids further power wastage.

- For BB/DSP and SSTRX in pico-cell context, the challenge is to use the best suited components and to enable dynamic power management and voltage scaling, as well as to implement algorithms and signal processing in energy efficient ways by means of high performance processors (ASIC, ASIP, FPGA).

4.2.3 Energy Efficiency in Radio Interface and Protocol Based Techniques

The LTE standard unlike many other radio access technologies does not require continues transmission of reference signals (or pilots) allowing periods where the system does not need to transmit. These can efficiently be used for reduction of system energy consumption. Another key feature of LTE are multi-antenna techniques which also need to be addressed in terms or energy efficient network operation.

- The discontinuous transmission (DTX) and sleep mode techniques adapt the operation of PA to the required transmission periods. DTX and sleep modes can be classified depending on their time scales. *Micro DTX* is working on a fraction of a subframe (<1 ms), *short DTX* is designed to adapt to the radio frame structure (<10ms), and *long DTX* operates for longer than one radio frame. Finally more traditional *sleep modes* are designed for periods >10ms.
- The technique of adaptability to system dynamics takes in advantage all the flexibility of the LTE standard to use the best configuration in each moment. This means that Adaptive Modulation and Coding (AMC), MIMO pre-coding and rank adaptation algorithms are evaluated for the energy saving potential of re-assigning the transmission mode as a function of scenario.
- Retransmission schemes and HARQ protocols are considered to find optimal scheduler algorithms that enhance the usage of the PA, reducing the power consumption.

5 Conclusions and Future Developments

For the deployment of future high speed mobile communication networks the system power consumption and the CO_2 footprint are critical, both from environmental point of view and for the operational cost of mobile operators. In the framework of EC FP7, the EARTH project is set to study energy saving measures, to assess their limits and potentials and to provide an integrated solution with at least 50% of energy savings compared to today's baseline system.

The project was started in January 2010 and has already delivered a consistent estimation of the global power consumption of mobile networks up to 2020. It has achieved the definition of a baseline system and scenarios for the evaluation of energy efficiency improvements. Currently EARTH is working on the specification of meaningful metrics and an evaluation framework.

EARTH aims for improvements beyond current development trends. Those improvements will be yielded by energy efficient deployment strategies including heterogeneous networks, relays and cooperative base stations; energy aware network

re-configuration and resource management; and transceivers or base station equipment with high adaptability to the traffic situation. Further improvements will be realized by employing advanced radio transmission techniques for energy efficiency instead of purely for spectral efficiency improvements and by energy efficiency enabling enhancements of radio interfaces. This position paper describes the main tracks and first results of the project.

Such techniques for improved energy efficiency of mobile broadband communications will only be deployed if they do not unduly impact on the system performance and the user's perceived "quality of service". The challenge of EARTH is thus to analyse the trade-offs between performance and energy efficiency. In this sense, metrics and figures of merit that account for both the energy efficiency and the capacity are studied during this first phase of the project. An evaluation framework that combines contributions and interactions from component to system level is being developed to enable a fair comparison of different concepts and technologies and to judge the gains of each standalone solution on the integrated system.

The results of EARTH will be widely disseminated and (where required) taken to the appropriate standardisation bodies. Several papers have already been presented [11][12][15] or accepted for publication [16][17]. Due to the industry driven consortium it is expected that the results will be quickly leveraged for efficient products, providing operators with sustainable equipment and operation of their future mobile broadband networks.

Acknowledgements

The authors gratefully acknowledge the contribution of the EARTH consortium.

The work leading to this paper has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement n° 247733 – project EARTH. This publication made by participants of the EARTH project reflects only the author's views and the European Union is not liable for any use that may be made of the information.

References

1. Climate change: Commission welcomes final adoption of Europe's climate and energy package. Press release IP/08/1998 of the European Commission (December 2008), http://europa.eu/rapid/pressReleasesAction.do?reference=IP/08/1998
2. Long Term Evolution (LTE): an introduction, Ericsson White Paper (October 2007)
3. SMART 2020: Enabling the low carbon economy in the information age, The Climate Group, GeSI, http://www.smart2020.org/
4. EU Commissioner Calls on ICT Industry to Reduce Its Carbon Footprint by 20% as Early as 2015, Press release of the European Commission, MEMO/09/140 (March 2009)
5. Vodafone Corporate Responsibility Report (2008)
6. Green Radio – Sustainable Wireless Networks, Mobile VCE (February 2009), http://www.mobilevce.com/dloads-publ/mtg284Item_1503.ppt

7. Optimising Power Efficiency in mobile RAdio Networks (OPERA-Net), EUREKA CELTIC project, http://opera-net.org/default.aspx
8. Energy Aware Radio and neTwork tecHnology (EARTH), EC FP7 project INFSO-ICT-247733 (January 2010-June 2012), http://www.ict-earth.eu
9. Fettweis, G., Zimmermann, E.: ICT Energy Consumption – Trends and Challenges. In: Proc. of WPMC 2008 (2008)
10. Biczók, G., Malmodin, J., Fehske, A.: Economical and Ecological Impact of ICT. Public Deliverable D2.1 of EARTH (June 2010), To be released on http://www.ict-earth.eu in Q4 2010
11. Auer, G., Gódor, I., Hévizi, L., Imran, M., Malmodin, J., Fazekas, P., Biczók, G., Zeller, D., Blume, O., Tafazolli, R.: The EARTH Project: Towards Energy Efficient Wireless Networks. In: Future Networks & Mobile Summit 2010, Florence (June 2010)
12. Arnold, O., Richter, F., Fettweis, G., Blume, O.: Power Consumption Modeling of Different Base Station Types in Heterogeneous Cellular Networks. In: Future Networks & Mobile Summit 2010, Florence (June 2010)
13. Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9), 3GPP TR 36.814 V9.0.0 (2010-03) (March 2010)
14. Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation (Release 9), 3GPP TR 36.211 V9.1.0 (March 2010)
15. Ferling, D., Bohn, T., Zeller, D., Frenger, P., Gódor, I., Jading, Y., Tomaselli, W.: Energy Efficiency Approaches for Radio Nodes. In: Future Networks & Mobile Summit 2010, Florence (June 2010)
16. Fehske, A., Malmodin, J., Biczók, G., Fettweis, G.: The Global Footprint of Mobile Communications – The Ecological and Economic Perspective. To appear in IEEE Communications Magazine, issue on Green Communications (November 2010)
17. Correia, L.M., Zeller, D., Blume, O., Jading, Y., Auer, G., Van der Perre, L., Ferling, D., Gódor, I.: Challenges and Enabling Technologies for Energy Aware Mobile Radio Networks. To appear in IEEE Communications Magazine, issue on Green Communications (November 2010)
18. LTE and the Evolution to 4G Wireless: Desing and Measurements Challenges, Agilent Technologies

A New Perspective on Mobility Management Scenarios and Approaches

Tiago Condeixa[1], Ricardo Matos[1], Alfredo Matos[1], Susana Sargento[1], and Rute Sofia[2]

[1] Instituto de Telecomunicacões, University of Aveiro, Portugal
{tscondeixa,ricardo.matos,alfredo.matos,susana}@ua.pt
[2] IAN, UTM, INESC Porto, Porto, Portugal
rsofia@inescporto.pt

Abstract. Currently and in the future, users demand for ubiquitous network connection to interact with the world. Moreover, the mobile devices are increasingly widespread and are better equipped in terms of access connections and services support. Mobility management is therefore an intrinsic feature of mobile networks; however, it is not yet ready to support the so called User-provided networks (UPNs), where the network elements can be devices controlled by regular users, or providers, and share subscribed access. In these scenarios, mobility has to be rethought to consider user-centric approaches. This paper discusses the efficiency and applicability of current mobility assumptions in user-centric scenarios, discussing their requirements and solutions addressing several types of networks that may exhibit user-centric characteristics. It also identifies the fundamentals of a user-centric mobility management architecture able to efficiently deal with the dynamicity of the aforementioned scenarios.

Keywords: mobility management, user-centric, scenarios, requirements, assumptions, challenges.

1 Introduction

Today, ubiquitous access is not only a commodity for Internet users, but it is also essential to interact with the world. Moreover, mobile devices are increasingly widespread and are better equipped in terms of access technologies and service support. Moreover, mobility is an essential key feature which holds specific network requirements. For instance, any movement across different network segments should be transparent to the user, so that the network should be able to maintain service continuity independently of location and access media, supporting the heterogeneity of today's networks and technologies.

Mobility management is today an intrinsic feature of mobile networks. However, it is still a feature left aside in what concerns the most recent trends in wireless networking, namely, in user-centric environments. In these environments, the social behavior inherent to humans is also impacting the way network access is perceived;

K. Pentikousis et al. (Eds.): MONAMI 2010, LNICST 68, pp. 340–353, 2011.

they consider that users develop spontaneous wireless networks simply based upon cooperation and access sharing on particular communities. Such user-centric architectures bring in several challenges to the traditional and tightly controlled mobility management schemes. First, in user-centric scenarios, the network elements are usually devices either carried or controlled by regular Internet users or providers. Second, users share subscribed Internet access. Third, users (and hence, the devices they own) tend to be mobile, with large dynamicity. Fourth, in these networks, some access control features may be transferred by the provider to the control of the user (as it is the case of femtocells).

Current mobility management solutions are not optimized for the previously described aspects, mainly personalization and dynamicity. It is therefore of major importance to re-think mobility management from an out-of-the-box perspective, and in particular, to consider user-centric approaches and how these can assist not only the individual user but also the provider in terms of mobility management coupled to the day-to-day living of Internet users. Such re-thinking has to consider trends on personalization and dynamicity: it is required to pro- vide a mobility process applied to distinct users with different mobility patterns, and also to different services and its characteristics, and forming different social networks.

As a first step towards a better re-thinking, this paper identifies and describes main user-centric scenarios, their assumptions, and mobility management requirements. We aim at assessing the suitability of current mobility models when applied to the described scenarios, and we show why a new approach and paradigm for mobility is required. For the sake of simplicity, one part of the discussion is divided in three blocks that we consider essential from a mobility management perspective: i) binding definition, what is the binding information and its initial discovery; ii) binding maintenance, how to maintain the translation/mapping update and at which cost; and iii) forwarding data problem, the required data plane techniques to keep up with the mutating control plane. We identify the main problems derived from the integration of these three steps in user-centric scenarios and propose some ideas to improve it. We analyze the cur rent problem of the use of IP address for both identification and location in the scope of mobility management and how it interferes with personal mobility, services mobility and multihoming. We further discuss initial ideas on requirements and characteristics of a mobility management architecture aiming at decentralizing the global management in user-centric scenarios, and exploiting how the context of users, networks and services can assist in optimizing mobility management. Finally, we analyze the distribution of mobility control points in the network according to different models and how they could be reallocated, based on adaptable principals that react to network changes.

This paper is organized as follows. Section 2 covers related work, also explaining our contribution in regards to previous work. Section 3 describes a set of user-centric scenarios, their mobility management assumptions and requirements, and their integration with current mobility solutions. Section 4 discusses challenges that we have identified and potential solutions to the identified gaps. The paper concludes with a summary and future work in section 5.

2 Related Work

To understand potential mobility patterns related to spontaneous environments, Huang et al. addressed the specific scenario of incident scenes, such as disaster networks and the need to provide connectivity on such environments [1]. The authors provided an analytical categorization of parameters that are related to mobility in such environments, and provide a set of recommendations to follow in regards to mobility in self-organizing networks. A study on an Urban landscape based on Google WiFi mesh network [2] provides a good basis for the analysis of wireless usage. In particular, the authors show that such usage is splitted into three classes mostly based on user devices, namely, traditional mobile computers (notebooks), APs, and PDA-like smart-phones. The authors also show that the urban mobility patterns exhibit the property of geographic locality. Specifically regarding accounting of mobile users in wireless environments, a solution considers the application of agents that track node mobility, the Mobile Agent (MA) middleware [3]. Such solution is based on having agents sent on demand to administer nodes. The central block works on the control plane only, in contrast to centralized mobility management solutions of today. A few proposals [4] [5] have considered the application of overlays to deal with mobility from a global perspective. This gives the means to consider mobility management from a distributed perspective, where the mobility anchor point may be placed within the user premises. However, these solutions do not consider de-centralization nor decoupling of mobility functionality. A proposal for a spontaneous environment mobility architecture based on the definition of more adequate addressing schemes, and hence, of more adequate routing [6], combines the notions of geographical routing based on ballistic trajectories with a location service based on Distributed Hash Tables (DHT) to achieve seamless mobility management in a k-neighborhood. Mobility management is based on the definition of an identifier that identifies the node on its constructed pseudo-geographical space and which associates the node with a k-neighborhood, thus providing an identifier to its mesh area.

3 User-Centric Scenarios

Broadband Internet access is in its majority complemented by wireless technologies in the last hop. Such wireless deployment, added to the low-cost and open-source firmware available, lead to a paradigm change in the user role in terms of networking architectures: the user today can contribute and assist in increasing the reach and support of the Internet broadband access. Hence, such networking scenarios correspond to user-centric scenarios in the sense that the user is capable of controlling its own wireless devices, which become part of the network.

The main aspects of user-centric scenarios relevant from a mobility management perspective are the following: high mobility frequency (users adhere to such wireless infrastructures mostly to be able to move freely across additional spots); nodes in the network that provide access to other nodes may change frequently; users share connectivity (share Internet subscription). To better analyse the implications of these aspects in terms of mobility management, we describe examples of scenarios with

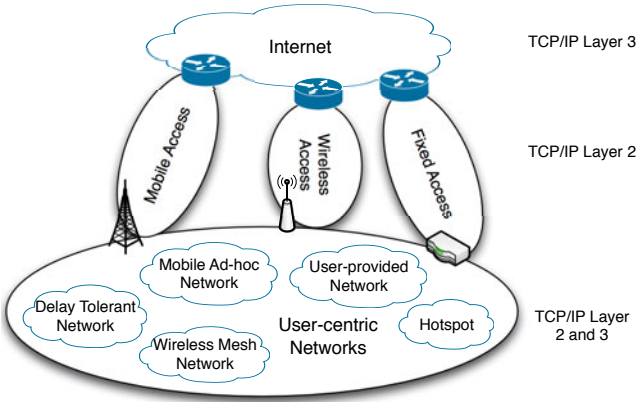


Fig. 1. Current Network Scheme

user-centric characteristics, as illustrated in Figure 1. We have considered five different architectures and we will analyze their main characteristics, common aspects and potential gaps to be filled from a mobility management perspective.

To support the comparison analysis, we provide Table 1. This table contains a set of main features that, jointly, characterize mobility management functionality: identification and user profile database, access control/authorization, service portability, resource management adaptation, and the potential need for local and global handover optimization.

Table 1. Systematization of the scenarios

Features/Scenario	Hotspot	WMN	MANET	UPN	DTN
Identification	Global user credentials, IP address	IP address (1 interface)	IP address (1 interface)	Community credentials, depends on local trust management system	Community credentials; may be provided spontaneously; local identification may not exist
User Database	Provider (access or service)	Community, provider	Provider (access or service)	Community, potentially distributed across several locations	Inexistent
Access control/ authorization	Centralized, provider	Distributed across a set of specific nodes	Distributed across a set of gateway nodes	Distributed and spontaneous	Decentralized
Mobility anchor point location	Edge node or Service Provider	Static gateway nodes	Static gateway nodes	Edge Node, Service Provider, Micro Provider	a few nodes on site, e.g. due to higher levels of residual energy
Portability	Local or dependent on 1 provider	Within a community (most likely tied to 1 provider)	Within a community (most likely tied to 1 provider)	Within a community (sometimes tied to a Virtual Operator)	Within a local area
Resource management adaptation	inexistent	inexistent	manual(nodes)	automatic	inexistent
Intra-handover frequency	low	medium	medium	high	high
Inter-handover frequency	low	low	medium	medium	low

3.1 Hotspot

The Hotspot is today the most popular architecture available with user-centric capabilities. Its main purpose is to expand and to complement current Internet broadband access. Wireless hotspots abound around us in residential households and public establishments (e.g. universities, shopping centers, hospitals, hotels). A Hotspot is composed by at least one wireless Access Point (AP) connected to the Internet through an Access Router (AR), being these devices usually provided by a Network Access Provider or an Internet Service Provider. Usually, the AR and AP are co-located (e.g. as part of the Residential Gateway on a household). In public Hotspots several APs may be connected to the same AR to provide a wider coverage area. Although in the hotspot scenario there is no connectivity relaying from one user to the other, the AP media is shared.

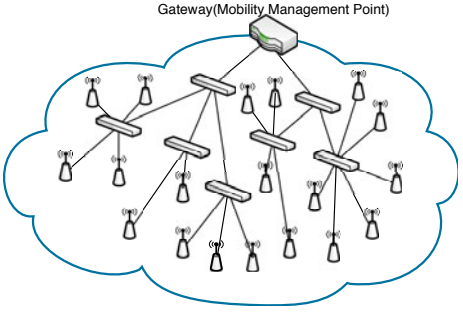


Fig. 2. Hotspot Scenario

The hotspot is controlled and owned by a provider. This is the case for our residential household, where we have an AP placed and controlled by our provider. Given that its purpose is to expand capillarity, hotspot users normally stick around and hence are said to visit a few preferred locations mobility is limited in scope (cf. first column of Table 1). Node movement speed is low (e.g. users walking on a house or to a coffee-shop), and connectivity while users move is intermittent.

Within a specific hotspot, movement is usually taken care of by the MAC layer. Moreover, user identification and authentication is provided by regular means based on MAC and IP identification, and controlled by the provider the user or a hotspot (for the case of pre-paid access) subscribes to. Hence, a hotspot is centralized, from an access control perspective. Service portability is an aspect that is only considered for hotspots belonging to the same provider, and resource management is usually inexistent, being the service provided on a best-effort basis. Consequently, the need to perform handovers normally relates to inter-hotspot scenarios, where users arrive or leave a hotspot. When inter-hotspot handovers occur, they are dealt with from an OSI Layer 3 perspective.

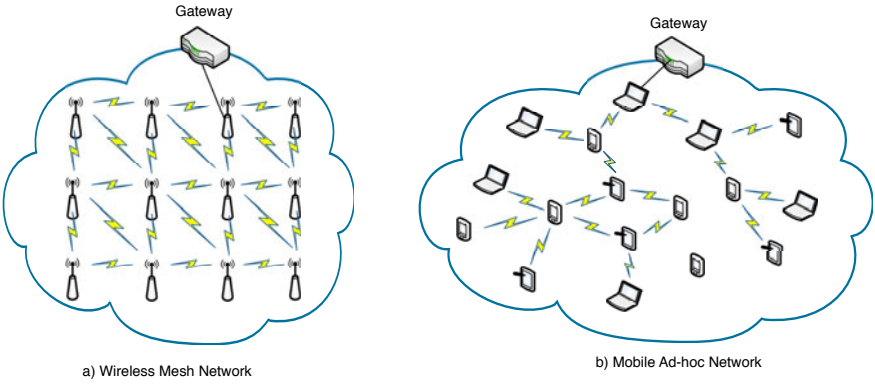


Fig. 3. MANET and WMN Scenarios

3.2 WMN and MANET Scenarios

Wireless Mesh Networks (WMN) and Mobile Ad-hoc Network (MANET) scenarios, which are respectively covered by the second and third columns of Table 1, are both sub-cases of ad-hoc networks. A main difference to highlight between WMNs and MANETs is that, while in WMNs all nodes are static, in MANETs all nodes may move, even though today most of the gateways are static. Both these scenarios have intrinsic characteristics that must be considered when addressing mobility management.

Similarly to the hotspot scenario, these scenarios are also applied to provide connectivity expansion in an autonomous way. However, it should be noticed that for this specific case (and in contrast to the hotspot model), the provider has no control on the MANET growth/operation. Hence, key aspects to take into consideration in this model are the clear split of network management functionality between the access and the MANET (Customer Premises) region. Moreover, there is no centralized control and hence, ad-hoc transmission brings both benefits and disadvantages (e.g. overhearing). Central to the deployment of this scenario is the need to consider a dynamic routing (multi-hop) protocol, to ensure reliable transmission across several hops.

The gateway nodes are, from a mobility management perspective, crucial nodes in the sense that the mobility anchor point may be co-located with theses nodes. Moreover, gateway nodes are also relevant in terms of access control.

Portability is also a feature that is limited to the scope of a specific community or tied to a provider.

3.3 User-Provided Networks

User Provided Networks (UPN) aspects are presented in column four of Table 1. UPNs relate to a recent trend in spontaneous wireless deployments where individual users or communities share subscribed access in exchange of incentives. Hence, the user becomes a provider of services given that by sharing Internet access users devices become networking devices, for instance, they relay connectivity. This

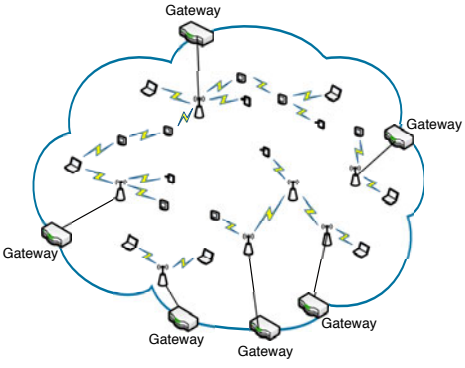


Fig. 4. UPN Scenario

aspect is not completely synchronized with the Internet end-to-end principle, which describes a clear functional splitting between end-systems and the network. A key aspect to UPNs is that they rely on existing network topologies, as illustrated in Fig. 4.

We can divide the UPNs according to two main types: infrastructure-based and mesh-based. Both types are being applied as complement to existing access networks: they allow expansion of such infrastructures across one wireless hop. For both models, there is usually one individual or entity (the Micro-Provider, MP) which is responsible for sharing its connection with N-1 other users (out of a universe of N users, who today belong to a single community). Moreover, a user is, in a specific community, simply identified by a virtual identifier (a set of credentials username and password) which is stored by a Virtual Operator (VO) and rely upon whenever the user decides to access the Internet by means of a specific community hotspot.

To better illustrate examples of UPNs, we here consider three different scenarios. In the first scenario, a hotspot owner willingly shares the Internet access with specific friends, using a local authentication procedure. The requirements and assumptions of this scenario are similar to the ones of the Hotspot model.

Another scenario corresponds to the regular municipality Wireless Fidelity (WiFi) case, being the user authentication local. In this scenario, the network adopts a mesh topology, so the characteristics and requirements of this scenario are similar to WMN, previously analyzed. The last UPN corresponds to a residential scenario, where a regular user at home decides to open access to a specific community, which is managed by a VO (cf. Table 1). The only relation the MP has to the VO is that the MP belongs to the community coordinated by the VO.

In terms of mobility management, UPNs are expected to exhibit more variability than the previously described scenarios given that user equipment is part of the network. For instance, a user can turn off his equipment at any time without previously notifying users profiting from the relayed connectivity. The impact of having users controlling portions of the network is highly related to the underlying network architecture(s): such impact may be local or propagated to the whole network. From a mobility management perspective, the MP is the crucial point to consider. The VO (in contrast to scenarios where the provider performs mobility management) is simply a coordinator for access control.

In terms of handovers, the UPNs privilege the inter-UPN handovers among MPs of users of the same community. The idea of the UPN is to provide connectivity in a large area to a user that belongs to the community. The user can move inside of a certain area, maintaining the connectivity from several access points. In general, the size of each UPN is small, so the requirement of intra-UPN handovers is not high.

3.4 DTNs

Delay Tolerant Networks (DTN) are used for chaotic conditions, like natural disasters, wars, accidents or space networks. These networks have completely different paradigms comparing with the previous ones. DTN is a concept an not a network topology or technology, since it can be applied to current network structures in disruptive and chaotic scenarios, as presented in Fig. 5. DTNs present intermittent connectivity, long and variable delay, asymmetric data rates and high error rates. The main purpose of DTNs is to deliver vital messages without losing any information, independently of the delay and connectivity. In these scenarios, some messages are stored in some nodes along the path until being possible to forward them. The traditional end-to-end notion of the transport protocols, like Transport Control Protocol, is impossible to be applied to DTNs.

In DTNs, the access control and authorization is decentralized through the network elements that form the DTN. The dimensions of the DTN depends on each scenario, from small spaces with disruption of connectivity, such as rural places, to the large and sparse places when natural disasters occur. DTNs could exist across different access technologies, and it is important to exploit all connectivity points and resources available, since they are reduced and sparse.

In DTNs, the identification of the user is provided spontaneously by the community, since the user database is inexistent. In this scenario, the identification of each user, regarding his community role and qualification, is extremely important. Besides the importance of the identification, its relevance depends on the location tracking process that is vital for the mobility management.

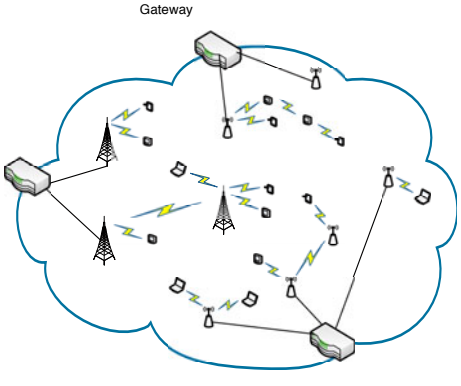


Fig. 5. DTN Scenario

In DTNs, users usually transport information, while they move along the entire network. When a user finds a connectivity point in the network, he delivers all stored information to the next storage point (other user or network element). The handovers between the DTN and the outside are not common and rarely happen, so the mobility management should consider the intra-DTN as the main priority.

3.5 Mobility Issues

In this section, we discuss the efficiency and applicability of the current mobility assumptions, requirements and solutions in user-centric scenarios. We discuss the main requirements and assumptions to improve the mobility management in the chosen scenarios, analyze the main advantages and drawbacks of the existing mobility solutions for each scenario, and identify the aspects that can be changed to improve each scenario.

Hotspot is implemented under a tree topology network model; therefore, gateways are strategical and natural points where communication messages travel, especially in inter-hotspot communications. The implementation of a mobility control point in the gateway is obviously an efficient solution, since the distance in hops between the gateway and a generic mobile node is, in average, the smallest in the entire hotspot. Moreover, data traffic between the hotspot and the outside is forced to cross the gateway, so it can be easily forwarded. In general, gateways are present in almost every user-centric scenario, being good elements for mobility control and even traffic forwarding. It is important to remember that most of the current hotspots already use bindings through Network Address Translation (NAT) to map the private IP/port addresses into a global IP/port address. The current mobility solutions can deal with the hotspot scenario, but not in a efficient way. The fact that it is required to send traffic to the home agent ([7]) every time a user wants to communicate with another one is one of the problems. Moreover, when the MN changes from one hotspot to another near the previous one, the user needs to inform its correspondent node of its new location, so that the correspondent node sends the packets directly to the new location. Another issue with current Mobile IP (MIP)-based solutions ([7], [8], [9], [10]) is that the home agent is associated with the gateways, so if the gateway fails, the home agent gets unavailable and the nodes belonging to that home agent will not be located by other nodes. In specific situations, where the hotspot has large dimensions and the user moves inside a small part of the hotspot, a control point in a network element that covers only the area visited by the user can be introduced, optimizing the mobility process, but also introducing a new level of control.

WMNs and MANETs implement mesh topologies, so they are substantially different from the Hotspot model. Although the traffic between the ad-hoc network and the outside needs to cross the gateway, interior traffic does not. Thus, gateways usually are implemented in the border of the ad-hoc network. In this sense, it is possible that a mesh node with fewer hops to a certain mobile node than the distance to the gateway, can take part on the local mobility management, specially in large WMNs. This solution is possible only for WMNs, since the mesh routers are stable and fixed. Specially in large WMNs, a control unit inside the network can decrease the overhead of control and improve the time to react to changes. Data packets are routed through wireless hops, delaying the delivery, which depends on the load to

access the medium, in a considerable time. The current MIP-based solutions only implement the mobility control up to the gateway, so these assumptions require significant changes. In a MANET, it is difficult and dangerous to select a control point in the middle of the network, since the routers, composed by the mobile nodes, are constantly moving and changing routes and topology. Therefore, it is better to implement the control functions only up to the gateway or in specific situations where the user has appropriate patterns that relate to stability. The current mobility management solutions do not cope with the dynamics and adaptability of the mobility control points along time. It is important that a user is able to subscribe a new mobility control point, not only when the previous one goes down, but also when a new one offers better conditions.

The UPNs bring different paradigms in mobility management, since a mobile node can use a wireless access network according to the user that becomes a Micro-Provider (MP); therefore, a MP can switch off its network equipment and the users in this UPN will lose the internet connectivity. In this sense, the users that become MP need to have specific stability patterns. We can have several MPs near each other, sharing their internet connection with other members of the same community, so, it is useful to exploit mobility control mechanisms among them for users that spend their time using these MPs to obtain connectivity. In some scenarios, these MPs can create a mesh network with their devices, being connected by two different ways. These cases, together with specific patterns of the users, provide the substrate to introduce mobility control inside the UPNs. The current solutions of mobility are not prepared to cope with these dynamic scenarios, where gateways and other network elements are constantly changing. Thus, home agents and mobility anchor points should not be statically defined in a certain network element; they should have the capability to be transferred to other points, reacting to the network behavior. In these scenarios completely focused on the user as the last access, the identification of each user should be taken into account, since most of the current mobility solutions do not use it, defining the IP for both location and identification.

DTNs present different characteristics and assumptions when compared to the previous ones. These networks can have both structured and unstructured parts forming the entire network. This scenario then covers several concepts of the previous scenarios, where spontaneity, dynamics, social and priority are the main paradigms. These networks are not controlled, but it is important to assure that different entities have different priorities to properly operate, such as in a natural disaster. The messages must be delivered without losing packets in the communication. The location part of mobility management is of great importance for this kind of applications, since we need to find as fast as possible the designated user and send him the vital information. In this scenario, the mobility control points should be very dynamic, since the network changes all the times, according to network, users and environment. In this case, it becomes important to introduce the control points near the most important users and in strategical places according the environment, inside the DTN. The search for control points needs to be very fast, since some devices carry with them vital information that should be forwarded to the network as soon as possible. Currently, no mobility solution can handle with chaotic scenarios, since it was not envisioned for these applications. The mobility management solution for this

scenario should be analyzed from a different perspective, integrating adaptability, spontaneity and personalization (e.g. identification and qualifications) in a overall mobility management solution.

4 Challenges of Mobility Management

In this section, we discuss the main issues in mobility management when dealing with user-centric scenarios, which will then be used to derive the fundamentals of a user-centric mobility management architecture. Fig. 6 illustrates the main challenges and initial ideas to improve the mobility management.

The current Internet model uses the IP address for both identification and location. In the Future Internet, it is expected that the IP address will only represent the location of a certain device according to the distribution of the IP addresses. The identification may not be connected to the IP address, since the user may change access network, use multiple interfaces and several devices. In this sense, binding definition needs to be re-thought to integrate the user's identification in the binding mechanism: this would allow the update of all IP addresses of the different interfaces, independent of the location and device of the user. Moreover, it provides the base to develop personal mobility, exploit multihoming, and increase the flexibility in the mobility management supported, enabling the user to be connected to different networks and mobility control points according to the services, network conditions and user requirements.

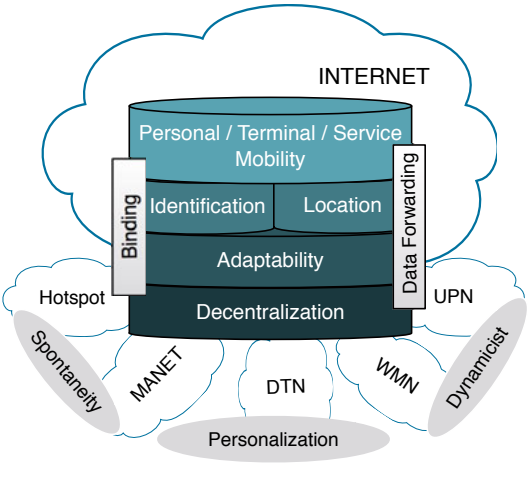


Fig. 6. Challenges for Mobility Management

Current mobility solutions define mobility control points, like the Mobility Anchor Point (MAP), in order to maintain the bindings of the mobile nodes. The control plane is separated from the data plane, since the binding update process is independent of data forwarding. However, it is not clear if control plane shall be completely

separated from the data plane. Considering that the routing and similar mechanisms are part of the control plane, current mobility solutions are more concerned with the control plane, specially the bindings maintenance; data forwarding is treated with the same approaches in the different scenarios. One possible solution for this problem is the development of two separate structures, one dealing with binding and another dealing with the data forwarding. These two structures need to be integrated in order to provide the best results for the general mobility management approach. This idea provides flexibility in the selection of both structures, since control binding elements require less resource than the data forwarding elements. Another possible solution is to develop an approach that deals with both parts, defining elements that, not only maintain bindings, but also decide on the traffic forwarding. This approach does not need the communication of the two structures, reducing overhead and time to react to events. However, it is much more rigid in the selection of the elements, since they will receive and send traffic not destined to them.

User-centric scenarios, as explained before, present several requirements, which dynamicity, spontaneity and personalization are some of them. In order to provide mobility management in these scenarios and according to the previous discussion on control and data traffic, the control points (bindings and data) must adapt to network changes, as illustrated in Fig. 7. These control points should be dynamically selected according to current network information, users information and services. When a control element fails (it moves or shuts down), the control mechanism needs to be moved to another element of the network. Besides failures, the control point could be moved to optimize resources and improve the general performance of the network (Fig. 7). A network element shall be able to recognize its neighbors that are available to be control points. Each control element should be classified according to context information in order to produce a ranked list of the neighbors' control elements (several ranked lists may exist according to user and services context). If a control element is shutting down, it selects the best one of the neighbors' ranked list and interacts with the selected control element to transfer the stored information and functions. As discussed in subsection 3.6, the distribution of mobility control points depends on the user-centric scenario. If binding and data control are aggregated in an unique structure, mobility control points will be advantageous when implemented in the gateway or edge node. However, if control points for binding and data are separated, the control binding points can be implemented inside the user-centric network, specially in the WMNs and DTNs.

Finally, we discuss the method to place the control points through the network. The four main approaches are: centralized, hierarchical, decentralized and distributed, sorted from the most central to the most distributed. Each one has advantages and disadvantages, regarding binding maintenance and traffic forwarding. In the centralized approach, it is easier to update and search for bindings, since the central point is well-known, never changing. However, the centralization of the binding storage increases overhead and latency when a user wants to communicate with a near one, since each user needs to constantly update his location to the central point. Centralization has another disadvantage regarding tolerance to failures, since it is a unique point of failure. Regarding the data forwarding, it has several disadvantages,

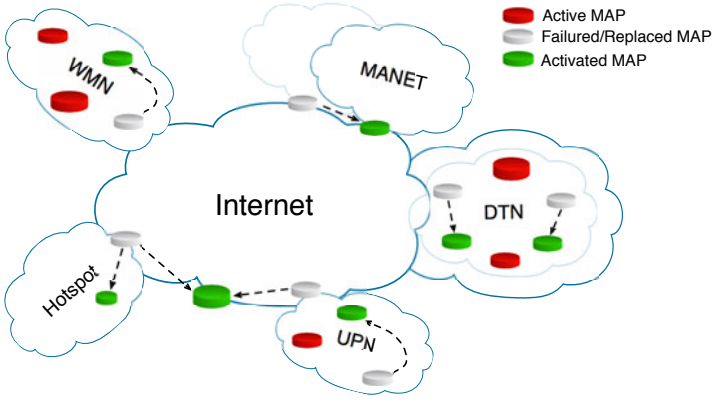


Fig. 7. Principals for Mobility Management

where one of them is the unique point of failure. Forwarding traffic with a central point implies that all traffic crosses this point, increasing the packets load near the central point and causing an unload-balanced traffic in the network. Another problem is the time to redirect traffic to the new location of the user, since the central point is far away from user. The hierarchical approach presents the same problem of failure, since we have several levels and the higher level usually is a central point that controls several lower levels. However, this solution introduces scalability, since it allows a more efficient binding search and update when a user wants to communicate with a neighbor, since it only uses lower levels for updating and searching. The data forwarding in hierarchical approach presents similar problems to the centralized approach, since it needs to cross the central point of the higher layer of the hierarchy. The decentralized approach is the most balanced, since it distributes the bindings across several points, according to a predefined criteria. So, it is much more tolerant to failure than centralized or hierarchical. Depending on the size of network, this solution can use different number of control points; they are relatively near to users, being easier and faster to update and search for users. However, decentralization implies an efficient method to distributively search for user location, or to synchronize the decentralized nodes. Regarding data forwarding, decentralization allows to redirect data to the current user's location in a fast way, without a significant impact in the overhead and load balancing. The last method is the distributed, where all nodes of a network participate in the mobility management. This method is the most tolerant do failures, since, even with several failures the network continues to work. However, this approach requires a large overhead and intelligence to search for a user's location. The simplicity in updating process increases the complexity in the searching process. Another problem is the update of information among the entire nodes, that increases not only the overhead but also the time of synchro- nization. This time needs to low, since several events (search and update) are constantly happening and the answer to these requests should be according to the latest network information.

5 Conclusions

This paper assessed the efficiency and applicability of current mobility assumptions in user-centric scenarios, addressing their requirements and solutions when applied to several types of networks that may exhibit user-centric characteristics, such as hotspots, wireless mesh and ad-hoc networks, user provided networks and delay tolerant networks. This study showed that current solutions of mobility are not prepared to cope with most of the scenarios: they may not be efficient (hotspot), they do not cope with the dynamics and adaptability of the mobility control points along time (WMNs, MANETs and UPNs), or they are not prepared to specific scenarios (DTNs). This paper also identified the fundamentals of a user-centric mobility management architecture able to efficiently deal with the aforementioned scenarios, such as: the integration of the user's identification in the binding mechanism; the coupling and decoupling of both control and data planes; the support of dynamic control points according to network conditions, user and service requirements; and the decision on where to place the control points in the network. The definition and specification of the user-centric mobility architecture, addressing the several issues discussed here, will be left as future work.

References

1. Latré, S., Simoens, P., De Vleeschauwer, B., Van de Meerssche, W., De Turck, F., Dhoedt, B., Demeester, P., Van den Berghe, S., de Lumley, E.G.: An Autonomic Architecture for Optimizing QoE in Multimedia Access Networks. Comput. Netw. 53(10), 1587–1602 (2009)
2. Akyildiz, I.F., Xie, J., Mohanty, S.: A Survey of Mobility Management in Next-generation All-IP-based Wireless Systems. Wireless Communications 11(4), 16–28 (2004)
3. Hussain, S., Hamid, Z., Khattak, N.S.: Mobility Management Challenges and Issues in 4G Heterogeneous Networks. In: InterSense 2006: Proceedings of the First International Conference on Integrated Internet Ad Hoc and Sensor Networks, p. 14. ACM, New York (2006)
4. Kumar, B.P.V., Venkataram, P.: Prediction-based Location Management Using Multilayer Neural Networks. Journal of Indian Institute of Science 82(1), 7–21 (2002)
5. Samaan, N., Karmouch, A.: A Mobility Prediction Architecture Based on Con- textual Knowledge and Spatial Conceptual Maps. IEEE Transactions on Mobile Computing 4(6), 537–551 (2005)
6. Lei, Y.-X., Kuo, G.-S.: Impact of MAP Selection on Handover Performance for Multimedia Services in Multi-Level HMIPv6 Networks. In: IEEE WCNC 2007 Wireless Communications and Networking Conference, pp. 3901–3906 (11-15, 2007)
7. Johnson, D., Perkins, C., Arkko, J.: Mobility Support in IPv6 (MIPv6). RFC3775 (Proposed Standard) (June 2004)
8. Soliman, H., Castelluccia, C., Malki, K.E., Bellier, L.: Hierarchical Mobile IPv6Mobility Management (HMIPv6). RFC 4140 (Proposed Standard) (August 2005)
9. Koodli, R.: Fast Handovers for Mobile IPv6. RFC 4068 (Proposed Standard) (July 2005)
10. Gundavelli, S., Leung, K., Devarapalli, V., Chowdhury, K., Patil, B.: ProxyMobile IPv6. RFC 5213 (Proposed Standard) (August 2009)

Author Index