Cross Layer Simulation: Application to Performance Modelling of Networks Composed of MANETs and Satellites

Riadh Dhaou[1], Vincent Gauthier[2], M. Issoufou Tiado[1], Monique Becker[2], and André-Luc Beylot[1]

[1] INPT-ENSEEIHT/IRIT
[2] GET-INT/SAMOVAR
{dhaou,tiado,beylot}@enseeiht.fr, {gauthier,becker}@int-evry.fr

Abstract. The Cross layer concept is a new way to see the quality of service in the network. It consists in adapting the current mechanisms at one level to the underlying levels and the definition of information to share between, not necessarily adjacent, levels and the global optimisation instead of multiple optimisations. Performance optimisation of the whole system is a crucial step in the process of design and validation of new systems. Models of large, complex and dynamic systems can often be reduced to smaller sub-models, for easier analysis, by processes known as decomposition or aggregation methods. Those techniques have to be implemented in a dynamic simulation tool. It has to be dynamic because simulations are driven by dynamic data and entail the ability to incorporate additional data (either archived or collected online) and reversely the simulator will be able to dynamically steer the measurement process.

Keywords: Cross layer, simulation, heterogeneous networks, network modelling.

1 General Introduction

The performance of wireless systems is extremely sensitive to the mobility, to the dynamicity and to the environment. Those systems may be extended on a large scale (long delay satellites, delay tolerant networks, IEEE802.20…) and are more and more complex.

The concept of layered network stack is the main idea which allowed the heterogeneous development of networks and standard protocols. But the specificity of the wireless networks, their interconnection with fixed wired networks and the use of adapted QoS mechanisms to wireless conditions have an impact on the performance.

The Cross layer concept is a new way to see the quality of service in the network. It does not consist in the addition of reservation mechanisms of bandwidth or any other Ad Hoc mechanism, but in adapting the current mechanisms at one level to the underlying levels; the definition of pertinent information, to share between not necessarily adjacent levels; and the global optimisation, instead of multiple optimisations.

Performance optimisation of the whole system is a crucial step in the process of design and validation of new systems. Models of large, complex and dynamic systems

D. Kouvatsos (Ed.): Next Generation Internet, LNCS 5233, pp. 477–508, 2011.

can often be reduced to smaller sub-models, for easier analysis, by processes known as decomposition or aggregation methods. Certain criteria for successful decomposition can be established. Models have to be designed on several levels and all the levels cannot be studied simultaneously. Decomposition has long been recognised as a powerful tool for the analysis of large and complex systems.

Those techniques have to be implemented in a dynamic simulation tool. It has to be dynamic because simulations are driven by dynamic data and entail the ability to incorporate additional data (either archived or collected online) and reversely the simulator will be able to dynamically steer the measurement process.

The first part of this paper provides an overview of the time, space and level dependence of highly complex networks. Then a state of the art of cross-layer design for today networks is presented.

The paper then designs cross-layer simulations which are aggregation methods, where an aggregate is a layer model. Those simulations are dynamic, autonomic and multi-layer. Examples of applications are presented: they concern the performance evaluation of MANETS. For the evaluation of routing algorithms MAC and network layers have to be simultaneously simulated. For transport layer MAC layer has to be modelled. Those cross-layer simulations happen to be difficult to operate, but they are promising methods.

In the second part we are interested in performances issues of routing in Ad Hoc networks and IP over GEO satellite links. Ad Hoc networks are collections of high rate randomly moving wireless devices within a particular area. These devices vary dynamically in their resource-richness. Such highly dynamic systems rise specific problems such as routing, media access control (MAC) and QoS mechanisms.

Many routing algorithms are proposed and standardisation is in process. Route is made difficult by the continuous moving of the devices as nodes in the networks. Access to the medium is made difficult by the so called hidden node problem. QoS mechanisms need to be designed by taking into account several techniques such as load balancing, MAC-level and IP-level differentiating mechanisms. Battery energy may be also saved by optimising the emission power. The distribution of the bandwidth in these mobile Ad Hoc Networks (MANETs) and GEO satellite networks can be very unfair. The behaviour of TCP and the introduction of QoS mechanisms at different levels and their interdependence remain to be studied more in detail in those systems.

Part 1 - Time, Space and Level Cyclic Dependence of Highly Complex and Dynamic Networks: Application to Dynamic Cross Layer Design of Mobile Systems

2 Introduction - Part 1

Performance evaluation is a crucial step in the process of designing and validating new complex telecommunications systems. Complexity comes from mobility, dynamicity, extension on a large scale and interaction between layers and elements. But performance evaluation of such systems is also a technical challenge. Out of the three classical methods for evaluating performances (analytical study, simulation study and experimental study) [26], only one may be applied in this context: the simulation. Indeed, analytical studies can only cope with systems of a reasonable complexity. This is definitely not the case for the new generations of telecommunications systems (satellite, mobile and sensor networks). And, as long as the whole system is not operational (this is obviously the case at the early stage of the design), experimental studies are either impossible, or restricted to small parts of the system.

Studies based on simulation techniques require both a computational model of the system and a simulation environment. Computational models are particularly difficult to elaborate because the system is large, complex and highly dynamic. At the physical level, for example, propagation effects are hard to model and often need to be accurately simulated in order to get realistic results. Furthermore, radio interfaces are now offering sophisticated mechanisms like power control or adaptive coding. Those particularities are added to the traditional complexity of mobile networks based on IP networks and protocols.

The motivation for improving the performance of networks has increased the interest in systems which rely on interactions between different levels. Cross layer systems appeared recently specially for wireless networks. The research, specific to cross layer systems, aims to optimise the overall performances of the protocol stack of a network rather than a local optimisation of each layer. The technique consists in taking into account, at the same time, information available from different levels [10], not necessarily adjacent, in order to create a system much more sensitive to wireless specificity.

This approach is advantageous with reserve to respect a minimal layered generic model [7], because the separation between layers allows a better interoperability between network elements (IP over everything) and allows to introduce relatively easily new local protocols without changing the other layers of the stack. There is a need to set up a generic element that interfaces with existing elements and using the current network interfaces [9].

Out of the challenges in designing cross layer systems, a major interest of performance evaluation community is:

- To design and develop a simulation environment introducing innovative features to ease model reuse and hierarchical modelling activities, as well as implementing powerful computing techniques;
- To define generic cross layer system network models, and implement their simulation within the framework; all mechanisms shall be addressed in order to provide a global model of the system;

- To identify and evaluate appropriate simulation techniques in cross layer network context.
- To implement mechanisms in one of the levels of the protocol architecture, while modelling in a rather accurate way the other levels (even if all the protocols are not yet standardised). For a study on network level, physical layer, data link and network level models must be developed; these models having to coexist (since the parameters of some models may influence the other hierarchical levels).

For these reasons, new simulation techniques called "Cross-layer Simulations" are set up. Those techniques have to be implemented in a dynamic simulation tool. It has to be dynamic because simulations are driven by dynamic data and entail the ability to incorporate additional data (either archived or collected online), and reversely the simulator will be able to dynamically steer the measurement process.

The generic cross layer network model, together with cross layer simulation model facilities offered by the tool, shall allow studies on cross layer systems.

This paper presents a state of the art on cross layer systems. Section 2 is emphasising on modelling problems that are specific to wireless cross layer networks simulation. Section 3 is presenting some original approaches that will be of use within a cross layer simulation environment.

3 Wireless Networks Complexity

Wireless networks have particularities that distinguish them from wired networks.

The wireless channels vary over time and space and has short term correlation due to multi-path. These variations are caused by either motion of the wireless device or changes in the surrounding physical environment, and lead to detector errors. This causes bursts of errors to occur during which packets cannot be successfully transmitted on the link.

Small scale channel variations due to fading are such that states of different channels can asynchronously switch from “good” to “bad” within a few milliseconds and vice versa. Furthermore, very strong forward error correction codes (i.e., very low rates) cannot be used to eliminate errors because this technique leads to reduce spectral efficiency [5].

In addition to small-scale channel variations, there are also spatiotemporal variations on a much greater time scale. Large scale channel variation means that the average channel rate conditions depend on user locations and interference levels. Thus, due to small scale and large scale changes in the channel access time than others based on their location or mobile velocity, even if their data rate requirement is the same as or less than other users.

4 Cross Layer Design

The concept of layered network stack is the main idea which allowed the heterogeneous development of networks and standard protocols. On these principles is that Internet networks could be born and with it the birth of all applications on IP and the true heterogeneous development of the networks. This model of this development gets many advantages such as:

- Each layer is independent.
- One generic model would be able to solve multiple problems
- The layers are quite interchangeable (different MAC layers exist)

But the growth of technologies in the field of the networks and the multiplication of the needs made the basic model more complex and these interactions (interfaces) which was formerly well structured, are now more complex and are not clearly defined.

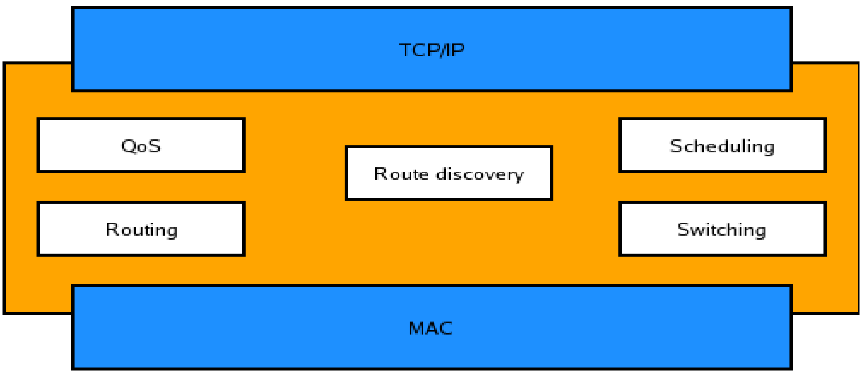

Fig. 1. Layers functionality

To avoid these problems in the GSM networks packets, multiple encapsulations is used to interface correctly GSM network and IP network. In the case of the ADSL access network, the same solution is applied to this problem. But, this technique increases enormously the overhead and the complexity the network topology. In the case of the MPLS networks, the interface between the networks layer and MAC layer was complex and then the limit between MAC layer and IP layer became very trouble (the IP addresses are used both by MAC layer and IP layer). Current researches in wireless networks, ad hoc networks and sensor networks find out new challenges which must be solve. In this type of networks, the problems of radio resources which are rare and time varying make problems which are very difficult to adapt with a layer structure where information between the different layers are independent and completely separate.

The evolution towards one Internet network which is able to interface with ad hoc networks and sensors networks in the same addressing map take the following problem: how to preserve the compatibility of the existing model taking in account the currently defined interfacing problems.

4.1 Inter-Layer Interactions

The diversity of the media and the different problems, which are likely to appear, cannot be solved in a generic and total way. Then it is necessary to create algorithms closer to the needs of each resource. The layers between themselves have to be able to exchange information about each problem. The concept of Cross-Layer makes possible to exchange information between the various networks layers. It allows improving the algorithms available at each network layer, thanks to information available in the other ones. One can for example use information about the link quality in the MAC Layer to create more powerful algorithms of routing. That's why the structure of networks stack must be modified so that information exchanges between the layers are possible. The layer structure itself does not have therefore been modified, that's why the creation of elements external to the networks stack, is used for this aim. Network Status is a component of the stack that interfaces the different layers between themselves. It acts as a database where each network layer can put or get information.

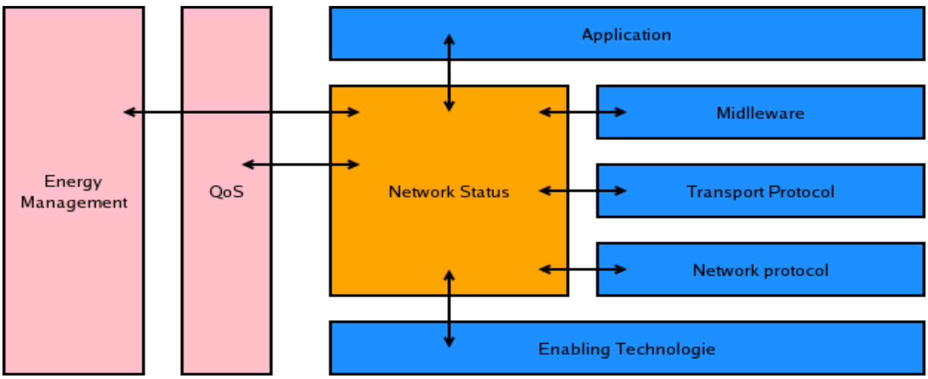

Fig. 2. Inter-layer interactions

The main think to be taken into account is to choose which information will be shared and which one will impact on the various algorithms (of routing for example). In wireless networks the element that we have to take into account is the radio channel quality and the energy consumption. But these elements cannot be managed simply at a local level but must be managed in a distributed way in the network. That's why a new network's metric must be created, instead of looking only to the number of hops between the transmitter and the receiver; it will also look at the quality of the radio channel along the entire path between the transmitter and the receiver. The assumption of channel quality in the metric can be considered as a new way of doing quality of service. Because it is now seen as an internal component of the networks, as for example the ad hoc layer in the grid systems. Its main advantage is the overall improving of the network capacity, and also the decrease of the QoS management complexity.

4.2 MANETS or Sensor Networks

The sensor networks are collaborative networks and very directed towards a precise application field, which implies that their structure is more easily controllable than the ad hoc network one. Because in the ad hoc networks the mobility and the very changeable nature of flows make very difficult any forecast on the nature of flows to transport. The structure of the network by its nature (mobile) makes the connectivity very uncontrollable. In the ad hoc networks the network stack is well defined and it is quite difficult to make changes to improve it. Whereas in the sensor networks, we can easily avoid this lack of flexibility by considering other models of protocol stack, more adaptable and arranged un-hierarchically. That's why the sensor networks are excellent candidates to implement a Cross-layer structure in their network stack for the exchange of information between the different layers.

It is obvious that cross layer design creates interactions, some intended, and others unintended. There is a need to examine the dependency relations, and to enforce time scale separation. Authors, in [7], argue that proposes of cross layer design must therefore consider the totality of the design, including the interactions with other layers, and also what other potential suggestions might be barred because they would interact with the particular proposal being made. They must also consider the architectural long term value of the suggestion. Cross layer design proposal must therefore be holistic rather than fragmenting.

Generally, properties of wireless networks are different from those of wired networks. In fact the interference created by the nodes emitting packets slow down the global throughput. The versatility of the wireless medium and the variability of the capacity of transmission of nodes have also an impact on the network throughput.

Another important factor in wireless networks is related to the problem of hidden station, where to avoid collisions a station that is emitting reduce the other neighbouring stations to silence [9].

Furthermore, routing algorithms used in wired networks such as shortest path are not adapted to wireless Ad Hoc networks [7].

Finally, the performances of the transport protocol TCP is not really adapted to wireless networks, or to Ad Hoc systems prone to high BER and mobility. Those factors induce connection interruptions, interpreted by TCP as a congestion indication, and decrease connections throughput. Congestion control mechanisms, including the congestion avoidance phase in all variants of TCP, have recently been shown to be distributed algorithms implicitly solving network utility to be distributed algorithms implicitly solving network utility maximisation problems, which are linearly constrained by link capacities that are assumed to be fixed quantities. However, network resources can sometimes be allocated to change link capacities, therefore change TCP dynamics and the optimal solution to network utility maximisation. For example, in CDMA wireless networks, transmit powers can be controlled to give different Signal to Interference Ratios (SIR) on the links, changing the attainable throughput on each link [1].

4.3 Examples of Cross Layers

The cross layers should allow to consult in real time available information on each network layer, nowadays this is not possible. Authors of [6] observe that some network functions, such as energy management, security and co-operation are cross layer by nature. The authors propose cross layer architecture, presenting a network status plan, which divides every shared network's functionality in different layers. Table 1 presents the levels implicated in different cross layer studies.

Table 1. Panel of cross layer studies at different levels

Physical layer	Channel state is useful to adapt the throughput to [1],[15],[20],[21].
MAC layer	The retransmission number at the MAC level may indicate the quality of the link, timer [2],[3],[4],[8],[11],[15],[20],[21],[23].
Network layer	Using quality of transmission information in routing algorithms [2],[3],[4],[8],[23].
Transport layer	The MAC layer could adapt the error control scheme among TCP retransmission timers and update the information on the delay for better QoS. [1], [2], [3], [11].

4.3.1 Physical-MAC Layers

A set of novel PHY-MAC mechanisms based on a cross layer dialogue was proposed. System efficiency improvement is achieved by means of automatic transmission rate adaptation, trading off generic packet switched CDMA access network [15]. The rate adaptation mechanism improves spectrum efficiency while keeping packet delay minimised. On the other hand, power dependent strategies reduce power consumption and inter-cell interference.

The joint effect of MAC and PHY layers on power efficiency was investigated. Specifically, Authors in [20] present a study of the link adaptation for a power efficient transmission by selecting a proper transmission mode and power level with the aid of a derived power efficiency model for IEEE 802.11a WLAN.

Another study, presented in [21] considers a reservation based medium access control MAC scheme where users reserve data channels through a slotted ALOHA procedure. The base station grants access to users in a Rayleigh fading environment using measurements at the physical layer and system information at the MAC layer.

4.3.2 MAC Layer

In [19] authors concentrate on the problem of medium access for wireless multi-hop networks. They first study CSMA/CA, and find that its performance strongly depends on the choice of the accompanying routing protocol. They then introduce two protocols that outperform CSMA/CA, both in terms of energy efficiency and achievable throughput. The Progressive Back-Off Algorithm (PBOA) performs medium access jointly with the power control. The Progressive Ramp-Up Algorithm (PRUA) sacrifices energy efficiency in favour of a tighter packing of transmissions and higher throughput.

The authors observe that: (i) power control can be very helpful in terms of energy efficiency, but its gains are limited in multi-hop environments. (ii) The First In First Out (FIFO) queuing discipline is sub-optimal for use in wireless ad hoc networks. A more relaxed rule can lead to a tighter packing of transmissions. (iii) The choice of routing protocol can significantly reduce the capacity of the network. Specifically, any routing strategy that assigns a single route to a node pair typically suffers a penalty in its performance. (iv) CSMA/CA is not well-suited for communication over weak links. Therefore, it should be coupled with routing protocols that avoid, if possible, such links. (v) Medium Access protocols that are distributed inherently carry a penalty in their performance since, to achieve capacity, it is necessary that nodes that are separated by an arbitrary large distance coordinate their transmissions. By definition, this is impossible in a distributed algorithm.

Authors in [5] observe that multi user diversity gains can substantially improve wireless network throughput. However when giving different types of QoS constraints, most of the studies so far are confined to single-cell scenario.

4.3.3 MAC-Network Layers

Channel reservation control packets employed at the MAC layer can be utilised at the physical layer in exchanging timely channel estimation information to enable an adaptive selection of a spectrally efficient transmission rate. In particular, the size of a digital constellation can be varied dynamically based on the channel condition estimated at the receiver which can be relayed to the transmitter via the control packets [2]. In addition this channel adaptive information gathered at the MAC layer can be communicated to the routing layer via different routing metrics for optimal route selection.

Many authors propose energy efficient schemes for wireless ad hoc and sensor networks. [4] proposes a scheme that utilises cross layer interactions between the network layer and MAC sub-layer to achieve energy conservation. In order to improve cross layer optimisation, [8] propose a multiple access collision avoidance protocol that combines RTS/CTS with scheduling algorithms to support the multicast routing protocol. The protocol avoids collision by including additional information in the RTS. Proposed scheme, together with extra benefits, such as power saving, reliable data transmission and higher channel utilisation compared with CSMA or multiple unicast, enables the support of multicast services.

Currently approaches solving problems associated with multicast for ad hoc networks solve them at the network layer. Authors in [8] propose a procedure for which, branch nodes are generated. In a simple multicast MAC protocol for branch node's transmitting multicast data is proposed, but there is no consideration of the hidden node problem. In [8], authors propose a multiple access collision avoidance protocol for multicast services that resolves the hidden node problem in ad hoc network.

The awareness of higher layer requirements is addressed in [16]. In fact, fundamentally applications look for services, not for IP addresses. This is especially so in ad hoc zero configuration application scenario. The "demand" for on-demand operation comes in the form of a service request from the user. The way the lower layers function should be determined by what the application requires. This suggests that having some information exchange across the layers would be of use. Specifically, the link-layer can use information about the "demand" to decide whether or not to form a particular link, or to decide which of a set of links to form.

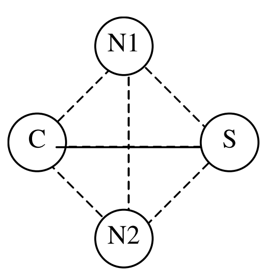

Fig. 3. Link layer topology

Consider the scenario depicted in figure 3. Node C is a client looking for a service S, N1 and N2 are two other nodes. All nodes are in physical proximity of each other. The link formation layer has to decide which links to form. Without any information on what the layer above requires, the algorithm to form links would be direction less. It could result in an inefficient route between C and S. Worse, it could unnecessarily activate extra links (the dotted ones in the figure), effectively voiding all the effort that has gone into the careful definition of low power link modes.

Path coupling involves MAC layer interactions that impact the performance of network layer paths that are otherwise disjoint. These interactions are shown to have significant impact on energy efficiency, throughput and delay [23].

4.3.4 MAC-Transport Layers

In [1] authors present a distributed power control algorithm that couple with the TCP protocol to increase the end-to-end throughput and energy efficiency of multi-hop transmissions in wireless ad hoc networks. The authors prove that the nonlinearly couples system converges to the global optimum of the joint congestion control and power control problem.

[3] Incorporates user feedback into the protocol stack by which the TCP throughput of desired set of applications running on the mobile host can be dynamically controlled. In addition the authors use the lower layer connection and disconnection information, to improve TCP performance.

4.3.5 MAC-Transport-Application Layers

In wireless networks, FEC protection is required at the application layer regardless of the underlying transport layer protocol in order to deliver high bit rate multimedia. [11] show that the amount of FEC overhead required, by transport protocol such as UDP Lite is considerably less than traditional UDP. Hence UDP Lite provides improvement in bandwidth utilisation in order to deliver loss less multimedia. [11] illustrates the suitability of cross layer protocol strategies supporting application-specific multimedia.

4.3.6 Application Layer

Cross layer information help also application level to understand network status. Many authors propose cross layer techniques for adaptive video streaming over wireless networks [12] and adaptive filter based on image content and dynamically changed threshold based on cross layer information [13].

It is shown that application adaptation can provide significant energy reductions over both fixed and adaptive hardware. Furthermore, [14] achieve joint hardware application adaptation while preserving the logical separation between layers. In this paper, authors propose a set of end to end application layer techniques for adaptive video streaming scheme over wireless networks. The adaptation is done both with respect to channel and data. Authors demonstrated the effectiveness of the application layer adaptivity combined with the RLP layer granularity.

4.3.7 Cross Layer Design

The paper [6] presents the IST-2001-38113 Mobile Man Project. The paper present the Mobile Man architecture (depicted on Figure 4). Some network functions, such as

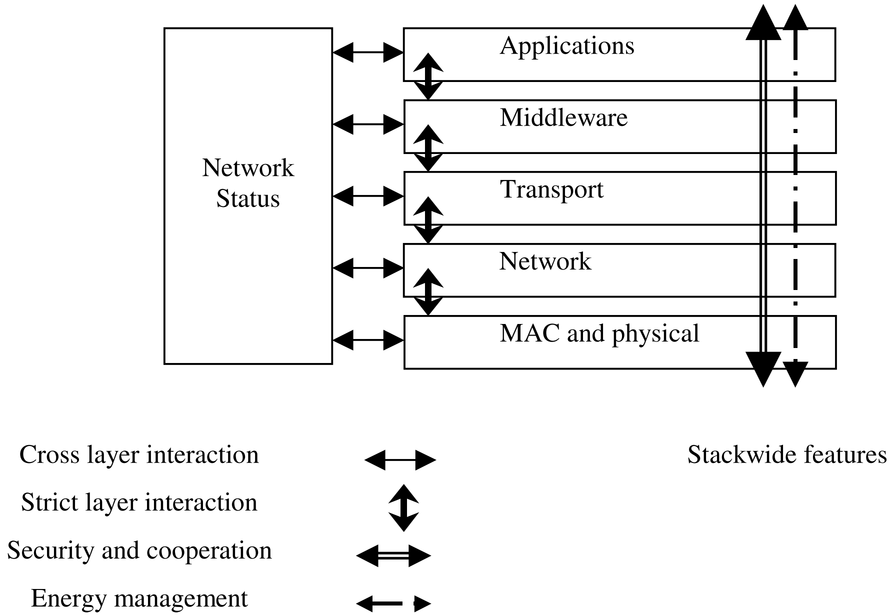

Fig. 4. Cross layer architecture (proposed in MobileMan project)

energy, management and security and cooperation are cross-layer by nature. Mobile Man seeks to extend cross-layering to all network functions through data sharing.

4.3.8 Cross Layer Information

Authors in [17] note that ad hoc networking is a multi-layer problem. The physical layer must adapt to rapid changes in link characteristics. Such as, the multiple access control layer needs to minimise collisions, allow fair access, and semi-reliably transport data over the shared wireless links in the presence of rapid changes and hidden or exposed terminals. The network layer needs to determine and distribute information used to calculate paths in a way that maintains efficiently when links change often and bandwidth is at a premium. Analysis above imply that each of protocol layer in ad hoc network is dependent each other, whether they directly connect or not. So a cross layer protocol design that supports adaptability and optimisation is needed.

However, authors note that, there remain many open questions in the understanding and implementation of this design philosophy. Such as, when building such cross

Table 2. Extend sense topology information

Layer	Information
Application layer	Topology control algorithm, Server location, Network Map
Transport layer	Congestion window, Timeout clock, Packet losses rate
Network layer	Routing affinity, Routing lifetime, Multiple routing
MAC/Link layer	Link bandwidth, Link quality, MAC packet delay
Physical layer	Node's location, Movement pattern, Radio transmission range, SNR information

layer architecture, complexity issues are introduced that, if not managed properly, can be insuperable. How to balance the power control and other system constraint such as delay, mobility and longevity in ad hoc network is still another problem.

4.3.9 Cross Layer Signalling

[22] propose an efficient, flexible and comprehensive scheme defined as Cross Layer Signalling Shortcuts (CLASS) and observe that layer by layer propagation approach of signals follows the data propagation mode. Consequently, the intermediate layers have to be involved even if only source layer and the destination layer are actually targeted. This causes unnecessary processing overhead and propagation latency. Direct signalling between layers adopted in CLASS is presented in the table 3.

Table 3. QoS adaptation and information exchange (for CLASS)

	APPLICATION	TRANSPORT	NETWORK	LINK	PHYSICAL
APPLICATION Real/non real time services	O	↔ Packet loss ratio, jitter, etc.	↔ delay constraints, etc.	↔ Desired value of QoS parameters	←
TRANSPORT TCP/UDP/RTP		O		↔ Joint error control using BER, Handoff notification	
NETWORK IP/ IntServ/ Diffserv	↔		O	← Joint delay control	
LINK Link quality; FEC/ARQ	↔ Environment measurements reports: SNR, RSS, etc.	↔	→	O	←
PHYSICAL Channel conditions	→ Environment measurements reports: SNR, RSS, etc.			→	O

5 Wireless and Mobile Cross Layer Systems Modelling

5.1 Modelling Goals

There is a need for a dynamic and global simulator. It has to be dynamic not only because nodes are moving fast and handoffs are often but also because simulation granularity has to be modified according to the level of the context. It has to be global because the topology is permanently changing and it is necessary to take into account every node. Also, models at different levels are required. This leads to the need for an extension of classical simulators, which were not designed for dynamic and global complex system.

The goal is to evaluate the capacity, the availability, the quality of service, and the spectral effectiveness of such complex systems. The metrics used to estimate these parameters are the following:

Capacity is estimated by the maximum instantaneous traffic, the traffic mean, the total user count, as well as the average number of connections established during a given period.

Availability is evaluated from a probability of communication establishment, or a probability of communication maintenance.

Quality of service is evaluated at several levels, through criterion of delay, jitter, BER, FER, ATM QoS parameters (CTD, CDV, CMR, CER, CLR, SECBR...), communication establishment overhead, connection outage time...

Evaluating the **spectral effectiveness** derived from an estimation of the supported traffic as a function of the number of subscribers, the geographical zone and service classes.

The studies concern the availability, effectiveness, the routing, the handoff, and the quality of service… We want to be able to simulate and compare several of these mechanisms and to integrate them in other studies. For example, a study of end-to-end QoS will integrate a handoff model but also a resource allocation model. Let us detail later these levels.

5.2 Main Modelling Problems

Models of large and complex systems can often be reduced to smaller sub-models, for easier analysis, by a process known as decomposition. Certain criteria for successful decomposition can be established. Models have to be designed on several levels and all the levels cannot be studied simultaneously. So an aggregation technique may be used [42]. The complex system is shared into subsystems, which are studied independently. Then the global system is studied taking into account the subsystem dependencies. For example, to study a handoff mechanism, the failure problems are neglected and to study propagation problems, we do not have to simulate the whole system. Nevertheless, most of the mechanisms and phenomena are correlated and in several ways.

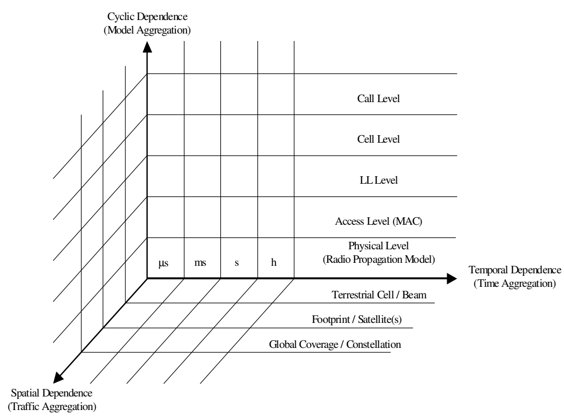

Fig. 5. Aggregation dimensions

5.2.1 Cyclic Dependence

A cyclic dependence occurs when several parts of the model, detailed on Figure 5, are correlated. For example, the performance of the radio channel depends on the traffic while the traffic itself depends on the performance of the radio channel, as shown on Figure 6. On another side, the interference calculation utilises the concurrent traffic.

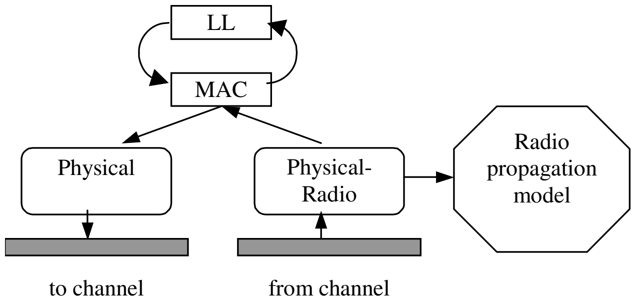

Fig. 6. Cyclic dependence

That leads to problems, for the representation of this concurrent traffic, for the relevance of its representation and the importance to take it into account.

In a limited space, many connections are handled simultaneously. Each of these connections generates a large amount of data, which results in a huge number of packets and a huge quantity of packets. But for some studies, the communication needs to be simulated both at the network and radio levels, because they are closely dependent. However, it is nearly impossible to carry out simulations, simultaneously, on two different levels.

In this case, the problem we have to solve is the choice of the model smoothness. The performance of a simulation has a significant impact on its feasibility, and its complexity has a significant impact on the confidence intervals. Studies focusing on a given level in general require a relatively coarse description of the other levels. For example, for a study on the MAC layer level, an evaluation of the BER is required from the radio layer. But it is not necessarily useful to include detailed models of orbitography in these BER computations: approximating the movement of nodes by simple equations makes it possible not to have to integrate movement equations and to replace bulky files handling by some simple calls to mathematical functions.

5.2.2 Temporal Dependence

The performance of the radio link at time t+1 depends on its performance at time t. This type of problem appears particularly when running step by step simulations with a very fine step and when running discrete event simulation with significantly correlated processes. That occurs when using detailed models of orbitography requiring, for their integration, a small step or when using long range dependence traffic models. In most cases, a trade-off has to be found between the smoothness of the model determining the step of the simulation and the desired result precision. Simulating process with long-term correlation is dangerous, because of the various time scales correlation. It is as dangerous as simulating non-stationary processes.

5.2.3 Space Dependence

The performance on site S at time t is related to the performance on a close site S’ at t or near time t. The co-localised space zones are dependent, for instance, because they share a handoff mechanism or another mechanism or phenomenon. The complexity of

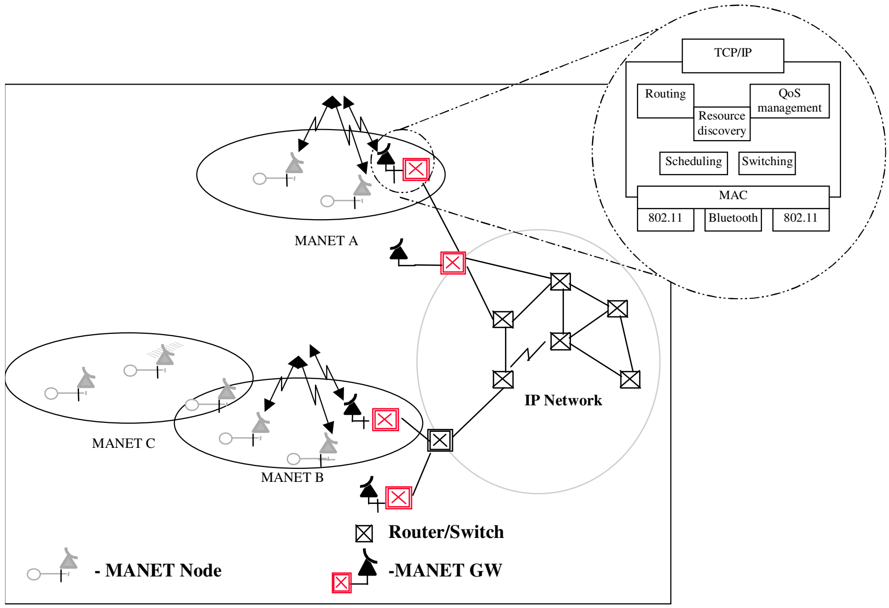

Fig. 7. Space dependence

global studies taking into account all the system can be reduced by undertaking the study at an elementary space granularity and by using the results of this study to estimate the parameters on higher levels: Iterative steps of traffic aggregation may be applied [24].

5.2.4 Multi-Scale Relations

There are also phenomena for which events at many scales of length or time do make contributions of comparable importance, or at least have a non negligible influence on each other. Any theory that describes these phenomena must take the entire spectrum of ranges into account in one way or another [25].

A typical example of multiple-time-level-space-scale dependence behaviour is provided by the highly complex, large and dynamic Ad Hoc and sensor systems. The high mobility of such systems make an experimental approach, based on measurements and testing, too much time consuming. Analytic or simulation models must be developed instead [26].

5.2.5 Choice of the Granularity

The choice depends on the studied topic. According to the type of study, various modelling levels of a same function or a same set of functions may be necessary. The modelling levels are represented on Figure 5. The multiple correlations are located; it is necessary here to identify these dependencies and treat them individually. We have to find and validate some efficient mechanisms in order to combine different time, space or model scales.

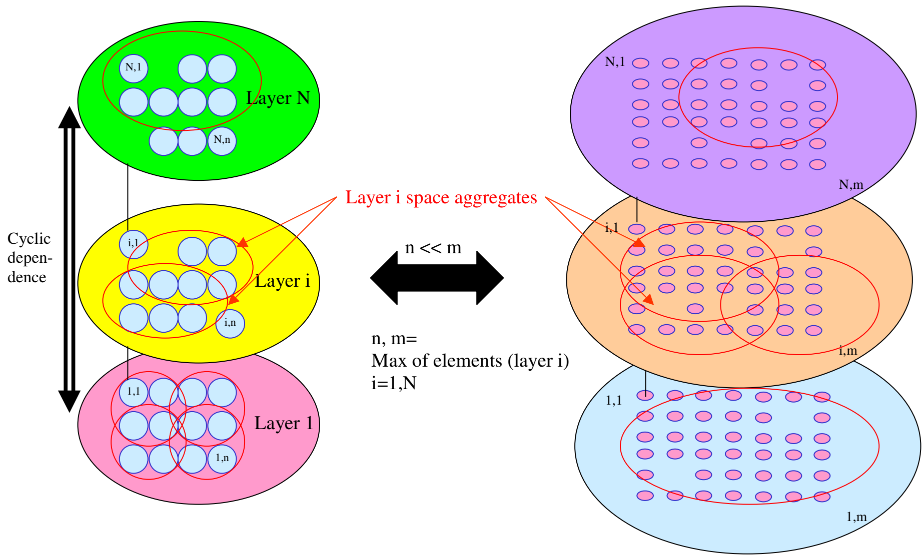

Fig. 8. Size and diversity of problems

Most of the problems we have mentioned are not treated in the traditional simulators (such as NS or OPNET), which leave the user to use some elementary block modules and to pile them up. But, the piled-up models cannot be simulated in a reasonable time to obtain realistic performance criteria. These problems are perhaps less crucial when local area networks or networks with a small number of nodes are simulated. Here the problem comes from the size (in terms of number of nodes or users) and from the diversity of the problems. It is thus advisable to consider them a priori to avoid disappointments.

Our point of view is that the only way to solve these problems is to provide for each object and modulus several visions that will depend on the level of necessary detail (time scale and space scale).

5.3 Cross Layer Simulation Approach

Setting up novel cross layer system results in the design of a completely novel in simulation environment. Usual approaches, for studying cross layer systems, are laborious and inefficient. In paragraph 0, we summarise and discuss these usual approaches. Then, in paragraph 0, we describe an approach that we propose to study cross layer systems.

5.3.1 Usual Approaches

When the need for simulating a cross layer system arises, the two usual approaches are either to build a new model of the targeted system on top of an existing simulation environment, or to build new and therefore proprietary simulation software, specifically designed for a given system. In the following, we give a short analysis of the advantages and drawbacks of these two approaches.

5.3.2 Reusing Existing Simulation Environment

Reusing an existing simulation environment, such as OPNET, STK, NS, Visualyze, COSSAP or SPW, has several advantages. First, existing environments usually include a set of common models or patterns out of which several may be reused. Second, the larger the user community of an environment is, the easier it is to find support and contributed models for this environment. Third, existing environments often come with a set of integrated or contributed tools, such as models animators and debuggers, plotters, data analysers, and so on, which improves their overall ergonomics and efficiency.

Unfortunately, cross layer systems are combining several aspects that used to be studied separately so far. Therefore, these separate kinds of studies lead to specific environments. Some are more specifically designed for network and protocol modelling, others for propagation and radio interference modelling, and others for space mechanics modelling. Of course, out of these specific environments, some have the ability to be extended to new areas. But, since a specific environment usually mean a specific and optimised design; such an extension is seldom easy: integrating new kinds of models often conflicts with the initial design and philosophy of the selected environment. The result is an added modelling complexity, as well as an added computing complexity, which are both critical points given the high complexity and large scale of the systems being studied.

5.3.3 Building Specific Simulation Software

Compared to the previous approach, building specific simulation software has the opposite advantages and drawbacks: the modelling complexity may be lowered to a minimal level and the computing complexity may be sharply optimised. But since the developments are specific and often proprietary, there is nothing or little to share, no community support, and a lot of additional effort is required to develop specific complementary tools or integrate existing ones.

5.3.4 Cross Layer Simulation Environment

Cross layer simulation environments have to provide, through an integrated user interface, all the functions required to achieve an experimental study based on simulation: model programming and assembly, experiment planning, simulation runtime support, and data analysis. This environment has to present several properties:

5.3.5 Elements of Design

Cross layer systems exhibit several levels of complexity: a complexity at a physical level with the radio transmissions, a complexity at an architectural level, with several kinds of nodes and links, a complexity at a scale level, with networks operating simultaneously up to thousands of terminals, and a complexity at a functional level, with several kinds of services (e.g.: real time or non real time services), several kinds of protocols (e.g.: IP stacks) and several kinds of procedures (e.g.: logon, power-control, CAC, hand-off). In order to cope with all these levels of complexity, the cross layer simulator has to implement and provide several innovative techniques.

5.3.6 Component Based Modelling Approach

The component based modelling approach is a key point in the cross layer design. This common pattern of object oriented programming [27], applied to the modelling area, allows for many interesting features, such as model genericity, variable granularity, and hierarchical modelling.

Genericity is a powerful property of cross layer models that allows the transparent reuse or exchange of any part of a given model in order to build new models. Variable granularity is the ability to switch the detail level of a given part of a model, from the most detailed level to the most approximate level. Hierarchical modelling is the ability to decompose complex parts of a model in simpler parts, until a reasonable complexity level is reached. An aggregation technique may then be used.

5.3.7 Open and Versatile Simulation Kernel

Cross layer simulator architecture is based on a framework approach. This will allow for a maximal flexibility and adaptability at all levels.

Typical complex simulators have shown to spend most of the computing time in event handling. Thus, this part of the simulator has to be carefully designed and optimised. But in the case of cross layer systems, the kind of designs and optimisations required depends on the part of the model being simulated. In order to allow several kinds of optimisation to be combined, cross layer simulator kernel will allow for transparent mixed-mode multi-level kernels.

6 Conclusions - Part 1

Cross layer is a promising research area, not yet completely exploited. Models based on cross layer could improve considerably wireless networks short term performance. Definition of pertinent information to share and global optimisation instead of multiple local optimisations are the key words for the future of those models. Nowadays research in this area concentrate on the model (cyclic) scale and have to take care of the cross layer design.

The Cross-Layer concept is a new way to see the quality of service in the network. It does not consist in the addition of reservation mechanism of band-width or any other ad hoc mechanism, but in adapting the current mechanisms of routing to the link. For that the sensor networks can take advantage from these exchange techniques of information, because their collaborative nature allows the use of information of other elements to improve the overall capacities of the networks.

Part 2 - Performance Issues of Routing in Ad Hoc Networks and IP over GEO Satellite Links

7 Introduction - Part 2

The ultimate aim of the Ad Hoc networks is to propose a fast deployment of new networks without presupposing the existence of an infrastructure. The aim is to have some machines successively assuming the function of host (transmitting and receiving packets) and of router to relay the packets towards other nodes of the network. The very active working IETF group MANET is interested in the standardisation of these networks and in particular in the routing algorithms. The ad hoc mobile networks are operational in local area networks, even if many studies on these algorithms are in progress...

Performance evaluation of ad hoc and satellite networks is a crucial issue. The specificity of wireless Ad hoc and long delay GEO networks has to be considered carefully. In this state of the art, we overview some routing techniques aiming to the minimisation and the optimisation the bandwidth use. Then, we present simulation results of some routing techniques considering several metrics such as fading, reception errors, overhead and delays. Finally we concentrate on the performance of IP on GEO satellite links. Section 7 is presenting a survey of ad hoc networks. Section 8 is presenting a performance comparison of the AODV and DSR routing protocols. Sections 9 and 10 are presenting some IP performance issues over GEO satellite links.

8 Ad Hoc Networks

Ad hoc network is a network spontaneously generated by a set of computers called mobile nodes to bring out their several mobility speed characteristic (pedestrian, automotive, plane, ship…). Mobile nodes use radio frequency broadcast/reception mechanism. To guaranty connectivity, those spontaneously network nodes need to work together: in the network, every node agrees to convey traffic information even if it is not the destination.

The possible mobility in space and in time of network nodes creates a dynamic topology that cause permanent route changing between any 2 nodes. This property generates the new "route installation" technique to exchange information between 2 mobile nodes. This technique allows activating the target location mechanism by the source. This "route installation" technique may be done at the network level through signalling. Over the traditional routing technique, it makes protocols to take in account environment parameters such as location, mobility, quality of service (QoS).

- Ad hoc networks are characterised by dynamic topology, limited bandwidth, energy consumption constraints, limited physical security with no centralised management. They are subject to many problems such as :
- Routing information technique choice;
- Configuration;

- Management;
- Security: signals detectors and passive receipts can spy radio communications if they are not protected;
- Low output: as the bandwidth is limited, the network management slice must be underplayed to let a maximum bandwidth to communications;
- Transmission errors : more frequent than in wired networks;
- Interference: wireless link are not isolated, two simultaneous broadcasting over the same frequency or two near frequencies came interfered. Some other equipment, like micro waves furnaces, may cause also interference;
- Changeable links: transmissions are subject to propagation condition. Link quality mandatory to operate properly radio communications;
- Signal strength: it decreases with the distance, it is subject to strong regulation by countries authorities;
- Limited radio capability: wave propagation depends on environment transmission (obstacles…);
- Hidden nodes: in some cases, for example in presence of obstacle between 2 nodes, they can broadcast simultaneously and thus can cause collision;
- No possibility to detect collision within a transmission : because nodes must broadcast and listen simultaneously;
- Energy : battery autonomy is limited;
- Mobility and dynamic topology;

The routing information technique choice problem and the configuration problem must ensure connectivity in the network. Routing information technique in ad hoc networks must be solved to ensure communication between mobile nodes.

8.1 Main Characteristics of Ad Hoc Network Routing Information Protocols

Ad hoc network protocols algorithms have to present some characteristics:

- be robust against topology changing
- be implemented over limited computing and memory equipments
- avoid routing loop
- work on limited bandwidth and try to use it in best
- be simple to be implemented

8.2 Ad Hoc Network Information Routing Mechanism Classification

Wired networks use regularly routing protocols to establish routes reaching one point of the network to another in advance, based on periodic routing information exchange between nodes. Those are pro-active protocols because they create routing tables in advance. They can be classified into two types: distant vector algorithm based protocols and link state algorithm based protocols, all of them are part of most short path and distributed routing classes.

A distance vector protocol is based on information exchange between adjacent nodes. A network node communicates with its immediate neighbours to exchange the routing information it detains. Those information incorporate nodes to be reached and each link cost. In normal operation, that scheme converges and produces the shortest path from one node to another of the network. In a link state protocol a node broadcasts its routing information to all network nodes. In normal operation, every node will get network global topology and must perform an additional algorithm to deduce all network paths with the best cost (example of Djikstra most short path algorithm).

Routing information in ad hoc network are similar to those used in wired networks. But because wireless networks are less efficient than wired networks in bandwidth, new signalling routing techniques have been designed to improve the bandwidth. Those techniques are bonded to dynamic network topology characteristic and are used on demand, route is installed when a node need to send an information to another node. In most of the cases, this can solve the problem of expiration of a route install in advance when the network topology has changed. The routing signalling is transported into new kind of protocols such as reactive protocols that establish routes when applications need them.

When a node has a packet to send to another node of the network, it sends a route request through the network to obtain a route that reaches the search destination. This protocol family contains: AODV (Ad hoc On Distant Vector) protocol [31] [32] [33], DSR (Dynamic Source Routing) protocol [40] [41] [42] [43], TORA (Temporally Ordered Routing Algorithm) protocol [34] [35], ABR (Associatively Based Routing) protocol [36] which have being studied by the IETF MANET group.

Contrary to reactive protocols, ad hoc network proactive protocols convey periodic control packets through the network to update their network topology knowledge. Actual works on this family of protocols regroup: DSDV (Destination Sequenced Distant Vector) protocol [28], OLSR (Optimised Link State Routing) protocol [29] [30], FSR (Fisheye State Routing) protocols [37] are submitted to the IETF MANET group [38] [39].

8.3 Ad Hoc Network Routing Protocols Comparative Classification

Table 4 summarise the main characteristics of main ah hoc network routing protocols. All protocols avoid loops either by using sequence numbers, or by source routing paradigm such as envisaged in DSR protocol. AODV is too much similar to DSR except the utilisation of the source routing notion in DSR. However, DSR offers several possible paths to a given destination.

None of the presented protocols considers security management issues, very important in wireless networks.

None of presented protocols assume QoS. Nevertheless, OLSR tries to avoid bad unreliable links. Finally none of the presented protocols takes into account the management of nodes power, this leads to an unbalanced network if some routes are much more used than others: battery exhaustion and links overload.

8.4 Two Examples of Routing Protocols

DSR: is a dynamic routing information protocol based on route discovery and route maintenance. Route discovery takes place when a node has a packet to send to another but does not know the path to use. The original node sends a route request, intermediate node that does not have the search path add their address in the address list field, intermediate that have the path or the destination node return the complete path to the sender.

Table 4. Characteristics of ad hoc routing protocols

Criterion	AODV	DSR	OLSR	FSR	CBRP	LANMAR	TBRPF	ZRP
Without loop	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Many routes possible	No	Yes	No	No	Yes	No	No	Yes
Distributed	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kind	reactive	reactive	proactive	proactive	hybrid	hybrid	Proactive	hybrid
Security	No	No	No	No	No	No	No	No
Periodic messages control	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Unidirectional links	No	Yes	Yes	Yes	Yes	No	No	Yes

CBRP: Cluster Based Routing Protocol.
LANMAR: Landmark Routing Protocol.
ZRP: Zone Routing Protocol.
TBRPF: Topology Broadcast Based on Reverse – Path Forwarding.

Route maintenance mechanism uses acknowledgement, route error messages, packet recovering by modifying source route of the packet and sending route error to the original sender, fragmentation to adjust packets to path size.

DSDV: is based on Bellman Ford method as RIP (Routing Information Protocol). Each node has a table containing destination nodes, number of hops to reach each destination, next hop to send packet, sequence number and last update date. Tables are regularly updated so that DSDV generates an important overhead. Another problem of DSDV is that the mobility must be estimated to find an adapted update time, but there is very less work over mobility. Table size is another problem for large networks.

8.5 MAC Routing vs. Network Routing

Mobile ad hoc routing can be located at different levels. In fact, HIPERLAN recommend a MAC routing whereas MANET recommends network routing. ANANAS is in favour of an intermediate level.

HIPERLAN (High Performance Radio Local Area Network) type 1 is an ETSI standard for local area networks without neither base station nor fixed infrastructure. HIPERLAN defines its own physical layer and routing algorithm, situated at the MAC layer. This European standard has reserved the frequency band 5.2-5.35 GHz and offer throughput attaining 23 Mbps. HIPERLAN defines its own MAC layer called EY-NPMA (Elimination Yield Non-pre-emptive Priority Multiple Access). This layer offers a QoS with five priority levels. The physical layer uses GMSK (Gaussian Minimum Shift Keying) as a frequency modulation technique. The routing, done a level 2, uses the physical address of the wireless cards. The routing table is constructed on the base of periodically disseminated neighbouring and topological information. The control information are diffused by the technique multihop relaying with higher priority with regard to data. The multihop relaying technique consists in diffusing a packet in the network using only a sub set of nodes in the network in order to save the bandwidth.

Advantages: from higher layers the network is seen as a local diffusing network. IP applications run normally without any change, and the HIPERLAN is transparent.

Drawbacks dedicated HIPERLAN cards have to be used. Otherwise, the implementation of the MAC ad hoc routing requires the modification of the drivers of the wireless cards.

MANET (Mobile Ad hoc NETwork) is an IETF working group. This group aims to specify and standardise the mobile ad hoc networks protocols, at the IP level. Protocols have to support the physical and MAC heterogeneity. In other words, the protocols have to be independent from lower layers. This is rather different from HIPERLAN that specify and consolidate the lower layers of the system.

The MANET network is defined as a network of mobile and autonomous plateforms. Thos plate-forms may have several shared hosts and interfaces. Thos plateforms may move freely without any constraint and should work as an autonomous network and support links to fixed networks and gateways. The MANET networks have dynamic multihop topologies and a variable size, from some ten hundreds nodes.

The first objective of this group is to choose one or several unicast routing protocols and the interaction between high and low layers. Then to study the problems of QoS and of multicast in mobile MANET environment.

ANANAS (A New Ad hoc Network Architectural Scheme) is an architecture that locate the management and the calculation of the routing at an intermediate level between the physical routing level and the IP routing level. The architecture of ANANAS is divided in three levels:

The physical level: it is the set of the network physical interfaces. Those interfaces may belong to different physical networks, using different technologies.

The ad hoc level: this level constructs a unique logical ad hoc network composed of the different technologies used by the physical layer cards. Each node has a unique identifier (ad hoc address). The ad hoc routing protocol is implemented at this level and is based on the unique identifiers of each node. The routing is considered as a commutation between physical interfaces and networks.

The IP level: at this level the network is seen as a classical Ethernet local network. IP see only the virtual interface that masks physical architecture. Thus, all applications work normally without any modification. For example, the packets intended to the address 255.255.255.255 reach all the destinations present in the network without any IP routing (packets intended to this address are not relayed by the IP routers).

This architecture do not specify any particular routing protocol, theoretically, it is possible to use one the proposed MANET protocols with light modifications in the packets format and in translation schemes between IP addresses and ad hoc addresses.

9 Performance Comparison of Routing Protocols

Both studies presented in this part are relatively old (1998 and 2000) but are referenced in several publications [43, 48, 53]:

The first study, realised by J.Broch, D.Maltz, D.Johnson and Y.Hu, was supported by the Monarch project [50]. It is really the first performance evaluation study on different ad hoc routing protocols. This study presents the four routing protocols behaviour: DSDV (proactive), TORA, DSR et AODV (reactive). The algorithms were implemented in ns-2 and the used mobility model is Random Waypoint. The choice of parameters is interesting because several ulterior studies were based on those parameters. The simulations are based on CBR traffic sources in order to control more easily le network total load. The number of nodes was fixed to 50, the number of sources is variable (10, 20 and 30), and the simulation duration is 900 seconds. All links are bi-directional.

The studied metrics are:

- **The loss rate:** important because the retransmission of the data is managed at transport level and consequently can influence the maximum throughput that the network supports
- **The routing overhead:** must be minimised to optimise the band-width of the network. It is measured as a number of packets
- **The relevance of the path:** the difference between the path taken by the data and the existing shortest path between the source and the destination. This metric shows the capacity of the protocol to find most efficient paths in terms of a number of intermediate nodes.

The degree of mobility selected to compare the four protocols is the time of break (parameter of the model of mobility). In terms of rate of losses, AODV and DSR are most effective whatever the number of sources. Concerning the control traffic, DSR is far better than the three other protocols. However, it should be underlined that the overhead is measured in number of packets. However, the size of the DSR header is considerably important because it contains all the paths that the packet have to follow. Considering the load, AODV becomes better except for a high mobility (very short time of break). The most powerful protocols as for the relevance of the path are DSDV and DSR. Moreover, contrary to TORA and AODV, this metric is relatively independent of the degree of mobility. To summarise, the results are favourable to DSR protocol, and to a lesser extent to AODV protocol (among the four, they are the only ones for which standardisation is in progress).

This study was considered, later on, in many simulations. For this reason, it was discussed on several points. The first one, inevitably, the selected degree of mobility is not representative of the network dynamics. The control traffic would have been compared to data traffic transferred by the protocol. Indeed, a protocol can generate more overhead but transport more data. Furthermore, the ratio (control volume/data volume) would have been more judicious as in [51]. Finally, to define the relevance of the path, the criterion of distance is rather reducing (no result relating to the delay is presented).

The second study, carried out by S.Das, C.Perkins and E.Royer, is more recent and compares more in detail the performance of AODV and DSR protocols [51]. As in the simulations of the first study, the selected environment is NS-2. The sources are CBR sources and the mobility model is Random Waypoint. Here, two surfaces of different simulations (1500*300 and 2200*600) are used respectively accommodating 50 and 100 mobile units with a varying number of sources from 10 to 40. The studied performance criteria are the rate of losses (as in the Monarch study), the end to end delay and the ratio overhead on information (in a number of packets). The estimator of mobility is always the duration of breaks. The links are bi-directional as in the preceding study. The two protocols are first tested on the configuration of 50 nodes:

- **The loss rate:** With 10 and 20 sources, both protocols are equivalent. AODV becomes more powerful with a more significant number of sources (30 and 40): DSR looses between 30 and 50 % of packets more than AODV for high mobility (short break time).
- **End to end delay:** The DSR is faster with 10 sources (a factor 4 compared to AODV) and 20 sources (but with a tighter difference). Beyond 30 sources, the tendency is reversed: AODV is better, in particular with short break times (AODV is twice faster).
- **Overhead/data:** DSR appears broad better in all the cases (at least 4 times better). When the number of sources increases, DSR preserves a stable ratio relatively to AODV even if its other criteria worsen. It as should be noticed that the delays of the two protocols are very long with 40 sources, even with low mobility. This is explained by the absence of load balancing mechanisms and by the policy of selection of the paths (for example, in the DSR, the shortest path in a number of nodes is always the optimal way).

With 100 mobile units, the results concerning the losses rate and the end to end delay are similar with those of the configuration of 50 stations. On the other hand, the difference between the ratio Overhead/data is weaker. Moreover, the ratio DSR ratio is no more stable.

In summary, let us notice that AODV protocol is better than DSR, as soon as the number of sources increases, in term of delay and losses rate. Even if the DSR remains better in term of overhead, these values were taken in a number of packets without taking into account their size. This behaviour can be explained because, in ad hoc networks, the access to the medium is much more expensive than to add some bytes with an already existing packet.

The behaviour of the two protocols is then observed while varying the load (from 0 to 800 Kbits/s) and by considering only 10 sources. The examined metrics are the end to end delay, the reception throughput and the overhead (in Kbits/s). It is noticed that the DSR throughput saturates from 325 Kbits/s because of a high losses rate while the AODV threshold is at 700Kbit/s. The time of the DSR is much better with weak load (at least a factor 2) but longer with high load. The overhead of both protocols are equivalent with a small advantage to the DSR. By taking 40 sources, the behaviours are identical but the thresholds of saturation are weaker (150 for the DSR and 300 for AODV) and the overhead of both protocols becomes higher than their reception throughput!

This study shows the problem of the two protocols, in particular the DSR, when the number of sources and the load of the network increase.

10 IP over GEO Satellite Links

This section presents in a few words performance issues of IP in a specific context; geostationary satellites. But lets first settle the context of IP communication and GEO satellite, before focusing on performance questions: why IP over satellite?

Even if service integration is not a brand new idea, it is still an area of research, since ADSL and cable operators fight to deliver over a same support access to the web, voice and television (Free, Tiscali…) at lower prices. All these services are based over high bandwidth medium and integrated via IP, which is unavoidable today.

As GEO satellites experience a new success with digital television, numerous projects study it as a medium for IP (IBIS [52], GEOCAST [53], DIPCAST [54] …). Indeed its wide geographical coverage, its broadcasting nature, its unique capacity to feed remote areas and its high bandwidth are combined advantages contributing to its relevance for service integration. Therefore, satellite seems naturally an interesting support for IP, and more precisely for multicast or broadcast IP services.

However, high costs of satellite deployment and the standardisation issue have highly impaired investments. Our participation in a French project, DIPCAST (DVB for IP multicast, RNRT project), has leaded us to propose several solutions: architectures based on bent-pipe system, architectures with on-board processing, architectures based on a hybrid satellite (relevant to conventional bent-pipe systems as to the next-generation of GEO satellites).

In order to compare and evaluate IP over DVB architectures, performance studies are unavoidable. Nevertheless these studies are not quite simple, since satellite systems are really different from terrestrial networks. Indeed GEO satellite natural delays are so high that all over delay could be overlapped. Moreover access methods are difficult to evaluate since they are often private and specific to the system. A global standardisation issue in satellite domain has highly impaired the development of a real environment of performance evaluation.

In this short state of the art, we propose to present what are the standard IP over GEO satellite performance issues and how they are approached.

11 Satellite Performance Issues

11.1 Services over Satellite

One issue of communication over satellite is the service relevance and adaptation to such a system. Applications are designed, modified, adapted to be supported by the satellite media. In fact satellite specificity, advantages as drawbacks, imply a real study of services, more precisely real-time services and multimedia applications, and thus performance studies and QoS management [55].

11.2 TCP Studies and Transport Protocols

Since it has been a long time that satellites are used as backbone links, the question of TCP performance is a topic as old as relevant. Indeed TCP performance over satellite links known as bad [56] [57] since TCP is not relevant to:

Satellite high delay (long round trip time) impaired throughput usable as illustrated in figure 9. With GEO satellites, the propagation delay is thus on the order of 0.5 seconds. A total RTT value of 0.55 s would be appropriate [58]. The TCP sender must wait this length of time to receive ACKs, which is going to slow the throughput, hurting also interactive applications as well as TCP congestion control algorithms.

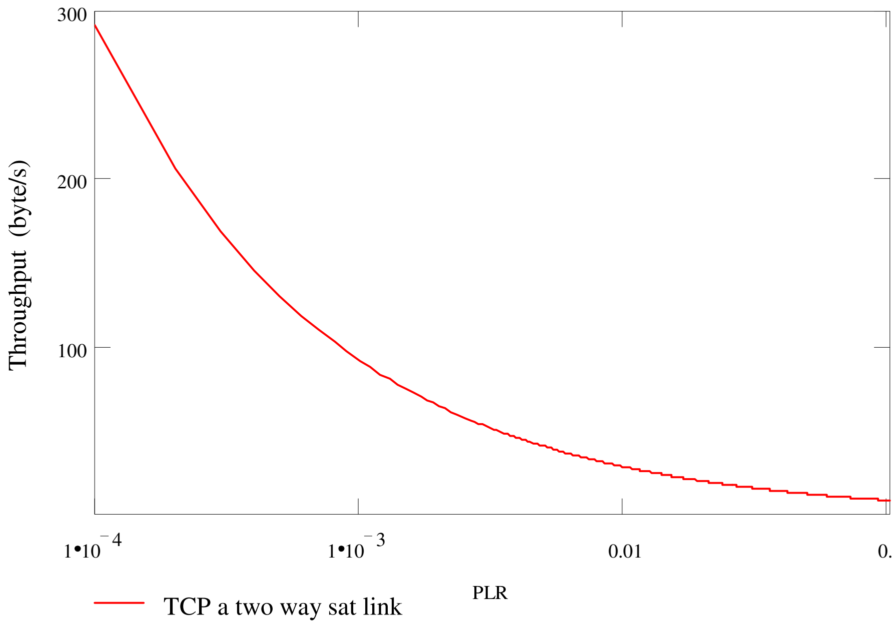

Fig. 9. TCP throughput over a symmetric satellite link

Transmission errors are more important in satellite channels than typical terrestrial network. Because TCP uses all packet drops as a congestion signal, its performance on satellite systems are decreased.

Asymmetric return link which are commonly used since end-to-end users cannot generally have their gateway. That may have an impact on TCP performance.

In this context many propositions have been made to fill TCP limits over satellite. Solutions and their studies are then a large work underway in IP over satellite domain. So many studies carry on focussing on performance evaluation of new solutions compared to old solutions. Here we present a panel of these solutions divided into two categories, external mechanisms and TCP improvements.

TCP improvements:

– CANIT algorithm in a satellite context in order to improve TCP congestion avoidance issue in satellite system [59].
– TCP-Proxy which enables full link throughput per connection [60].
– Fast retransmit.
– Fast recovery.
– SACK option.

External mechanisms:

- Path Maximum Transfer Unit Discovery [58] allows TCP to use the largest packet size without any risk of fragmentation. However its performances are criticised since it implies a longer delay before sending data.

- Forward Error Correction enables to lower PLR and thus to increase TCP maximum throughput. It seems a real opportunity nowadays [61] [62] but it is also used in other transport protocols [63] [64].
- Explicit Congestion Notification [65].

11.3 Access Mechanisms

Another general performance issue in IP over satellite is the access mechanisms and their studies. Indeed if the transmission delay in a satellite system is generally short compared to RTT, the connection to the media could induce an important and variable delay. Therefore many studies compare different access mechanisms and try to optimise them.

Such studies may integrate transport protocol issues [66].

11.4 Other Issues

There are many performance studies which seem interesting in the satellite domain. In a global architecture design, performance issues are numerous and must be considered step by step. For example in IP architecture over a next generation GEO satellite, On-board-processing performance has to be evaluated (switching, table updating…), then access mechanisms [67] and overheads. Indeed, overhead in IP over satellite is a common topic since in DVB solutions the common solution, MultiProtocol Encapsulation, has several limits [68] and satellite bandwidth remains precious. Moreover, IP encapsulation has to be rightly chosen as it often implies signalling protocols, which may impair routing schemes. Eventually, a global comparison between the new architecture and standard bent-pipes architecture have to be done, so as to validate routing schemes, on-board-processing interests…

12 Conclusions - Part 2

The distribution of the band-width in these Mobile Ad Hoc NETworks (MANETs) and GEO satellite can be very unfair. While the behaviour of TCP remains to be studied more in details in MANETs, it is necessary in addition to think about the introduction of quality of service mechanisms according to the routing between various users and according to the constraints of each one. A very dependent subject is the wireless Ad Hoc routing, in particular protocols such as DSR, AODV, OLSR..., which are conceived for the local area networks, but it is very interesting to study the impact for even larger networks (in term of space and of number of nodes). Many routing algorithms are proposed and standardisation is in process. However, QoS mechanisms at different levels and their interdependence is an open issue. Many interesting subjects, such as: (i) Scheduling and consensus algorithms for QoS in Ad Hoc networks. (ii) New architectures of Ad Hoc networks locate the management and the calculation of the routing in an intermediate level between the routing at the physical level and the routing on IP level.

The QoS mechanisms are very dependent on the underlying protocol layers. In particular, if we look at the evolutions of the standards of the IEEE 802.11 family, it appears that some mechanisms are now based on traffic differentiation. These mechanisms are not yet completely standardised. In addition, the physical layer in these types of networks consists in using a radio support which by nature will have a very variable quality in space and in time.

References

1. Chiang, M.: To layer or Not To Layer: Balancing Transport and Physical Layers in Wireless Multihop Networks. In: IEEE INFOCOM 2004 (2004) (to appear)
2. Yuen, W.H., Lee, H.-n., Andersen, T.D.: A Simple and Effective Cross Layer Networking System for Mobile Ad Hoc Networks. In: IEEE PIMRC 2002, pp. 1952–1956 (2002)
3. Raisinghani, V.T., Singh, A.K., Iyer, S.: Improving TCP performance over Mobile Wireless Environments using Cross Layer Feedback. In: IEEE ICPWC 2002, pp. 81–85 (2002)
4. Safwat, A., Hassanein, H., Mouftah, H.: Optimal Cross Layer Designs for Energy Efficient Wireless Ad-hoc and Sensor Networks. In: IEEE IPCCC 2003, pp. 123–128 (2003)
5. Shakkotai, S., Rappoaport, T.S., Karlson, P.C.: Cross Layer Design for Wireless Networks. IEEE Communiction Magazine, 74–80 (October 2003)
6. Conti, M., Maselli, G., Turi, G., Giordano, S.: Cross layering in mobile Ad Hoc Network Design, pp. 48–51. IEEE Computer Society, Los Alamitos (2004)
7. Kawadia, V., Kumar, P.R.: A Cautionary Perspective on Cross Layer Design. IEEE Wireless Communication Marazine (2004)
8. Lee, K.-H.: A Multiple Access Collision Avoidance Protocol for Multicast Services in Mobile Ad Hoc Networks. IEEE Communications Letters 7(10) (October 2003)
9. Gauthier, V.: Cross layer et QoS dans les réseaux ad hoc sans fil. GET-INT-RR (2004)
10. Kaixin Xu, M.G., Bae, S.: Effectiveness or RTS/CTS handshake in IEEE 802.11 based ad hoc networks. In: Globecom 2002 (2002)
11. Khayam, S.A., Karande, S., Krappel, M., Radha, H.: Cross layer protocol design for real-time multimedia applications over 802.11b networks. In: IEEE ICME 2003, pp. II-425–II-428 (2003)
12. Shan, Y., Zakhor, A.: Cross layer techniques for adaptive video streaming over wireless networks. In: IEEE ICME 2002 (2002)
13. Chen, J., Hsia: Joint cross layer Design for Wireless QoS Video Delivery. In: IEEE ICME 2003, pp. I-197–I-200 (2003)
14. Sachs, D.G., Adve, S.V., Jones, D.L.: Cross layer adaptive video coding to reduce energy on general purpose processors. In: IEEE 2003, pp. III-109–III-112 (2003)
15. Alonso, L., Agusti, R.: Automatic Rate Adaptation and Energy-Saving Mechanisms Based on Cross-Layer Information for Packet-Switched Data Networks. IEEE Radio Communications, S15–S19 (March 2004)
16. Raman, B., Bhagwat, P., Seshan, S.: Arguments for Cross Layer Optimisations in Bluetooth Scatternets. In: SAINT 2001, pp. 176–184. IEEE Computer Society, Los Alamitos (2001)
17. Li, X., Bao-yu, Z.: Study on Cross-layer Design and Power Conservation in Ad Hoc Network. In: IEEE PDCAT 2003, pp. 324–328 (2003)
18. Shan, Y., Zakhor, A.: Cross layer techniques for adaptive video streaming over wireless networks. In: IEEE ICME 2002, pp. 277–280 (2002)
19. Toumpis, S., Goldsmith, A.J.: Performance, optimisation, and Cross-Layer Design of Media Access Protocols for Wireless Ad Hoc Networks. In: IEEE ICC 2003, pp. 2234–2240 (2003)

20. Zhao, J., Guo, Z., Zhu, W.: Power Efficiency in IEEE 802.11a WLAN with Cross-Layer Adaptation. In: IEEE ICC 2003, p. 2030 (2003)
21. Maharshi, A., Tong, L.: Corss layer Designs of Multichannel Reservation MAC under Rayleigh Fading. IEEE Transactions on Signal Processing 51(8) (2003)
22. Wang, Q., Abu-Rgheff, M.A.: Cross layer signalling for next-generation wireless systems. In: IEEE WCNC 2003, pp. 1084–1089 (2003)
23. Fang, Y., Bruce McDonald, A.: Cross layer performance effects of path coupling in wireless ad-hoc networks: power and throughput implications of IEEE 802.11 MAC, pp. 281–290 (2002)
24. Becker, M., Beylot, A.-L., Dhaou, R.: Aggregation Methods for Performance Evaluation of Computer and Communication Networks. In: International Symposium: Performance Evaluation - Stories and Perspectives, Vienna, Austria (December 2003)
25. Courtois, P.-J.: On time and space decomposition of complex structures. Communication of the ACM 28(6) (June 1985)
26. Jain, R.: The art of computer performance analysis: Techniques for Experimental Design, Measurement, Simulation and Modelling. Wiley, Chichester (1991)
27. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object Oriented Software. Addison Wesley, Massachusetts (1994)
28. Perkins, C., Bhagwat, P.: Destination Sequenced Distance Vector (DSDV). Internet – Draft, IETF MANET Working Group (November 2001)
29. Jacquet, P., Clausen, T.: Optimized Link State Routing Protocol (OLSR). Internet – Draft, IETF MANET Working Group (October 2001)
30. Laouti, A., Muhlethaler, P., Najid, A., Plakoo, E.: Simulation Results of the OLSR Routing Protocol for Wireless Network. INRIA Rocquencourt (April 1998), http://menetou.inria.fr/~muhletha/medhoc.pdf
31. Royer, E.M., Perkins, C.E.: Multicast Operation of the Ad hoc On Demand Distance Vector Routing Protocol (August 2000)
32. Das, S., Perkins, C., Royer, E.: Ad hoc On Demand Distance Vector Routing (AODV). Internet – Draft, IETF MANET Working Group (November 2001)
33. Royer, E.M., Perkins, C.E.: Ad hoc On Demand Distance Vector Routing (AODV). In: Second Annual IEEE Workshop on Mobile Computing Systems and Applications, February 1999, pp. 90–100 (1999)
34. Corson, S., Vincent, P.: Temporally Ordered Routing Algorithm (TORA). Internet – Draft, IETF MANET Working Group (August 2001)
35. A performance comparison of TORA and ideal Link State Routing (February 1997), http://tonnant.itd.nrl.navy.mil/tora/tora_sim.html
36. Zinin, A.: Alternative OSPF ABR Implementations. Internet – Draft, IETF MANET Working Group (February 2001)
37. Gerla, M., Hong, X.: Alternative OSPF ABR Implementations. Internet – Draft, IETF MANET Working Group (December 2001)
38. Corson, S., Macker, J.: Mobile Ad hoc Networking (MANET): Routing Protocol Performance Issues and Evaluation Considerations (Janvier 1999)
39. Mobile Ad hoc Network (MANET), http://www.ietf.org/
40. Jetcheva, J.G., Hu, Y., Johnson, D., Maltz, D.: The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR). Internet Draft, IETF MANET working group (November 2001)
41. Johnson, D.B., Maltz, D.A.: Dynamic Source Routing in Ad hoc Wireless Networks (Mars 1998), http://www.monarch.cs.cmu.edu/monarch-papers/

42. Demir, T.: Simulation of Ad hoc Networks with DSR Protocol (May 2001), http://netlab.boun.edu.tr/papers/Iscis2001-DSR-TamerDEMIR+.pdf
43. The Dynamic Source Routing Protocol for Mobile Ad hoc Networks (DSR) (February 21, 2002), http://www.ietf.org/proceedings/02mar/I-D/draft-ietf-manet-dsr-07.txt
44. Farid, T.: Les modèles de mobilité dans les réseaux ad hoc. présentation interne
45. Ndoye, R.: Qos dans les réseaux mobiles ad hoc. présentation interne
46. Réseaux locaux et Internet, des protocoles à l'interconnexion, Laurent Toutain, Editions HERMES, série réseaux et télécommunications (1996)
47. RFC 2328, OSPF version 2 (April 1998), http://www.ietf.org/rfc/rfc2328.txt?number=2328
48. The Network Simulator – ns2: links to help getting started, http://www.isi.edu/nsnam/ns
49. The Network Simulator: Building ns, http://www.isi.edu/nsnam/ns/ns-build.htm
50. Broch, J., Maltz, D., Johnson, D.B., Hu, Y.-C., Jetcheva, J.: A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols. In: Proc. ACM/IEEE International Conference on Mobile Computing and Networking (Mobicom 1998), pp. 85–97 (1998)
51. Das, S.R., Perkins, C.E., Royer, E.M.: Performance Comparison of Two On-demand Routing Protocols for Ad Hoc Networks. IEEE Personal Communications 8(1), 16–28 (2001)
52. Chacón, S., Casas, J.L., Cal, A., Rey, R., Prat, J., de la Plaza, J., Monzat, G., Carrere, P., Miguel, C., Ruiz, F.J.: Networking over the IBIS System. IST Mobile & Wireless Telecommunications Summit (June 2002)
53. sun, Z., Howarth, M.P., Cruickshank, H., Iyengar, S., Claverotte, L.: Networking Issues in IP Multicast over Satellite. International Journal of Satellite Communications and Networking (2003)
54. DIPCAST, "le Dvb comme support de l'IP multiCAst par SaTellite", RNRT project, http://www.dipcast-satellite.com
55. Mobasseri, M., Leung, V.C.M.: A new buffer management scheme for multimedia terminals in broadband satellite networks. In: Proc. 35th Hawaii International Conference on System Sciences, Hawaii (January 2002)
56. Akyildiz, I.F., Jeong, S.H.: Satellite ATM networks: A survey. IEEE Communications Magazine (1997)
57. Roddy, D.: Satellite communications, 3rd edn. (2001), ISBN 0-07-137176-1
58. Allman, M., Glover, D., Sanchez, L.: Enhancing TCP over Satellite Channels using Standard Mechanisms. RFC 2488 (January 1999)
59. Benadoud, H., Berqia, A., Mikou, N.: Enhancing TCP over satellite links using CANIT algorithm. International Journal of SIMULATION 3(3-4), 81–91 (2002)
60. Wu, L., Peng, F., Leung, V.C.M.: Dynamic Congestion Control for Satellite Networks Employing TCP Performance Enhancement Proxies. To be presented at IEEE PIMRC 2004, Barcelona, Spain (September 2004)
61. IETF Reliable Multicast Transport Group, http://www.ietf.org/rmt-charter.html,motlabs.com
62. Celandroni, N., Ferro, E., Potorti, F.: Goodput optimisation of long-lived TCP connections in a rain faded satellite channel. VTC (Spring 2004)
63. Nonnenmacher, J., Biersack, E.W., Towsley, D.: Parity-based loss recovery for reliable multicast transmission Networking. In: Proceedings of ACM SIGCOMM 1997 (1997)

64. Byers, H., Luby, M., Mitzenmacher, M., Rege, A.: A Digital Fountain Approach to Reliable Distribution of Bulk Data. In: Proceedings of ACM SIGCOMM 1998, Vancouver, BC (1998)
65. Ramakrishnan, K.K., Floyd, S.: A proposal to add explicit congestion notification (ecn) to IP. RFC 2481 (1999)
66. Li, Y., Jiang, Z., Leung, V.C.M.: Performance evaluations of PRR-CFDAMA for TCP traffic over geosynchronous satellite links. In: Proc. IEEE WCNC 2002, Orlando, FL (March 2002)
67. Acar, G., Rosenberg, C.: Performance Study of End-to-End Resource Management in ATM Geostationary Satellite Networks with On-Board Processing. Space Communications Journal 17(1-3), 89–106 (2001)
68. IETF working-group IP/DVB, `http://www.erg.abdn.ac.uk/ip-dvb/`