Chapter 7
Hybrid Broadcast Broadband for the Delivery of 3D Video

Asimakis Lykourgiotis, Tasos Dagiuklas, Ilias Politis, Hugo Marques, Jonathan Rodríguez, and Hélio Silva

Abstract Fuelled by the need to increase their market share, broadcast service providers enhance their services with media and social networking. Hybrid broadcast broadband (HBB) television (HbbTV) has emerged as a promising means of content distribution for enriched and interactive television services. Specifically, the convergence of terrestrial digital video broadcasting (DVB) services and peer-to-peer (P2P) overlay networks is an up-and-coming telecommunication design. In this chapter, we discuss how this vision came into fruition, within the context of the European-funded project ROMEO. ROMEO project aims at delivering three-dimensional (3D) multi-view content to fixed, mobile and portable users with guaranteed minimum quality of experience (QoE). In particular, we elaborate on the P2P overlay construction process as well as on techniques to optimise its performance. Moreover, we present a packetisation scheme that facilitates the task of synchronising the streams that are being received from DVB and P2P networks. Additionally, ROMEO architecture supports mobility by exploiting the evolved packet core design and enhancing content delivery, particularly for mobile devices. Finally, we specify the life cycle of the ROMEO content from a QoE perspective from core network to access network including scenarios involving multiple internet service providers.

A. Lykourgiotis (✉) • I. Politis
Hellenic Open University, Parodos Aristotelous 18, Patras 26335, Greece
e-mail: alykourgiotis@eap.gr; ilpolitis@eap.gr

T. Dagiuklas
London South Bank University, 103 Borough Road, London, SE1 0AA, UK
e-mail: tdagiuklas@lsbu.ac.uk

H. Marques • J. Rodríguez • H. Silva
Instituto de Telecomunicações, Campus Universitário Santiago, Aveiro, Portugal
e-mail: hugo.marques@av.it.pt; jonathan@av.it.pt; heliosilva@av.it.pt

7.1   Introduction

We are moving to a new era of convergence between media consumer and content provider. This shift is due to the new methods of media consumption via sharing using technologies such as converged heterogeneous networks, new transport methods (e.g. peer-to-peer—P2P, multiple multicast trees—MMT, etc.), personal/user generated tools and social media [15]. As our lifestyles become more connected, even the passive behaviour of watching television is turning into a very active process, where viewers are multitasking on their mobile devices while watching television. This shift poses new challenges in jointly optimising media networking and sharing personalised content. Social TV refers to the convergence between broadcasting, media networking and social networking. This gives the capability to the end-users to communicate, publish and share content among themselves and interact with the TV program. Consumers have aggressively adopted online video services (e.g. Netflix, YouTube, etc.). As more providers, more content and more devices become available, consumers seem ready to take full advantage. More consumers also expect to see increased use of video on laptops, tablets and smartphones than on any other devices [5].

Looking from users' perspective, uploading their own content while enjoying the commonly shared content, as well as interacting either with the content which is being displayed or with their colleagues in the socially collaborating group, is an important parameter to make the whole social interaction truly enjoyable. A user on a mobile device (e.g. PDA, Tablet, etc.) should be able to access the content from the closest point (distance, signal strength and availability) where the service is available. This particular service point is expected to also provide media processing capabilities to match the needs of the mobile terminal by pre-processing the available content in accordance to the given device capabilities, e.g. processing capability, resolution, content format it supports, etc. This could be supported by media storage and processing clouds that either the users will subscribe to (as an additional service) or their mobile network operator will provide it (within the subscription package) for increased user QoE and hence increase the network's customer base. The TV experience is enhanced by additional services.

As demonstrated by the introduction of HDTV through satellite and digital terrestrial television , broadcast TV networks have pioneered the transition to high quality TV services and consistently play a leading role in the introduction of new TV standards. However, broadcast networks do not benefit from the native return path, which is required to deliver advanced TV services. Broadband operators can deliver on-demand content and full-scale interactivity but their geographic reach for TV services is limited. A number of standards, which can be natively implemented in connected TVs or stand-alone devices, are also available for hybrid broadcast broadband (HBB) TV integration [7]. HbbTV combines the advantages of broadband for delivering individual choice of on-demand content with the efficiency of broadcasting (satellite, digital terrestrial) for making high quality TV

simultaneously available to a large audience. Hybrid TV solutions also require very short time-to-market implementation of linear and non-linear TV services for consumers. The European Commission funded integrated research project ROMEO [18] introduces a novel architecture for HBB networks in order to support 3D multi-view content [14]. In this chapter, the architecture and main functional entities of the ROMEO platform are being presented, along with a discussion of the life cycle of the multimedia content.

The remainder of this chapter is as follows. Initially, in Sect. 7.2 we present an overview of the ROMEO concept, briefly describing the architecture and the basic supported use cases of the ROMEO platform. Subsequently, in Sect. 7.3, we elaborate on the different modules that compose the server and the peers and enable the timely 3D content delivery and consumption. Thereafter, in Sect. 7.4, we describe how the 3D video content is generated as well as the content packetisation scheme that enables synchronisation between the broadcast and the broadband access network. Afterwards, in Sect. 7.5, we elaborate on the life cycle of the content from a QoE perspective from the core network to the access network and the fixed and mobile users. Finally, in Sect. 7.6, we conclude the chapter.

7.2 Hybrid Architecture: The ROMEO Concept

ROMEO project aims at delivering 3D multi-view content to fixed, mobile and portable users with guaranteed minimum QoE. To achieve this, DVB [8] technology is combined with the P2P Internet technology to take advantage of both DVB and Internet Protocol-based (IP) networks. DVB provides a reliable and deterministic content delivery path and by exploiting this characteristic of DVB, ROMEO users retrieve the content with a quality of DVB stream in the worst case. On the other hand, IP network allows delivering multiple high quality views, which cannot be delivered via DVB due to its bandwidth limitations. As a result, the broadcaster can deliver high quality stereoscopic 3D content to end-users via DVB and at the same time stream a set of supplementary 3D multi-view content (e.g. additional viewpoints with their respective spatial audio), through the P2P network. ROMEO also exploits and augments the evolved packet core (EPC) [16] functionalities to deliver 3D multi-view content to mobile users. As another capability, the remote users are provided with a real-time audio-visual (A/V) communication channel so that they can share their experiences while watching the high quality 3D media. These collaborating users can also share their own content between each other.

In Fig. 7.1, an overview of ROMEO services is provided. As can be seen, a user is connected to the ROMEO system as a member of P2P topology (MMT). ROMEO topology is managed by the central ROMEO server, so the user must be connected to the server as the first step to join the ROMEO system. ROMEO server is also responsible for distributing the P2P streams corresponding to multiple view 3D video and audio to the whole ROMEO system via super-peers. Delivered streams are then delivered through the MMT [3] and peers receive these streams from their
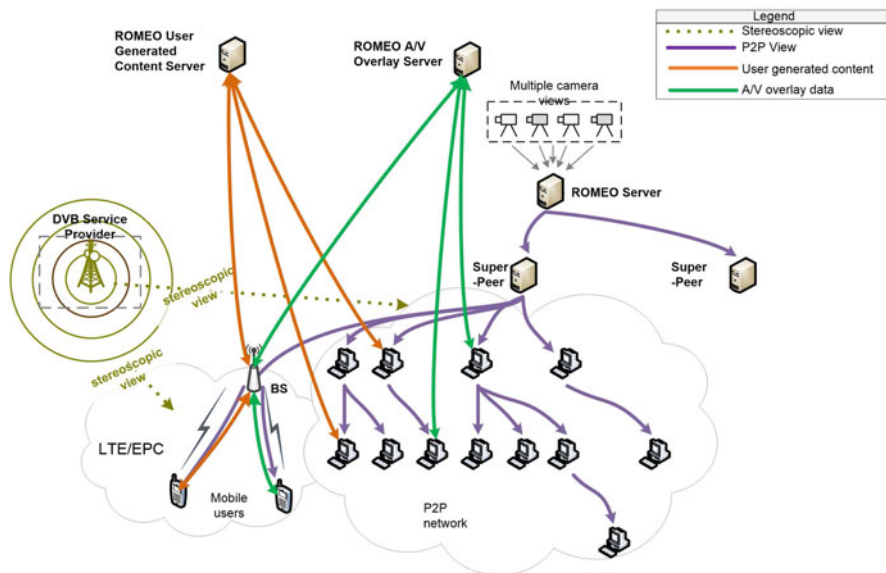
Fig. 7.1 ROMEO architecture

parents. ROMEO P2P streams are shown as violet in Fig. 7.1. It is assumed that each user will receive main stereoscopic view from DVB, which is shown as beige in Fig. 7.1. This stereoscopic view is then combined and synchronised with the additional views received from P2P to provide users high quality 3D multi-view experience. The assumption of DVB availability for each of the users guarantees the seamless content consumption in case of deteriorations in the network conditions.

During the consumption of the live stream, ROMEO users can collaborate and share their own content as mentioned before. In order to minimise the impact of these services (such as transfer of the user generated content—UGC) on the P2P live streaming media during the collaboration, UGC (shown as orange in Fig. 7.1) and A/V overlay (shown as green in Fig. 7.1) data paths are separated from the P2P delivery path with an independent client–server-based architecture.

The incorporation of DVB and IP service providers at the same time requires the coordination between the content received from these two providers. An operator having both DVB- and IP-based content services is an ideal business model for providing the ROMEO services which may not be the case for most of the time. As an alternative business model, virtual media operators would emerge as a new entity that will provide media services to end-users by contracting with companies providing broadcast and broadband services. With the development of ROMEO and similar hybrid media delivery platforms, these virtual operators may come into play to provide services exploiting hybrid delivery technologies in near future.

To sum up, the ROMEO system mainly involves two use cases. The first use case scenario is the consumption of 3D audio/video over hybrid delivery paths. In this

scenario, the hybrid network is delivering a 3D live stream using either only one of the available paths or a combination of them (i.e. DVB and P2P over IP) at a given time to maximise the level of QoE offered to the user, including adaptation in response to varying network conditions, triggering different media delivery options (adjusting number of views, quality of views, etc.). The second use case scenario incorporates the collaborating group communication and sharing of UGC. In this scenario, the hybrid network is providing a real-time audio-visual communication channel to connect collaborating users while they are receiving the live stream via their DVB and/or P2P connections. In this scope, users can initiate a collaborating group and specify the members of the group. Users can generate or use existing content and share it with the members of the collaborating group for further social interaction.

7.3   A Closer Look at ROMEO Building Blocks

The ROMEO approach for network delivery of synchronised 3D media networking with real-time audio-visual communication support involves a variety of building blocks. Initially, there is the main server, where the ROMEO 3D media content is originated and stored. It belongs to a specific DNS domain (typically associated with the service brand) and it is property of the service owner. Additionally, UGC server is responsible for storing and managing UGC, that is, content created by ROMEO users. Moreover, there is the multiple coding server, which converts UGC to multiple file formats for compatibility and quality of experience (QoE) reasons. The A/V overlay server provides a real-time audio-visual communication channel which connects all collaborating users. Furthermore, there are the super-peers, that are proxies/replicas of the main server and are placed at the premises of every internet service provider (ISP) that has an agreement with the service owner. The super-peers are responsible to serve peers from a specific geographical area or ISP. For large scale deployments, ROMEO super-peer nodes, which are also shown in Fig. 7.1, behave as the replicas of the ROMEO server. Lastly, there are the peers, which are fixed, portable or mobile devices belonging to the ROMEO end-users, who will consume the ROMEO 3D media content.

Figure 7.2 shows the ROMEO enabler components which provide network agnostic behaviour to the ROMEO system. When a user wants to join the ROMEO system, its device and network capabilities are queried by the network monitoring subsystem (NMS) component (Sect. 7.3.2.5). The topology builder (TB) component inserts the peer in the most suitable position in the MMT topology by taking these network monitoring reports into consideration (Sect. 7.3.1.1). During run-time, the NMS peer module periodically reports the conditions of each peer to the multicast tree manager (MTM) server module, presented in Sect. 7.3.1.2, which is intrinsically related to the TB. By using the updated peer records and a complex algorithm, the TB decides whether it should promote or demote peers from their current position on each tree. Chunk selection (CS), described in Sect. 7.3.2.3, and media aware proxy (MAP), defined in Sect. 7.5.3, will perform network adaptation for the peers in case
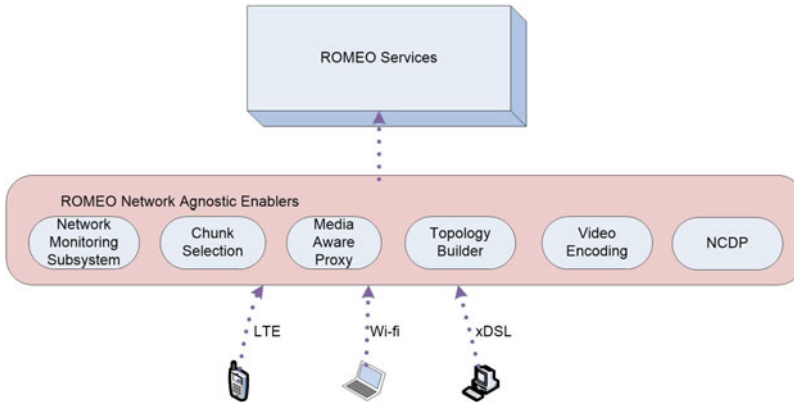
Fig. 7.2 ROMEO heterogeneous networks solution

of having insufficient bandwidth. In this way, it is aimed to provide service to each peer according to its device and network capabilities regardless of its network access method. This capability requires having suitable encoding scheme for the video to allow stream elimination which is implemented by video encoding component as specified in Sect. 7.4. In the following subsections, the components of these building blocks are analysed.

7.3.1 ROMEO Server

The three main building blocks of the ROMEO server are the topology builder, the MTM and the P2P transmitter. In the following sections, we briefly described these modules and their functionalities.

7.3.1.1 Topology Builder

The TB is a software module mainly responsible for creating MMT that effectively construct the P2P overlay network used for the distribution of the multiple streams associated with the visualisation of 3D content. In particular TB performs the following functions at the main server/super-peer. It listens for new peer connection requests. Also, it acts as an authentication proxy (authenticator) for user authentication with the main server. Moreover, it collects network monitoring data from each of the peers. Additionally, it creates multiple P2P application-level multicast trees for content distribution. Finally, TB computes peer insertion on the P2P application-level multicast trees.

When a peer is redirected to a super-peer, it is the responsibility of the TB to compute the peer position at the P2P multicast tree at the access network level.

Initialising the computation process, the TB groups peers according to the requested content. Next, it groups peers according to their proximity to edge routers (ER) geographical aggregation. Finally, it sorts the peers by evaluation.

After grouping and sorting operations, the multiple P2P multicast trees are computed, one for each requested content and ER. Traffic at the ISP core will be transmitted using IP multicast and it will be the ER's responsibility to map each requested content multicast address to specific parent(s) IP addresses(s)—the ER will effectively act as a replicator. To optimise the ER resources, the super-peer predetermines how many top level peers (parents) can be directly fed by one ER. This means that, when constructing each P2P tree, the super-peer arranges a predetermined number of highest ranking peers at the top of the tree and delegates in these the distribution of content to other peers in the same access network. Every time a peer is selected to forward content, its resources are diminished, and if its evaluation becomes lower than other peers, the new highest ranking peer will take the role of parent for subsequent content streaming.

To minimise the issues associated with peers joining and leaving the system, also known as churn, the TB uses the grounding, graceful leaving and redundancy mechanisms. The grounding technique inserts new joining peers always at the bottom of the P2P tree. The algorithm is executed periodically to maximise the efficiency of the P2P tree-promoting and demoting peers. In the graceful leaving concept, the peers, whenever possible, inform the TB about their imminent disconnection, and only stop forwarding content to their children when instructed by the TB or upon a time-out. Lastly, in the redundancy scheme, all peers will be informed of their active parent and a backup parent when inserted on a tree. If the active parent is not reachable within a time-out the peer switches to the backup parent and informs the TB.

The TB specification brings resiliency, scalability and higher performance to ROMEO platform. Resiliency because by using tree separation, a major fault in a specific part of the network will not affect other parts. By grouping peers by common ER, tree depth is significantly reduced, since peers sharing the same access network have improved downstream and upstream bandwidth, which allows more children per parent and thus supporting higher scalability. Finally, improved performance is achieved when the total number of hops between top level parents and their children is significantly lower, which contributes to reduce the packet/chunk delay, jitter and packet loss. Recovery from minor faults (such as peer churn) can also be achieved in a faster way, since the alternative link (backup parent) is on the same access network and ready to start forwarding packets immediately.

7.3.1.2 Multicast Tree Manager

The MTM also runs at the main server/super-peer and it is intrinsically related with the TB operations. MTM collects/aggregates network monitoring data (percentage of packet loss, average delay, jitter and available bandwidth), from all connected peers, provides the TB with peers' updated network conditions. Furthermore, MTM
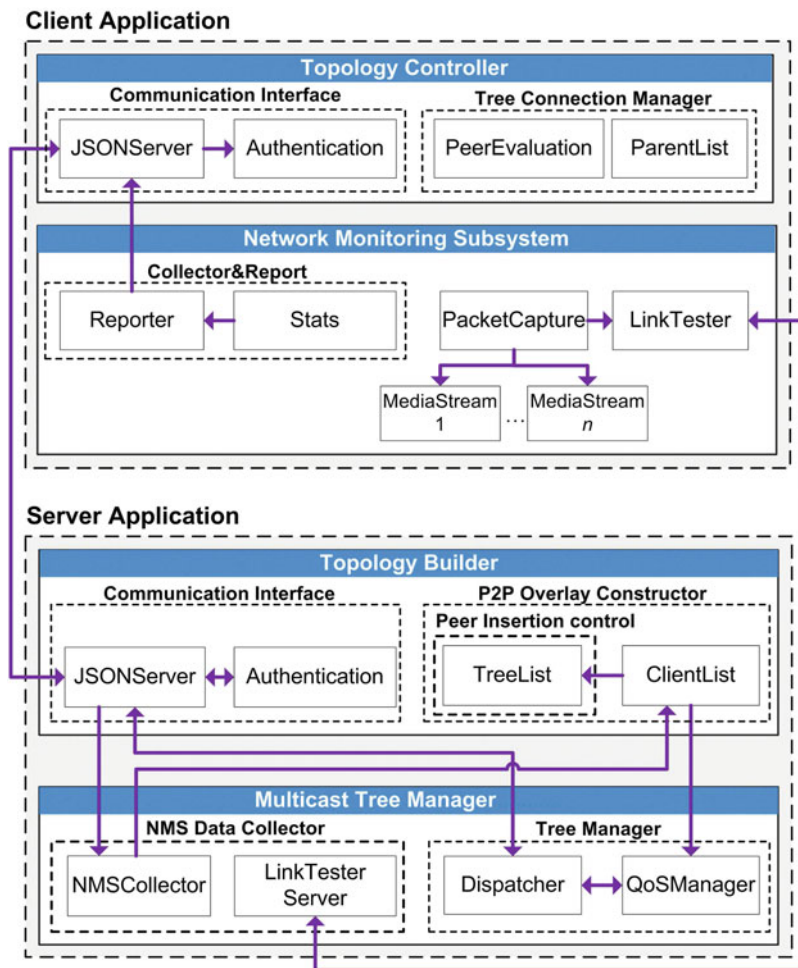
Client Application

Fig. 7.3 Client (peer) and server software modules and their interactions

allows peers to perform bandwidth tests with the super-peer. Finally, MTM informs the ISP's QoS mechanisms on the endpoints of IP multicast trees at the ISP core network (between the super-peer and the ERs serving the peers).

The MTM performs its functions using the NMSCollector, LinkTesterServer, QoSManager and Dispatcher sub-modules, as depicted in Fig. 7.3. The NMSCollector collects network monitoring data periodically sent by the NMS running at every connected peer. This information is also shared with the TB for quick P2P tree maintenance operations. The LinkTesterServer allows authorised peers to perform bandwidth tests. For the download test, it sends a predefined fixed size binary file, while for the upload test, it expects to receive the exact same file. The results based on this transfer are then used on a composite metric, to compute its evaluation. The
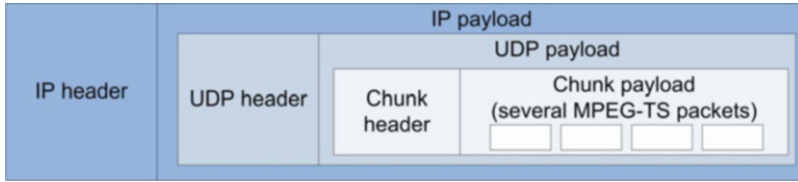
Fig. 7.4 IP packet structure

bandwidth test should be performed at the first time a peer connects (new PeerID) and upon super-peer request (for troubleshooting/maintenance reasons). Moreover, as peers connect to the TB, the QoSManager sub-module is responsible to signal the Internet Resource and Admission Control Subsystem (IRACS), an ISP QoS reservation mechanism on the new peers' addresses, to check if an IP multicast distribution tree exists from the super-peer to the ERs serving the peers (at ISP's core network level). If yes, no changes need to be done at ISP level. If not, either a new IP multicast tree is created or new branches are added to an existing tree. Finally, the dispatcher is used to interface with the TB's JSONServer [13] whenever messages need to be sent or received by the MTM sub-modules.

7.3.1.3 P2P Packetisation/Transmitter

MPEG-2 transport stream (MPEG2-TS) [7] is used for encapsulation since digital video broadcasting—terrestrial second generation (DVB-T2) [6] also uses the same encapsulation method, and the effective use of MPEG2-TS built-in program clock reference makes possible the synchronisation task of the streams received from DVB-T2 and P2P network. As shown in Fig. 7.4, MPEG2-TS packets are then encapsulated as P2P packets to be distributed over IP network. ROMEO employs a packetisation scheme, which aims at efficient and error resilient content delivery over P2P networks. As illustrated in Fig. 7.4, chunk payloads are formed to include MPEG2-TS packets that correspond to a single network abstraction layer unit (NALU)—which are self-contained and decodable units—and chunk headers will carry all the information needed for content aware P2P delivery, chunk selection and retransmission mechanisms. In this way, the P2P packetisation module aims at improving the data delivery performance since other components do not need to extract any specific information within the chunk throughout the streaming over P2P networks.

With the help of the unique chunk identifiers per stream, it is possible to detect missing chunks and request retransmission from the sender with the proposed intelligent chunk selection mechanism. Formation of the small-sized (< MTU size) chunks also improve the performance of the chunk recovery process. Because with the employed scheme, it is possible to request a small-sized single NALU with its unique identifier. Prepared P2P chunks are then delivered through the P2P network.

Table 7.1 P2P header fields

Header parameter	Purpose
Content identifier	Unique id for the content to match with DVB stream
Chunk identifier	Unique id, used for reporting defect P2P chunks
PID (stream id)	Packet ID, used for identifying different elementary streams
PCR	Program clock reference, used for synchronising P2P and DVB-T2 streams
Meta-data flag	If set, the payload contains meta-data (overrides audio flag)
View identifier	Identifies different video views
Descriptor number	Indicates the different descriptors
Layer flag	Indicates the layer type: either base or enhancement layer
Audio flag	If set, the payload contains audio data
Payload size	Size of the payload (number of bytes)
Signature	Digital signature of the payload
Priority	Indicates priority of the stream to decide if the stream is discardable

The P2P transmitter is the component, which delivers the IP streams to ROMEO users over P2P network. It is a multi-threaded UDP sender application that delivers each stream, composed of P2P chunks, in a separate thread that runs in parallel. This module is responsible for encapsulation and decapsulation of the media to be carried over the P2P network. While encapsulating, the data are separated into different streams to be distributed over different multicast trees. In two-layered coding using scalable video coding (SVC) [11], each frame has a base layer bit-stream and an enhancement layer bit-stream. It is possible to separate the original bit-stream and generate two bit-streams, one for base and one for enhancement layer. Besides, audio data and depth information are considered as different streams as well. In order to perform data distribution and synchronisation tasks, some mandatory data is needed for peers and these data, which is listed in Table 7.1 is embedded into the P2P header. As shown in Fig. 7.5, the modules that corresponds to packetisation tasks are part of the main server whereas the modules that depacketise the streams are part of the peer software.

7.3.2 ROMEO Peer

Within the ROMEO peer, the component responsible for topology related actions sends the request to join the ROMEO system with the terminal and network capability information to the ROMEO server. The P2P delivered content is received by P2P Receiver and main stereoscopic content is received by DVB Receiver. Both contents are fed into the synchronisation component for multiple view audio/video synchronisation. The synchronised content is then sent to the audio and video decoders to be decoded and then to the renderers to be rendered at the user display. As a content and network aware system, ROMEO has chunk selection and network
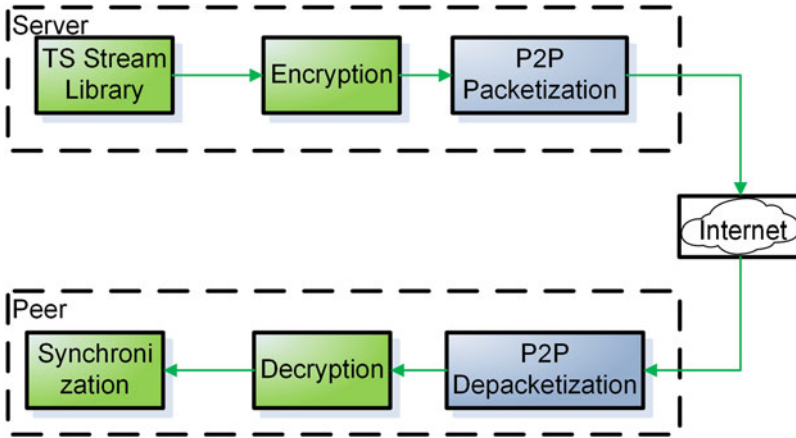
Fig. 7.5 Block diagram for packetisation

monitoring capabilities to monitor the dynamic changes in the network, and adapt to the changing conditions. Also a ROMEO peer can join the collaborating group with A/V communication overlay component and during the collaboration, can share its own content with the other collaborating users with the UGC component.

7.3.2.1   Topology Controller

The topology controller (TC) is a software module that runs at the peer and it is responsible for the P2P functionalities at the peer side. In particular, TC establishes the initial contact with the main server for user authentication and redirection to the nearest super-peer. Furthermore, TC computes the peer evaluation using peer's hardware characteristics and network statistics, as provided by the NMS. Within the responsibilities of TC is to inform the TB of its intention to consume specific 3D content and performing P2P tree operations as commanded by the TB (parent, parent/child or child). Finally, TC establishes connections with parents for content request and accepting connections from children peers for content forwarding.

In order to implement these functionalities the TC uses four sub-modules, as depicted in Fig. 7.3. The authentication sub-module interacts with the user interface and control component for user authentication interactions with the authentication, registration and security component running at the main server. The JSONServer module [13] is responsible for sending and receiving messages to/from the client application modules, whereas the PeerEvaluation module computes the peer evaluation. Lastly, the ParentList sub-module contains this peer's list of parents and backup parents (for each content type) as indicated by the TB.
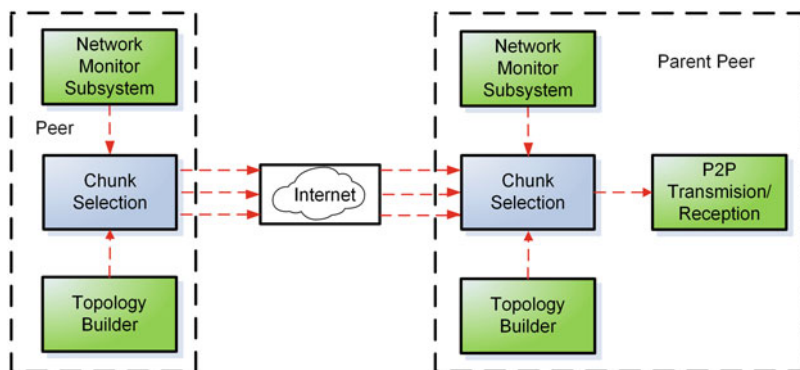
Fig. 7.6 Chunk selection interfaces

7.3.2.2 P2P Reception/ Depacketisation

In each ROMEO peer the P2P Receiver is the module responsible for receiving the
P2P streams from its parents. It is a multi-threaded UDP application, which receives
each stream by a separate thread in parallel. For each received stream (from P2P
network), the P2P Receiver sends the chunks to two different modules: the chunk
selection (to selectively forward the chunks to the children peers, as explained in
next subsection) and the P2P Depacketisation. The P2P Depacketisation extracts
the content from the P2P chunks and feeds the synchronisation component.

7.3.2.3 ROMEO Chunk Selection

The CS module is responsible for the control of the flow of chunks from a peer
to its children. The role of CS is two-fold, to adapt the streams to the network
conditions in conjunction with the upload capacity of the peer, and to implement a
fail-over mechanism that increases delivery of chunks during disconnection events.
In doing so, it uses information retrieved from the NMS and the TB. The decision is
then passed to the P2P Transmission/Reception module which in turn forwards the
selected chunks. As shown in Fig. 7.6, there are two interfaces at the input and two
interfaces at the output. These interfaces will transfer data from DVB and P2P and
deliver audio and video to decoders. Each view's base and enhancement layers will
be transmitted to the same decoder at the same time. This will enable decoders to
construct 3D image more efficiently.

The CS module consists of four sub-modules. The InterfaceToNMS requests and
retrieves the values of delay, jitter, packet loss and the upload capacity of the peer
from the NMS. On the other hand, the InterfaceToTC receives topology updates
from the topology controller. Then CS updates the list of peer's children accordingly.
Additionally, the BitrateMonitor sub-module measures the bit rate of all the streams
that are being received by the peer continuously. The values are used to determine
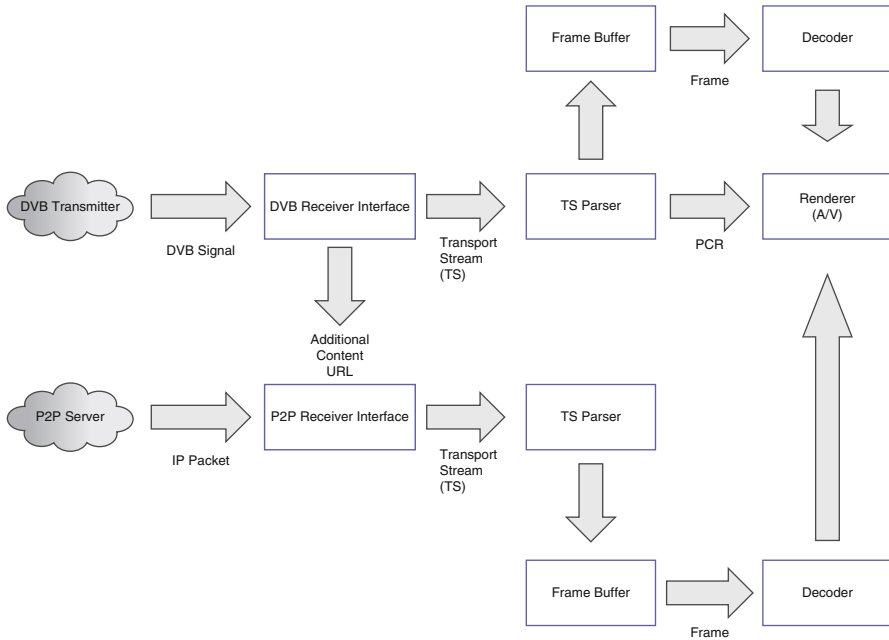
Fig. 7.7 ROMEO synchronisation module block diagram

which streams are going to be forwarded to the children. Last but not least, the ChunkScheduler is the core module of CS, as it determines which streams are going to be forwarded to the children in each tree. This decision-making process takes into account the properties of the stream (e.g. importance, bit rate), the peer capability (e.g. upload capacity), the network conditions (e.g. delay, jitter, packet loss) and the children of the peer in each tree.

7.3.2.4 ROMEO Synchronisation Module

In ROMEO, synchronisation is required at three levels. At a first level, the network synchronisation (play-out synchronisation) takes place, scheduling the play-out time of audio-visual data in order to reduce the jitter and packet loss incurred as packets traverse the network. As a second level synchronisation, the audio and video are synchronised at the terminal (temporal and spatial alignment of 3D video and spatial/object audio) in order to maintain their original temporal relationship. Finally, a third level synchronisation is focusing on the UGC of the A/V communication overlay, where the media playback of different users is synchronised, and the latency differences between users are catered for. Figure 7.7 illustrates the synchronisation module's blocks residing at the receiver.

In this section, we focus on the first and third case. In the case of network synchronisation or play-out synchronisation, the ROMEO peer receives one stream from DVB network and the other streams from the P2P delivery network. Even though both streams have identical presentation time-stamps (PTS), due to different transmission paths, they will not be received at the same time by the peer. The DVB stream packets will, under normal circumstances, be received before the delivery of corresponding stream packets through P2P. To synchronise DVB and P2P streams, a buffer is required. The basic unit extracted from the streams and kept in the buffers is a frame. In ROMEO, each frame has a unique time-stamp within the stream. In addition, in each stream there are also frames with identical time-stamps. In the peer terminal, there is one buffer for each stream. The number of buffers corresponds to the number of streams received by the peer. After all buffers receive their first frames, their time-stamps values are compared. The maximum of the compared time-stamps determines the stream with the maximum lag/delay. The buffers are synchronised with respect to the stream with the highest delay. In other words, frames with earlier time-stamps are dropped in peers that have lower delivery delay.

On the other hand, a clock reference is delivered within the transport stream of DVB transmission, which is used by the audio and video renderers of the 3D media player. Buffering of the frames for synchronisation causes a delay between time-stamps and the clock reference, and the time of some frames may elapse while they are stored in the buffer. To prevent such losses, the synchronisation block adds an offset to the time-stamps of frames at the time of forwarding to the renderers. The amount of this offset depends on the time difference between the arrival time and decoding time of a frame. After these operations, the play-out synchronisation is achieved and a basis for the temporal alignment of 3D video and spatial/object audio is formed.

In the case of user synchronisation, the collaborative experience, which was one of the main goals of ROMEO, raises an additional synchronisation challenge. Due to the nature of P2P transmission, each peer in a group may receive the same frame at different time instances, which causes a delay between peers. To provide a rich and seamless collaborative experience to a group of users, each user must watch the same frame at the same time, or with a negligible (unperceivable) delay difference. ROMEO offers a user synchronisation method similar to the deployed network synchronisation method. This method is implemented before the start of decoding and rendering. After synchronising the received streams for each individual user, the obtained maximum time-stamp values on each peer are sent to the collaboration (A/V overlay) server. The A/V overlay server computes the collaborative delay amongst all collaborating users, and then notifies/synchronises all the peers with the collaborative delay. This delay is computed according to the collaborator with the highest acceptable delay, and sent to each peer in the group.

However, network conditions of several users may differ, which may result in longer delays than practical to compensate. The value of the maximum compensable delay can vary between 5 and 10 s. More than that amount of delay may disturb other peers who have to wait the lagging ones in the group. Therefore, admission

Table 7.2 Updated structure for the NMS data report

Field	Description	Size (in Bytes)
ID	Message ID (identifies this is a report)	4
PeerID	The unique peer ID as given by the TB	32
LocalIP	The local IP address of the Peer (IPv4/v6)	16
NetMask	The local IP subnet mask	4
DL	The download capability in (Kbps)	4
UL	The upload capability (Kbps)	4
nChildren	The total number of children of this peer	4
nCPUcores	The number of CPU cores and its type	4
TotalMem	The size of RAM memory in the peer (MB)	4
FreeMem	The size of available memory (MB)	4
ConsMem	The size of consumed memory (MB)	4
OS	The operating system identification	variable
Delay	The average packet delay (μs)	4
Jitter	The delay jitter (μs)	4
PacketLoss	The packet loss (in %, content specific)	4

control is performed and the users with poor network conditions are not included or accepted in the collaboration group.

7.3.2.5 Network Monitor Subsystem

The NMS has the three important functionalities. It collects peer hardware and software characteristics, in order to determine the end device capabilities. Likewise, NMS collects network traffic statistics (i.e. packet loss, average delay, jitter and available bandwidth) for each received stream. Periodically, it reports the collected data to its parents (it chooses a different parent in each iteration using a round-robin approach) or to the MTM (in case this is a top level peer). In the end, NMS reports can also be triggered by a request from the MTM or when changes in the peer's network conditions cross a specific threshold, for an event-driven approach.

Table 7.2 shows the structure for the periodic NMS report. To save resources and simplify socket management at the receivers, reports are sent to one of the parents, who then collects all the received reports during a time-window and sends all collected reports to its own parent (one by one in a persistent TCP connection). This process goes on until the data gets to the highest peers in the P2P tree hierarchy, who then send the bundle of all collected reports to the NMSCollector sub-module of MTM.

The NMS is also responsible for detecting, reporting and possibly solving peer connectivity problems. If a major failure occurs in stream reception, the NMS of the peer first contacts the NMS of the parent responsible for streaming that specific

content, and if reachable, it is up to that parent to solve the problem. If the parent is not able to solve the problem in a pre-specified time-window, or cannot be contacted, the TC is notified in order to immediately switch to the backup parent and inform the MTM.

NMS functionalities are implemented through the four sub-modules, as depicted in Fig. 7.3. Initially, the PacketCapture sub-module passively collects network data such as connection history (start time, duration) or traffic statistics (packet loss, delay and jitter) by exploiting the libpcap library [12]. Moreover, the LinkTester sub-module provides link testing functions with the MTM to determine link characteristics such as the download and upload link capacity. The Stats sub-module is responsible for collecting hardware and software data as depicted in Table 7.2 and for computing the statistics associated with the network data collected by the PacketCapture sub-module. Finally, the so-called Reporter sub-module formats the information collected by the Stats sub-module according to a report template. It is also responsible for sending to the TC's JSONServer [13] the report to the selected peer parent. If this peer is a parent, this sub-module is also responsible for collecting NMS reports from all of its children and to send the report bundle to the selected parent in the hierarchy.

7.4 ROMEO Content Description

In the following section we will describe the 3D video generation process as well as the content packetisation process that enables synchronisation between the broadcast and the broadband access network. The scalable multi-view video encoding module that is placed in the multimedia content server is responsible for compressing the captured and pre-processed raw multi-view plus depth map videos (delivered in YUV 4:2:0 uncompressed format) and subsequently generating a number of bit-streams for each camera view and its sub-layers (e.g. video quality layers and depth). The raw video is provided by the content generation module. Well-known SVC standard [4, 11] is used to individually encode each camera view and its corresponding depth map sequence. The output of this module is sent to the media encapsulation block, where the elementary bit-streams are packed into MPEG2-TS, ready to be encrypted and encapsulated into P2P chunks.

ROMEO deploys a MMT P2P media delivery in a content aware fashion [3]. For increased delivery robustness and for exploiting the multi-path delivery platform, multiple-description extensions are taken into consideration. In order to conserve the underlying video decoding architecture, we employ multiple-description generation such that each description is self-decodable using the standard SVC decoder (each description can comprise multiple quality layers) and a full representation can be obtained easily by post-decoder processing of individual descriptions. This is illustrated in Fig. 7.8. In this hierarchical way, multiple-description feature can easily be added on/removed from the media encoding/decoding complex.
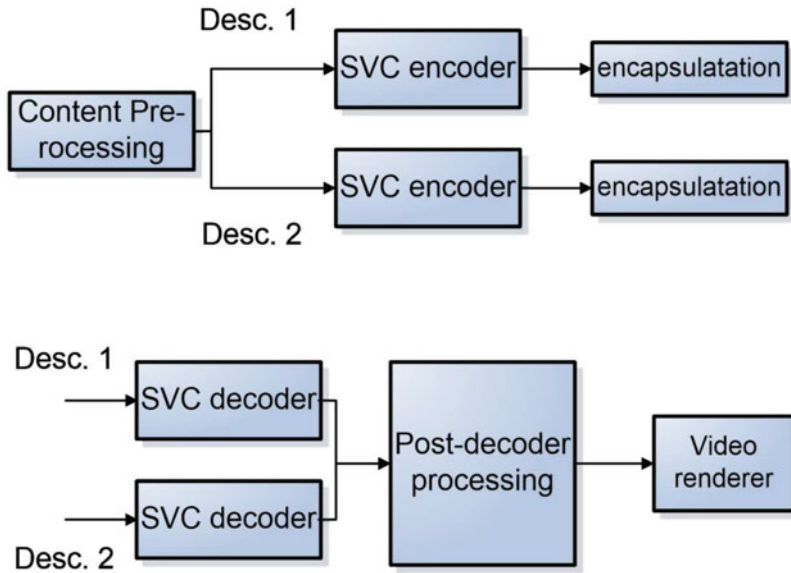
Fig. 7.8 Multiple-description generation

The scalable multi-view video coding approach is inherently scalable in view domain, i.e. any camera view can be decoded independent of any other camera view. In other words, any camera view can be discarded in the network freely, when the network conditions are deteriorating. Another envisaged addition to the scalable multi-view encoding complex is the generation of side information (view interpolation meta-data). The side information is used to assist optimal recovery of lost view frames in the video rendering. The side information is used such that the successfully received camera views' frames are blended in right proportion to obtain the lost/missing views' frames. Hence, the additional data generated in the encoder comprises coefficient sets for blending process and an index sequence to specific coefficient sets at all times. The generation process is independent of the scalable video encoding process, and so is the scalable video decoding process at the peer side. The encoded packets of meta-data blocks will be decompressed in a side process in the core video decoding process, without inter-process communication. The decoded side information (i.e. coefficients and index to them) will be delivered to the multi-view plus depth renderer to use in optimal view interpolation in case of losses.
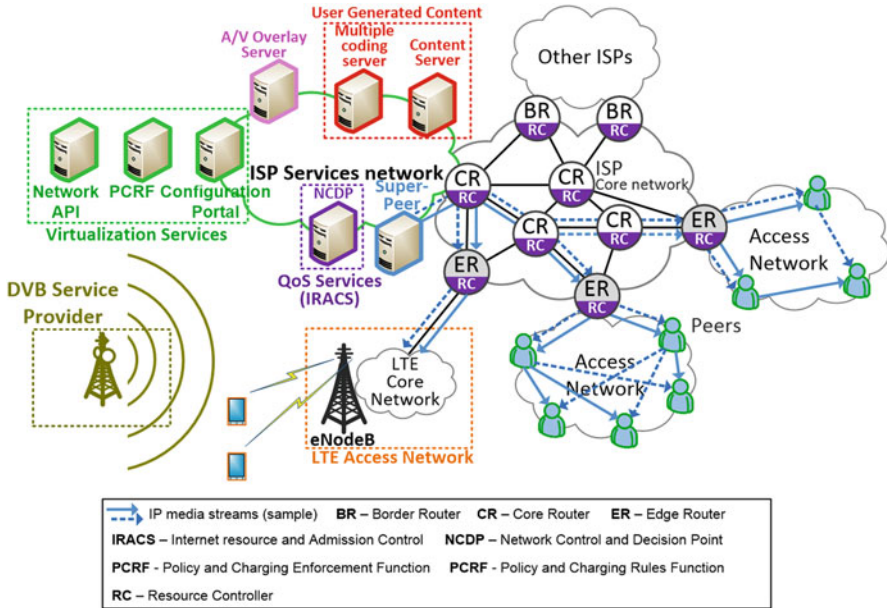
Fig. 7.9 ROMEO data delivery path overview

7.5 ROMEO Data Delivery

This section is dedicated to explain the life cycle of the content from a QoE perspective regarding both the core and the access network including the mobile users. Figure 7.9 provides an overall picture for the networking components architecture, which aims at maximising the experienced quality during the live streaming. ROMEO services are installed in the services network of each ROMEO enabled ISP. Content is then distributed from the ROMEO super-peer to the ROMEO users, passing through the ISP core network and the access network—which can be cabled or wireless. To safeguard user QoE, service requests to use the network infrastructure are subjects to admission control which is performed by the network control decision point (NCDP) and then enforced by the resource controller (RC) module in each ISP core network router. The inter-domain connections are performed based on contracted service level agreements (SLAs) to define the services to be provided and service level specification to include the traffic treatment and performance metrics between the network provider and customers. To ensure these QoS assurances are extended to the user access network, there is an intrinsic relation between the ROMEO super-peer and network API (neutralisation services) who then informs the policy and charging control function (PCRF) [17] to command the ER (also known as the broadband remote access server—BRAS) serving the user to enforce the desired QoS on the access network. The user can also personalise the QoS parameters through the use of the configuration portal. For the mobile users,
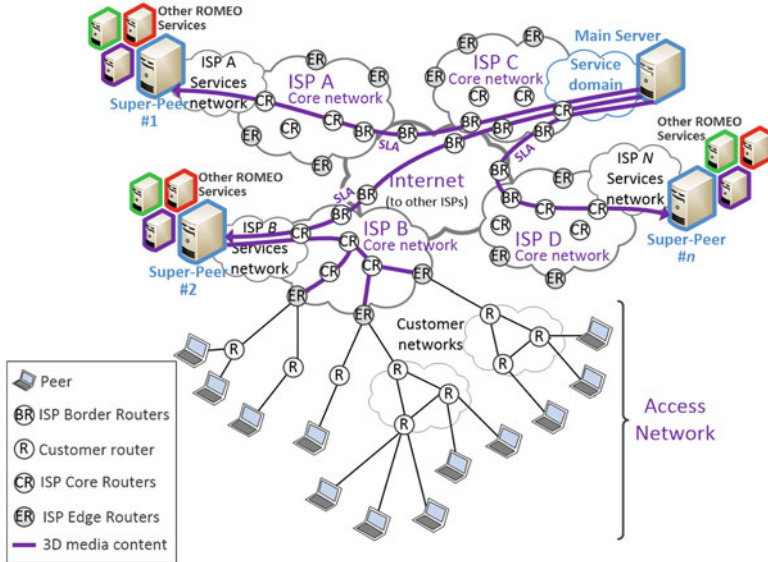
Fig. 7.10 Distribution of ROMEO content at the IP core network

the mobility aware proxy located at the long term evolution (LTE) core network
is used to perform the needed adaptations in order to guarantee the best possible
user QoE, as will be explained later. Additionally, two other services are provided
the A/V overlay server and the UGC component. The A/V overlay server is the
responsible to accept and host video conferences between ROMEO users. For each
user joining the video conference, the A/V overlay server signals the NCDP to set
up the needed QoS assurance at the ISP core network. If users are on different
ROMEO enabled ISPs, the QoS is subject to SLAs in place, and the procedures
for ISP-to-ISP signalling will be applicable. The UGC server stores and distributes
UGC to authorised users. Through the use of the multiple coding server, ROMEO
allows administrators to create multiple coding profiles that automatically encode
each uploaded content to a specific quality and bit rate.

In an initial stage (small scale deployment) the ROMEO services can be
implemented in a single ISP. This facilitates the implementation and dimensioning
of the system while providing ISP users the best QoE. Further expansion of ROMEO
services can then be achieved by establishing SLAs with additional ISPs. On a
larger scale, the ROMEO media streams will be forwarded from the ROMEO main
content server to multiple geo-located ROMEO super-peers, as depicted in Fig. 7.10.
Synchronisation between the media content database at the ROMEO main server
and the media content databases at different super-peers may occur during low traffic
conditions or upon request. SLAs between ROMEO enabled ISPs are of utmost
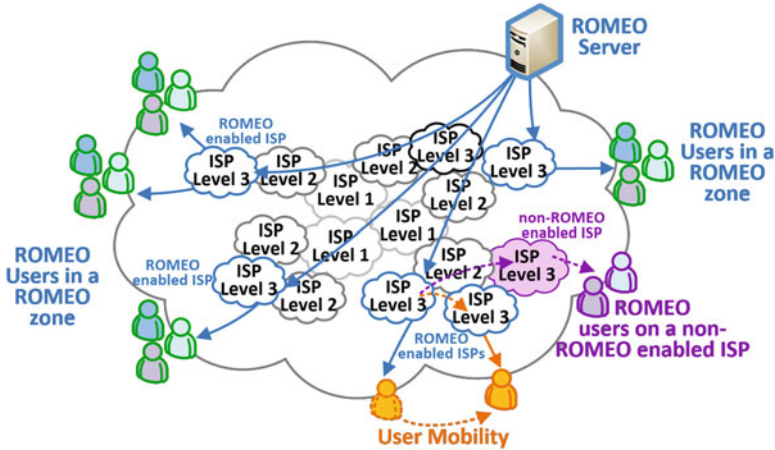importance to guarantee the best QoE to the ROMEO users.

Fig. 7.11 ROMEO user connection scenarios

7.5.1 ROMEO Content Delivery in the IP Core Network

In case a ROMEO user connects from a non-enabled ROMEO ISP, as depicted in Fig. 7.11, the super-peer (at each enabled ROMEO Level 3 ISP) from the closest location will be selected to feed that user. Since the targeted ISP is a non-enabled ROMEO ISP, the border router connecting the ROMEO enabled ISP to the next ISP towards the user will perform the multicast-to-unicast conversion. In such a case the QoS guarantees will be the ones established between the involved ISPs and admission control will be performed.

7.5.2 ROMEO Content Delivery in Access Network

In addition to the QoS enforcement procedures provided by ROMEO within the core network, QoS can also be warranted in the access network segment, in case there are agreements in place (expressed by an SLA) between ROMEO and the ISP operating the fixed access network. In these cases, the ROMEO super-peer notifies the ISP that the content is going to be delivered to a user located in its access network, so the ISP can enforce the expected QoS levels on the access network nodes. The proposed architecture is depicted in Fig. 7.12. The actual network architecture in the access network depends on the standardisation body the network operator chooses to comply with. The standards followed in the ROMEO approach are those furthered by the broadband forum (BBF). The advantage of BBF's standards is two-fold: on one side, most network operators managing fixed accesses comply with BBF recommendation; on the other side, the BBF model has been specified to ensure
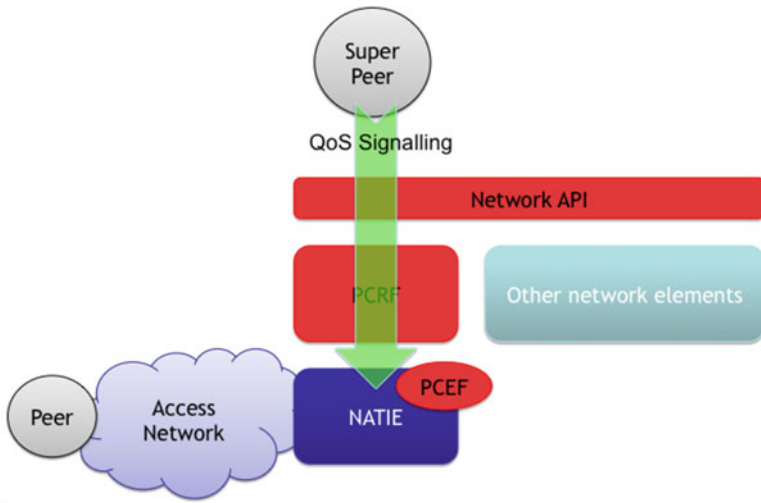
Fig. 7.12 ROMEO access network QoS enforcement

the convergence with the 3GPP standards, which are commonly followed by mobile operators.

To achieve a dynamic QoS assignation, BBF and 3GPP define a policy manager named PCRF [17]. In the ROMEO architecture, a PCRF receives QoS requests from ROMEO upper layer services through its northbound interface (Rx) and enforces them in the NATIE (IP edge router) through its southbound interface (Gx). Since the PCRF is a standardised network element, no new elements are to be introduced in the existing operator's network infrastructure, thus diminishing the impact of ROMEO deployment and therefore facilitating it. The PCRF Rx interface relies on Diameter, which is a complex protocol based on peer sessions. In order to abstract ROMEO from Diameter particularities and simplify the QoS enforcement procedures, an HTML-based overlay has been implemented over the Rx interface. Hence, the ROMEO modules can just request the enforcement of the expected QoS, based on a request/replay sessionless scheme.

7.5.3   Content Delivery for Mobile Users

The solution proposed by ROMEO combines the DVB-T2 broadcast network technology along with access network technologies using a QoE-aware P2P distribution system that operates over wired and wireless links. Inherently P2P networks provide services and functionalities with limited or no centralised infrastructure, as opposed to centralised systems like the IP multimedia subsystem [19]. Hence, ROMEO architecture overcomes the disadvantages of centralised systems, such as lack of redundancy and scalability and single point of failure. Moreover, ROMEO supports full IP mobility to mobile users through the EPC [16], which is the core network of

Fig. 7.13 ROMEO mobility modules into LTE/EPC architecture

LTE [9] technology. The protocol that provides IP connectivity using non-3GPP accesses is the proxy mobile IP (PMIP) [10] as described in [1]. The involved entities in the PMIP domain are the local mobility anchor (LMA) and the mobile access gateway (MAG). LMA is the home agent for the mobile node and the topological anchor point for the mobile node's home network IP address. MAG is a function on an access router that manages the mobility-related signalling for a mobile node that is attached to its access link. It is responsible for tracking the mobile node's movements to and from the access link and for signalling the mobile node's LMA. The ROMEO proposed solution architecture is given in Fig. 7.13.

According to 3GPP technical specification 23.402 [1], the LMA functionality is co-located with the packet data network gateway (PDN-GW). The MAG functionalities are implemented at the last hop entities based on the access technology the user equipment (UE) is attached to. In particular, for LTE access, the MAG is a functionality of the serving gateway (S-GW) for trusted WLANs. As such, MAG is implemented at the access gateway (AGW), while for the non-trusted access networks at the evolved packet data gateway (ePDG). Moreover ROMEO extends PMIP to support multi-homing in a similar manner to [2], which is an Active-Internet draft of the Network-Based Mobility Extensions Working Group of IETF. As a result, ROMEO client is not only mobile with the ability to roam in a heterogeneous environment but also capable of receiving simultaneously from multiple interfaces. In order to provide link-layer intelligence and other related network information, as well as to optimise horizontal and vertical handovers, ROMEO exploits the Media Independent Handover (MIH 802.21) protocol [20]. A set of

handover-enabling functions called MIH function (MIHF) is used for that purpose. MIHF detects changes in link-layer properties and implements a set of commands that are relevant to handover and switch between links, triggering the IP mobility process of PMIP. The other component is the media independent information service (MIIS) that provides details on the characteristics and services provided by the MIHF from both the mobile node and the network side. The information enables efficient system access and effective handover decisions. The MIIS is also located at PDN GW as this node lies at the heart of the EPC, connecting all other entities. The coupling of PMIP with MIH results in a robust mobility management that guarantees not only reliable handover management, but also efficient network detection and selection. Finally, ROMEO proposes media aware proxy (MAP) as a transparent user-space module used for low delay filtering of scalable multiple-description video streams. Its functionality extends from the network layer to the application layer providing video stream adaptation by taking into account the clients' wireless network conditions. Therefore based on a proposed machine learning model, MAP is able to either drop or forward packets that carry specific layers of a stream to the receiving users. The need for such a middlebox stems from the fact that in a real environment each client must be able to have a guaranteed QoE. The MAP function is integrated with PDN-GW and communicates internally with the MIIS database, ensuring fast transmission rate adaptation decisions, whenever required. It is noted that as PDN GW already implements packet lawful interception and routing, it can easily adopt the MAP functionalities. Subsequently, MAP inherits the EPC scalability and enhances existing processes to guarantee a minimum QoE level.

7.6 Conclusion

This chapter presented an innovative approach for immersive, social TV services as was implemented by the European-funded project ROMEO. In order to support the demanding requirements in terms of bandwidth and QoE of such services, ROMEO exploited the HbbTV concept. The broadcast DVB network was combined with the P2P broadband internet technology to deliver 3D multi-view content alongside UGC in a timely manner. For the P2P tree construction process, the MMT approach was selected for increased delivery robustness and for exploiting the multi-path delivery platform. The SVC standard was used to individually encode each camera view and its corresponding depth map sequence. For synchronising the streams of DVB and P2P networks, the MPEG-TS content packetisation scheme, and particularly the built-in program clock reference, was employed. Finally, the life cycle of the ROMEO content from a QoE perspective was explained, taking into account specific use cases to cater for content delivery to different ISPs as well as to mobile users.

References

1. Architecture enhancements for non-3GPP accesses (2015) 3GPP TS 23.402 V13.3.0
2. Bernardos C (2015) Proxy Mobile IPv6 Extensions to Support Flow Mobility. https://tools.ietf.org/html/draft-ietf-netext-pmipv6-flowmob-14
3. Birkos K, Politis I, Lykourgiotis A, et al (2012) Towards 3D video delivery over heterogeneous networks: the ROMEO approach. In: International conference on telecommunications and multimedia (TEMU), pp 60–65. IEEE, Chania
4. Birkos K, Tselios C, Dagiuklas T, Kotsopoulos S (2013) Peer selection and scheduling of h. 264 svc video over wireless networks. In: 2013 IEEE wireless communications and networking conference (WCNC), pp 1633–1638. IEEE, Shanghai
5. Cesar P, Geerts D (2011) Past, present, and future of social TV: a categorization. In: Consumer communications and networking conference (CCNC), pp 347–351. IEEE, Las Vegas
6. Digital Video Broadcasting (DVB) (2012) Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2) . TS 102 831 - V1.2.1
7. Digital Video Broadcasting (DVB) (2014) Transport of MPEG-2 TS Based DVB Services over IP Based Networks. ETSI TS 102 034 V1.5.1
8. Digital Video Broadcasting Project (2015) https://www.dvb.org/. Accessed 30 Nov 2015
9. Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (2015) Overall description; Stage 2. 3GPP TS 36.300 V13.1.0
10. Gundavelli S, et al (2008) Proxy Mobile IPv6. IETF RFC 5213
11. Information technology (H. 264/SVC) – Coding of Audio-Visual Objects – Part 10: Advanced Video Coding (2014) ISO/IEC 14496–10, 8th Edition
12. Jacobson V, McCanne S (2009) libpcap: packet capture library. Lawrence Berkeley Laboratory, Berkeley
13. libjson-rpc-cpp package (2015) https://launchpad.net/ubuntu/+source/libjson-rpc-cpp. Accessed 30 Nov 2015
14. Lykourgiotis A, Birkos K, Dagiuklas T, et al (2014) Hybrid broadcast and broadband networks convergence for immersive TV applications. IEEE Wirel Commun Mag 21(3):62–69
15. Montpetit MJ, Klym N, Mirlacher T (2011) The future of IPTV. Multimed Tools Appl 53(3):519–532
16. Network architecture (2015) 3GPP TS 23.002 V13.3.0
17. Policy and charging control architecture (PCRF) (2015) 3GPP TS 23.203 V13.5.1
18. ROMEO (2015) Remote Collaborative Real-Time Multimedia Experience over the Future Internet. http://www.ict-romeo.eu/. Accessed 30 Nov 2015
19. Service requirements for the Internet Protocol (IP) multimedia core network subsystem (IMS) (2015) Stage 1. 3GPP TS 22.228 V13.3.0
20. Standard I, Networks MA (2006) Media Independent Handover Services