Multimedia Communications: Technologies, Services, Perspectives

Part I. Technologies and Delivery Systems

Leonardo Chiariglione and Csaba A. Szabó

Abstract—This survey/position paper gives an overview of the state-of-the art multimedia communications technologies and services, analyses their evolution over the last decade, points out to their present significance and expected future role, and attempts to identify development trends. The paper consists of two parts. Part I deals with the technologies and systems for multimedia delivery. It covers the dedicated networks such as digital broadcasting systems and IPTV as well as the technologies of Internet based multimedia delivery. Networking issues including delivery over future Internet architectures and enabling technologies such as streaming and content delivery networks are dealt with in this part. Part II, to be published in the next issue of this journal, will address applications, services, and future directions.

Index Terms — Multimedia communication, IP networks, Internet, mobile communications.

I. INTRODUCTION

A decade ago, Stephen Weinstein and Alexander Gelman, recognized professionals in communications and media technologies, published a paper in the "Topics in Emerging Technologies" section of IEEE Communications Magazine, titled "Networked Multimedia: Issues and Perspectives" [1]. This excellent survey paper discussed the state-of-the-art of network infrastructures for carrying multimedia content, enumerated several existing and promising multimedia applications and proposed approaches that were supposed to lift the that time existing barriers on the way of the penetration of these applications and services. The authors said: "...resolution of several business models, public policy, and technical issues would enable a new era of networked multimedia services and, along the way, could revitalize the communications industry. It may take some time to get there, but we believe that the future broadband Internet, with both wired and wireless access, will carry the dominant mass market media services."

It is quite interesting to see where we stand now and what trends can be observed, after ten years since the paper was published, and, in particular, to address three questions: (i) how the networking infrastructures and services have developed, (ii) have the forecasted applications gained wide acceptance and implementations and (iii) are there any new trends not foreseen that time by Weinstein and Gelman. This paper attempts to answer these questions.

As for networking and services infrastructures, the authors stated: "Access networking is the bottleneck preventing us from using the optical core network to its full potential." Furthermore, "...the infrastructure for commercial quality audio/video streaming and interactive media communication is not yet in place." This paper discusses the progress that has occurred since then and tries to draw a necessarily high-level picture of the multimedia distribution and delivery networks and services of today and of the near future.

Let us refer to two other visionaries regarding the trends in multimedia networking:

Charles Judice, the father of JPEG, in his keynote speech [2], forecasted that digital storytelling could be a source of generating huge volumes of content on the Internet. Michael L. Brodie, that time Chief Scientist of Verizon, emphasized the rapidly growing user generated content [3].

The figures in recent forecasts for the expected growth of networked multimedia are really impressive. As an example, Intel said that there will be 12 billion connected devices worldwide in 2015, delivering 500 billion hours of TV and other video content. Note that the world population is expected to be around 7 billion [4].

Coming back to the forecasts by Weinstein and Gelman, they enumerated several that time existing or promising multimedia applications, including peer-to-peer exchanges of media materials, exchange of personal digital photographs and movie clips, web-based retailing of physical products, furthermore educational, government and medical services. In our paper, we address these, grouped into key application areas of networked multimedia, starting from entertainment applications through e-health, visual collaboration to smart city applications and services.

The rest of this paper is organized as follows. In Section II, we give an overview of multimedia coding techniques and standards that are of fundamental importance for digital video and sound broadcasting as well as for Internet-based multimedia delivery systems. Section III, titled "Multimedia delivery over dedicated networks" covers digital TV and

Multimedia Communications:
Technologies, Services, Perspectives

sound broadcasting (Sub-section A) and IP-based TV distribution over dedicated networks, commonly called IPTV (Sub-section B). The underlying technologies are briefly dealt with and benefits from the point of view of both service providers and customers are addressed. Sub-sections C and D discuss the issues around mobile multimedia and media delivery over heterogenous networks. Sub-section E completes Section III by an overview of IMS – IP Multimedia Subsystem – that supports service development, implementation and provisioning in IP-based multimedia networked systems.

In Section IV, we discuss some networking and access technology issues (in Sub-section A) and enabling technologies (in Sub-section B) that support the dramatic move of media distribution, delivery and consumption from dedicated systems to IP-based networks and to the public Internet. First, networking aspects will be dealt with, trying to answer the question whether we will have a totally new Future Internet network infrastructure or several incremental steps are being accomplished to satisfy the requirements posed by multimedia applications, including 3D and mobile. Challenges of providing ubiquitous Internet access are addressed next. Then an overview of some enabling technologies will be given, namely media streaming and CDNs – Content Delivery Networks.

This concludes Part I of this paper. In Part II, we shall discuss the service aspects of TV broadcasting, IPTV and Internet TV, the role and specific forms of the social element in multimedia applications, key application areas of multimedia communications, and, in the last section, which concludes this two-part paper, we shall point out to some future directions.

II. ENABLING MULTIMEDIA TECHNOLOGIES: MULTIMEDIA CODING

Studies of digitisation of multimedia information – essentially audio and video – started at the instigation of the global multi-decade plan hatched by telecommunication operators to convert their copper-based analogue networks to digital first and fiber optics-based networks later.

In the mid-1970s, European Action 211 of COST Area 2 Telecommunications became the focus of video coding activities that led to the development of a 1.5/2 Mbps videoconference codec that used DPCM and Conditional Replenishment and became the basis of the ITU-T Recommendation H.120. Later on, COST 211 became a major contributor to H.261, another video-related ITU-T recommendation for px64 kbit/s video coding (p-1, ..., 30) that used a more sophisticated and efficient linear transformation with motion compensated prediction algorithm.

ITU-T was also involved in speech coding since the early 1960s. The first standard in this area – G.711 – has two non-linear quantisation characteristics that take into account the logarithmic sensitivity of the ear to the audio intensity. Since then, ITU-T and other telecommunication-related standards organisations have continued producing speech coding standards.

With the appearance of MPEG, multimedia coding has become a high-profile area of endeavour, standardisation and exploitation. In its 25 + years of activity MPEG has produced five major generations of video coding standards and has pushed forward the frontiers of video coding performance.

At the target bitrate of 1.5 Mbps, MPEG-1 Video yields a quality comparable to the VHS cassette (comparison is made with the analogue version of video used at that time). The quality of MPEG-2 Video, measured in 1995, showed that at 6 Mbps the quality was indistinguishable from the composite (PAL or NTSC) original and at 8 Mbps the quality was indistinguishable from the component (YUV) original. The first deployments used a bitrate of 4 Mbps but the current operational bitrate is at 2 Mbps with approximately the same quality. In 1998, 4 years after approval of MPEG-2, MPEG-4 Visual yielded a reduction in bitrate of about 25% and 5 years later MPEG-4 Advanced Video Coding (AVC) yielded a further reduction of 30%. Finally, the latest MPEG video compression standard approved in 2013 yielded an astonishing 60% reduction in bitrate compared to AVC. Note that the H.264 standard specified in ITU-T is identical with MPEG-4 AVC. The two specifications are maintained jointly by MPEG and the Video Coding Experts Group (VCEG) of ITU-T. MPEG-H HEVC, too, has been developed jointly with VCEG, and it has the name H.265 within the family of ITU-T standards.

Compression is an important dimension because the spatial – but partly also temporal – resolution of video continuously increases. MPEG-1 Video was designed to work particularly for ¼ of the spatial resolution of regular television, MPEG-2 for standard definition (even though in the USA it was deployed for Digital Terrestrial Television HDTV). MPEG-4 AVC is typically used also for HDTV and the latest HEVC standard is poised to take over the so-called 4k (i. e. about 4000 pixels per line) application field.

However, the video application fields are manifold. In some cases scalability – i.e. the ability to extract meaningfully decodable sub-bitstreams from a bistream, e.g. 1 Mbps from a 2 Mbps bistream – is required. MPEG has continued working on this aspect of the video coding field for many years with increasingly better results. The MPEG-2 Video and MPEG-4 Visual scalable video compression modes save 10% of the bitrate compared to "simulcast" (i. e. transmitting two individual non-scalable bitstreams). In other terms, if the application needs two bitstreams one at 1 Mbps and another at 2 Mbps, the scalable coding mode enables the transmission of a single scalable bitstream at 2.7 Mbps. This is probably not a sufficiently high gain to justify the use of a scalable mode, but the AVC and HEVC scalable modes offer a saving of 25%. In the example above, instead of 2 bitstreams at a total bitrate of 3 Mbps the scalable bitstream has just 2.25 Mbps.

In other application domains the transmission of two signals from two slightly separated cameras are used to provide a stereo image at the receiver. This has been done in several attempts at deploying "3D TV services" by simply transmitting two separately encoded bitstreams. Starting from MPEG-2 Visual, however, MPEG has provided a "stereo mode" that saves up to about 15% for MPEG-2 and MPEG-4 Visual and up to about 25% for AVC and HEVC. The comparison for the

last case can thus be between 2 bistreams at 2 Mbps each for a total of 4 Mbps against a "stereo bitstream" at 3 Mbps.

3D Video is a world in itself whose surface MPEG has barely started scratching. Another technology to represent a 3D Video is "Texture + Depth". In this case every pixel of an image have the usual RGB or YUV values and are supplemented by a value that represents the distance of the pixel on the camera from the object that creates the pixel. This technology has only been applied to the more recent AVC and HEVC standards and offers an additional 20% saving compared to the stereo mode. Still in this space another possibility offered by MPEG standards is the ability of a user at the receiving end to define an arbitrary viewpoint of the scene and to use the available information to synthesize the missing image. Obviously this functionality entails an increase of the bitrate – minimal, at the cost of 5-10% more bitrate.

It should be noted that there is no absolute value in the numbers reported above, just a rough statistical and usually subjective assessment of the performance of the algorithms on which the standards are based.

So many things are common but also so many things are different in the field of audio, a word that is in this paper is used to mean "music".

The first MPEG attempt in the stereo audio coding field was MPEG-1 Audio (a standard approved in 1992) with a choice of 3 versions ("layers") of the standard: Layer 1, used for the now defunct Digital Compact Cassette (DCC); Layer 2, used for terrestrial, satellite and cable set top boxes; and Layer 3 soon christened as MP3, an acronym that needs no introduction. Tests carried out in 1992 showed that the 3 layers offered a "quality subjectively transparent with the original" at 384, 256 and 192 kbps, respectively. The 192 kbps of MP3 is a reference bitrate: transparency can be achieved at a higher bitrate or at a lower bitrate, depending on how "smart" the encoder is in exploiting the characteristics of the human hearing system.

The second attempt began with the extension of MPEG-1 Audio to multichannel, a kind of "bottom-up" scalability because the new multichannel audio coding had to contain the already defined MPEG-1 Audio stream. This did not provide sufficiently attractive results, so a new MPEG-2 Audio standard – Advanced Audio Coding (AAC) – was designed focused on providing broadcast quality performance for 5-channel music signals at a total bit rate of 320 kbps. This standard was further developed as MPEG-4 AAC which provides subjective transparency at 128 kbps and excellent performance down to 48 kbps. The MPEG-4 High Efficiency AAC (HE AAC) uses Spectral Band Replication (SBR) which encodes the lower frequency part of the spectrum using a waveform coder and reconstructs the high frequency part by transposing the lower frequencies. HE AAC further improves performance at lower bitrates.

Another MPEG Audio coding standards developed more recently is MPEG Surround which encodes multi-channel audio by adding a low-rate side-information channel to a compressed stereo or mono audio program. A stereo/mono player receiving an MPEG Surround bitstream still produces a useful output while new-generation players can produce the full multi-channel experience. Another MPEG Audio coding standard is Spatial Audio Object Coding (SAOC) which allows access to individual audio objects (e.g. voices, instruments, ambience etc.) in an audio mix, so that listeners can adjust the mix to suit their personal tastes. Finally Unified Speech and Audio Coding (USAC) achieves consistently state-of-the-art (as of 2011) compression performance for any arbitrary content composed of speech, music or a mix of speech and music in the sense that it provides better performance than individual codecs designed for either speech or audio and significantly improves state-of-the-art performance at bit rates ranging from 8 kbps for mono signals to 32 kbps for stereo signals, and for bitrates to 64 kbps for stereo and beyond.

The latest standard still under development is 3D Audio, an MPEG Audio coding standard suitable for all scenarios – such as in home theater, automotive, headphones connected to a tablet/smartphone – where a multi-channel audio program (e.g. 22.2) needs to be compressed and rendered to a number of loudspeakers that is not necessarily the same as used at the source.

The objective of this section was to cover video and audio standards developed within the MPEG community. Let us finally mention other audio compression formats, first of all the Dolby Digital technology, a.k.a. AC-3, which is widespreadly used in DVD and Blu-ray players and in digital broadcasting.

III. MULTIMEDIA DELIVERY OVER DEDICATED NETWORKS

Media delivery and consumption is in the process of transition from using dedicated – vertically integrated – systems, namely the radio and TV broadcast networks, through dedicated and managed IP networks, to the public Internet. This section deals with digital TV and sound broadcasting systems, and IP-based TV distribution over dedicated networks, commonly called IPTV. In this section, we will also discuss the issues around mobile multimedia and media delivery over heterogenous networks. Finally, the IMS – IP Multimedia Subsystem – that supports multimedia service development, implementation and delivery will be introduced.

A. Digital broadcasting systems

1) Digital television systems

The advantages of digital TV broadcasting, in comparison with the old analogue broadcasting, are obvious for all stakeholders. Broadcasters can broadcast more TV channels without having to buy new frequency bands. Regulators and governments can sell the bandwidth freed up by the digital switchover, the so-called digital dividend. And, last but not least, consumers get improved video quality, also in wide screen (16:9) format, mono, stereo and surround sound, several audio tracks plus new features and services (subtitling, EPG – Electronic Program Guide, interactivity...). The price the customer pays for these new features and services is not really significant as most new TV sets are already digital ones

and set-top boxes for analogue sets are inexpensive, although this may be a problem for low-income population groups. To help them, governments usually implement various support programs.

The history of digital TV broadcasting started about a decade ago, when, in 1993, the satellite system, DVB-S [5], shortly thereafter, in 1994, the cable system, DVB-C were standardized [6]. In 1996, FCC adopted the ATSC (Advanced Television System Committee) standard for digital television broadcasting in the USA. About the same time, in 1997, the ISDB (Integrated Services Digital Broadcasting) standard was adopted in Japan. In 2000 DVB-T, the terrestrial system was born [7], followed by the mobile version, DVB-H in 2004. During the years from 2005 through 2010 the 2nd generation of DVB-X standards were established: DVB-S2 (2005), DVB-T2 (2008), and DVB-C2 (2010) [8].

Digital television systems are rather interesting from the technology point of view because of the sophisticated communication and coding technologies used to take into account the specific properties of the satellite, cable or terrestrial channels. The common elements of all three systems are as follows.

• Transport stream (MPEG-2 TS). The input of the systems is the audio/video transport stream, coded and packaged according to the MPEG-2 standard, see e.g. [9].

• An energy dispersal module. This unit, also called scrambler or randomizer, is used to generate a flat spectral density and to eliminate long sequences of "0"s and "1"s, by pseudo-randomising the MPEG-2 TS packet stream.

• FEC module, also called "outer FEC", since, in DVB-T system, a second FEC module, called "inner FEC" is used. It applies a Reed-Solomon code with error correcting capability of 8 symbols in a 204-symbol MPEG2-TS packet.

• Interleaver. The purpose of this unit is to rearrange the bytes in order to randomize the channel errors and improve the error-correcting capability of the Reed-Solomon code. It uses a convolutional interleaver of depth 12, that increases the error correcting to approx. 12x8=96 symbols (bytes).

In the three digital broadcasting systems, different transmission methods and additional error correcting modules are used to take into account the different nature of the transmission channels in the three cases. In the satellite channel, only attenuation and thermal noise (AWGN) plays role, there is no multipath propagation, and the bandwidth is not as limited as in the case of the other two systems. In cable systems, the bandwidth per channel is more limited. The terrestrial transmission channel is the most challenging one, with noises and interferences and multipath propagation.

Fig. 1 shows a conceptual block diagram of the three DVB systems.

The digital TV systems in North America (ATSC) and Japan (IMDB) are built along the same principles, for a comparison see the textbook [10] and the recent survey paper [9].

In the second generation digital TV standards, further improved transmission and coding techniques have been incorporated. For example, in satellite systems the main goal

was to increase the data throughput in a given bandwidth (to increase the spectral efficiency). In the terrestrial system similar goals were set and modifications carried out. In cable systems, OFDM (Orthogonal Frequency Multiplexing) technique was incorporated instead of the single-carrier modulation schemes.
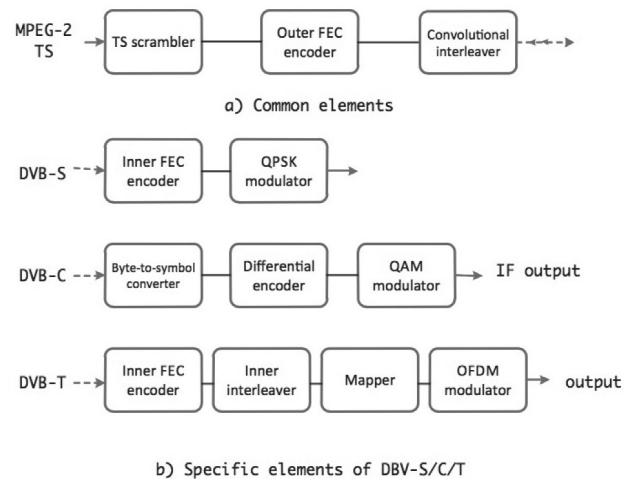

Figure 1 Conceptual block diagram of the three DVB systems

Currently the different countries around the globe have either already completed the switchover (true for most of the developed countries), or are in the process of completing it. Most of European countries completed the transition during the last years. In the United States, the switchover took place in 2009, in Australia and New Zealand in 2013. Mexico and Turkey will be among the last ones with planned switchover in 2015.

Finally let us mention the interesting member of the DVB-X family, the DVB-H (Digital Video Broadcast to Handsets), see e.g. [11]. While DVB-T was designed for use for living-room TV sets with rooftop antennas, DVB-H extends this terrestrial service to handheld devices. The technology is based on that of DVB-T which has been modified to take into account the specific properties of handheld devices, mainly the power consumption requirement but also the smaller screen and antenna, mobility and the like. The first commercial DVB-H service in Europe was introduced during the Football World Cup in 2006. After an initial fast growth of subscriber numbers (in particular in Italy where there were more than 1 million users in 2009) a decline followed and the DVB-H broadcasting was terminated in several European countries during 2000-2012.

2) Higher resolution or more dimensions in television?
HDTV, 3D and beyond

During the last few years, the HD quality, meaning 1920x1080 pixels ("full-HD"), has become ubiquitous in entertainment industry, in digital cameras, TV sets and digital television broadcasting (HDTV). It is incorporated in coding standards such as MPEG-2 and MPEG-4, monitors and TV sets are now HD-capable and most TV programs are being

broadcasted in HD format. The next step seems to be a more recent 4k (and 8k) technology also called UHDTV (adopted by CEA – Customer Electronic Association, USA - in 2012) and Super Hi-Vision (introduced by NHK in Japan). EBU, the European Broadcasting Union calls this new technology UHD-1 and UHD-2. This technology, providing a resolution of 3840x2160 pixels (thus almost four thousand pixels in horizontal direction, hence the notation 4k) and 7680x4320 pixels (8k), was integrated first in monitors and projectors starting from 2011, then in TV sets starting from 2013. An ITU-R Recommendation was approved in 2012 [12].

3D broadcasting technology has been around for several years, most TV sets in the market are 3D-capable (to be viewed with polarized glasses) and several broadcasters started 3D trials. For instance, BBC began a two-year 3D trial in 2011, and broadcasted several shows and events in 3D, including the Olympic Games. Half of the estimated 1.5 million households in the UK with a 3D-enabled television watched the opening ceremony of the 2012 Olympic games in 3D [13]. However, BBC has recently postponed the trials, and will make no further 3D programmes for 3 years. In the USA, ESPN have decided to suspend the use of 3D technology for broadcasting. The Australian Pay-TV operator Foxtel has also terminated its dedicated 3D broadcast channel [14].

Why 3D TV (based on current technologies) is not breaking through? Reasons include the viewing inconvenience due to the need of wearing glasses, and the sometimes not adequate image quality. Also 3D has added value only for a few genres, and the content offering is far from satisfactory. Why, on the other hand, it seems that ultra-high resolution 2D TV could eventually break through? It clearly offers enhanced viewing experience without a discomfort caused by a supplementary device (the 3D glasses), provides larger field of view, it is 2D but nevertheless offers a better sense of realness, and causes less fatigue for the eye and brain. The picture may change in few years from now when glassless 3D TV technology becomes available for public.

3) Digital sound broadcasting

According to ETSI, "Digital Audio Broadcasting (DAB) was conceived as a means of digitizing audio programmes in order to offer distortion-free reception and CD quality sound" [15]. Digital sound broadcasting standards include DAB, its more recent variant DAB+ and DMB. For a comprehensive treatment of digital radio broadcasting, refer to [16], and for up-to-date information, visit the website [17].

DAB specifies sound broadcasting with MPEG Audio Layer III (MP3) coding and DAB+ sound broadcasting with MPEG-4 (AAC) coding. DMB is about adding video/multimedia capabilities to audio broadcasting thus allowing DAB to become a digital mobile television platform. All three have the same physical layer just the transport etc. protocols are different and they offer different services. The main operating frequency band is VHF III (174-230 MHz / 240 MHz in some countries). In this band, a large area can be covered with an external antenna and good penetration into buildings can be achieved. L-band (1452-1479.5 MHz) is used in some

countries where Band III is not available yet or as the supplemental broadcasting band. In these frequency bands, no external antenna is needed which is an advantage particularly for mobile phones). This band is usable in urban areas where good reception can be achieved even in non-line-of-sight conditions. However, penetration into the buildings is limited and reception inside can be bad.

Advantages of digital sound broadcasting for consumers are CD quality, possibility of mobile reception, and enhanced receiver features. For operators and regulators, the advantages are spectral efficiency as compared with analogue broadcasting and lower transmitter power. Standardisation in Europe is well established.

In spite of these advantages, digital sound broadcasting is penetrating in a much slower pace than digital television has been. No country has done a complete switch-off of FM radio stations yet. Norway is the closest to that, it was announced that there will be 99.5% coverage in 2014, and that Norway was planning a switch-off of FM radio in 2017. There are signs of penetration in other countries as well. In the UK, 46% of households have DAB and the national coverage is 94%. 44% of new cars are equipped with digital receivers. Germany plans full national coverage by 2014 [18], [19], [20].

Let us finally mention DRM – Digital Radio Mondiale, which has been designed specifically as a high quality digital replacement for current analogue radio broadcasting in the AM and FM/VHF bands [21]. There is no significant penetration, many countries in Europe started then stopped their trials and did not launch commercial DRM broadcasting.

In spite of the standardization efforts in the aforementioned organizations and introduction plans in various countries, the future of digital sound broadcasting is at least unclear. Users can listen to a large amuont of radio stations on the Internet (we shall come back to this issue later), and as music is the primary genre in radio broadcasting, downloading MP3 songs from the Internet and enjoying them on mobile devices is just enough for most listeners.

B. Multimedia distribution over dedicated IP networks: IPTV

According to ITU-T Focus Group: „IPTV is defined as the service delivery of video/audio, text, graphics and interactivity over IP based networks managed to provide the required level of QoS/QoE, security and reliability".

IPTV is an opportunity for "classical" telecom operators to enter into the broadcasting business. Since they already play the role of an ISP by providing Internet access, typically over their xDSL networks, by adding TV they become a "triple play" provider of TV+Internet+Telephone services. IPTV offers services such as interactivity, time shifting (playback after the initial broadcasting of the content), VoD – Video-on-Demand - content consumption, program recording, and EPG – Eletronic Program Guide. The latter is an electronic program that allows intelligent selection and sorting of programmes as well as obtaining all kind of information about specific programs.
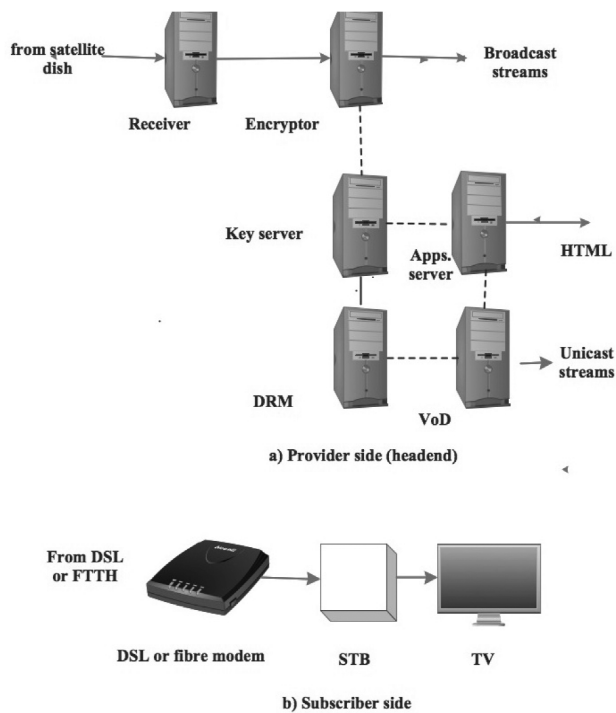
Figure 2 Functional diagram of an IPTV system

The technical aspects of the IPTV service are illustrated in Figs 2-3. On the functional diagram of Fig. 2, the headend (a term borrowed from cable TV systems) is where the content is collected and processed. Content can be live TV programs from a satellite or terrestrial distribution network, or can be a stored one from local media servers. Live or stored video then coded/transcoded, encrypted and transmitted to clients. Electronic Program Guide support is also part of the headend. The client side functional unit is the set-top-box (STB) which performs the media decoding, decryption, EPG client functions.
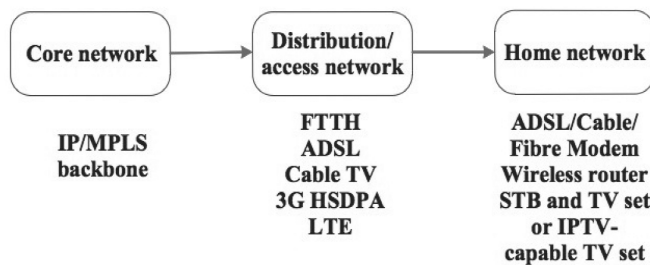

Figure 3 The IPTV delivery network

Fig. 3 shows the high level networking infrastructure used for IPTV delivery. The core part is the IP/MPLS backbone of the telecom service provider. The access network or first/last mile network is mostly xDSL or FTTH – Fiber-to-the-home, but it can be a Cable TV distribution network or a broadband wireless network (3G – HSDPA, LTE or WiMAX) as well. The home network usually consists of an ADSL modem, a wireless LAN router, IPTV-capable TV set with a set-top-box or with built-in IPTV capability and additional client devices.

The IPTV protocol architecture is shown in Fig. 4. The media stream coming from the application layer is coded into PES – Program Elementary Stream according to the MPEG standard, then it is packaged into MPEG Transport Stream packets (the same are used in digital television standards). Media transport is supported by RTP – Real Time Transmission Protocol that provides sequence numbering and time stamping services. RTP packets then carried in the payload of UDP – Universal Datagram Protocol packets. The protocol overhead added to the 188 bytes long MPEG TS packets is total 40 bytes plus the MAC/PHY overhead. For an extensive treatment of IPTV technology, see the textbook [22] and the paper [23].

To meet Quality-of-Service requirements and Quality-of-Experience expectations of the customers, a series of technical challenges have to be addressed. An IPTV system itself is a pretty complex one, so even if the input stream is ok, which is not always the case, sources of quality deterioration can be the failures in the core network (rarely), in the distribution and access networks (more frequently) and of course within the subscriber's home network. From the customer point of view, all this should be the service providers responsibility, however, the latter is not in the position of managing all the aforementioned components from a central place. (E.g. media streams are often sourced from third parties.)
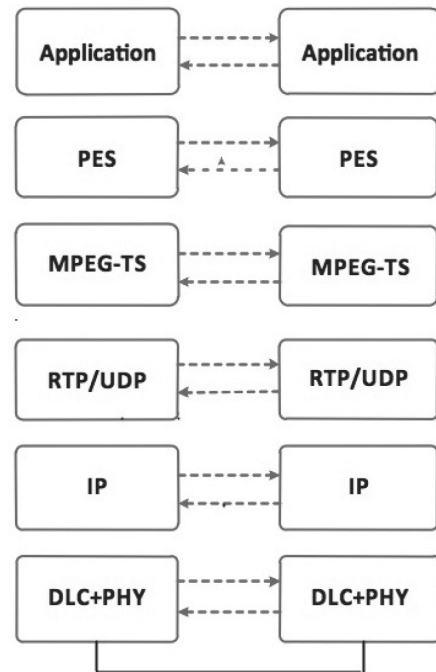

Figure 4 IPTV protocol architecture

Coming back to the customer side: What can IPTV offer (compared with digital TV broadcasting)?
- The same high quality picture and sound as in digital broadcasting.
- Time Shift - allowing playback of content after its initial transmission.
- EPG or Electronic Program Guide.

- Personalized interactive media consumption in many ways (recording programs, video on demand, alert messages for favorites programs etc.).

- Communication (video conferencing) and online training service.

In classical television broadcasting, there have been improvements in picture quality (HDTV), in the channel offering (multiplexes in digital broadcasting), however it remained basically a one-way distribution vehicle from the service provider to the end-user with a very limited inter-activity. At the same time, a large share of TV users, in particular the younger generations, having already accustomed to the freedom when consuming media, including TV prog-rams, on the Internet, are no longer satisfied with what the traditional TV broadcasting systems offer. For them, IPTV might be attractive.

C. Mobile multimedia

Providing multimedia services, such as the distribution of TV programs, for mobile users satisfies the growing demand for accessing these services anytime, anywhere and on any device. Mobile multimedia refers to transmission and delivery of multimedia information to mobile customers who access the Internet via cellular mobile services. Because of the spe-cific properties of the wireless mobile channels - high error rate and packet loss rate, lower bandwidth and bandwidth depending on location and the heterogeneity of access networks and user devices -, serious technical issues have to be solved, including coding and presentation of multimedia content for mobile devices, end-to-end error control, multicast transmission, mobility management and other network-related issues. For example, the H.264 multimedia coding standard provides specific coding technique called Flexible Macroblock Ordering to cope with error propagation and error accu-mulation. Scalable Video Coding (SVC), an extension of H.264/MPEG-4 AVC video compression standard, provides adaptation of the coding rate to the estimated bandwidth of the wireless channel [24].

TV broadcasting to mobile devices requires coding formats suitable for mobile screens (QCIP, CIF, QVGA resolution), although there will be more and more devices with enhanced resolution (full HD), thus adaptation to the aforementioned formats might become unnecessary in the future. In addition, mobility management is needed even at high speeds (usage in cars on motorways), and multicast transmission is required. As for the latter, the 3GPP Release 6 standard includes a service called MBMS – Multimedia Broadcast Multicast Service, which is a general point-to-multipont service for IP packets offering data rates up to 256 kbps. Subsequent releases extend it for 3G/HSDPA and 4G/LTE mobile cellular services.

Lastly, ensuring mobility needs sophisticated methods and protocols starting from mobile IP at the network layer, through transport layer mobility protocols to solving mobility in application layer using SIP, the Session Initiation Protocol. A specific case is when the user moves accross wireless and mobile networks that are based on different technologies. The handover between cells in this case is called „vertical

handover", to distinguish the task from the usual handover when the user moves across cells of the same mobile cellular network. Let us briefly explain it by the example of a past project that was carried out by one of the authors and his team.

D. Media delivery over heterogenous networks

In a multi-platform access network environment, the user has several physical connections to access the Internet, hence the same resource could be accessed via different wireless networks, even simultaneously. This opportunity could be utilized to achieve higher quality service, i.e. faster download or higher quality media streaming solution by using all access networks simultaneously or selecting the best access net-work(s) dynamically. On the other hand, the available wireless access networks have quite different characteristics and properties such as average and peak bandwidth, availability, delay and jitter, packet loss rate and bit-error rate, optimal packet size, and pricing. Furthermore, these properties usually depend on the actual state of the network and on the user's location. In the media streaming architecture outlined in [25], a best-effort single-connection scheme is used i.e. the media streaming system uses the best connection (active connection) to transmit the media stream and avoid the other (idle) connections. In the single connection scheme, the moment of the handover (namely the change of the active connection) must be invisible to the user and he or she becomes aware of the handover only by observing a degradation or improvement of the media quality, depending on the characteristics of the earlier network connection and the new one. The decision on the switching of streams is based on the client's measure-ments. Based on the measured parameters (current packet loss rate and the access network type), the optimal bandwidth is estimated, the ranking of the access networks are made, and the best bandwidth/quality version of the content is deter-mined and the switching is carried out in case of need. To accomplish this, the media server should provide the same media content in different resolutions continuously to allow the system to choose the appropriate resolution according to the quality of the active connection and the properties of the client device.

The media streaming architecture in [25] has the following key features:

- Vertical handover among different access networks, including 2G and 2.5G technologies (GSM, GPRS, EDGE), 3G cellular (UMTS), WLAN (Wi-Fi), WMAN (WiMAX) and even some wireline access such as xDSL.

- Horizontal handover, i.e. handover between the same kinds of wireless networks of different service providers.

- Content- and environment-adaptive charging, accounting, billing and payment schemes.

- Digital rights management schemes.

The generic system architecture is shown in Fig. 5. In the testbed, the UMTS and GPRS/EDGE access networks be-longed to the same service provider, whereas the WLAN, WMAN and xDSL access networks were provided by a different operator. The xDSL wireline network was accessed via a Wi-Fi wireless access router.

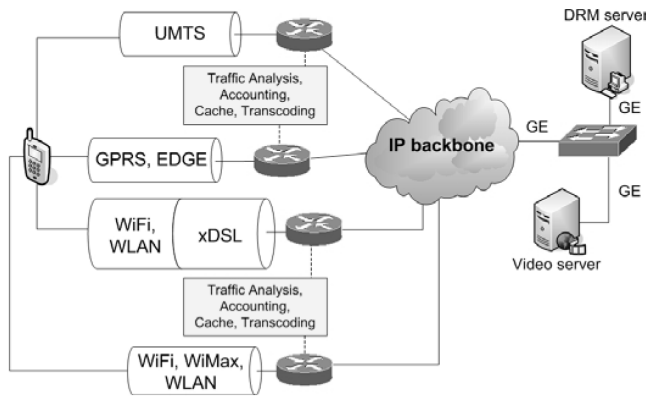Multimedia Communications:
Technologies, Services, Perspectives

Figure 5 The generic scheme of the media streaming testbed [25]

E. Supporting service development, implementation and delivery: IMS

Starting from the introduction of VoIP – voice over the Internet Protocol – in mid-late 90s, telecom service providers – both incumbents and new ones – have been gradually moving from circuit-switched to packet switched voice services. While in the first VoIP systems the signalling/session control protocol was ITU's H.323 [26], SIP or Session Initiation Protocol, developed within IETF [27], emerged almost in parallel. While their functionalities are similar, SIP is a more flexible and better scalable protocol that can be easily integrated into web-based applications. At this point SIP seems to be the future. In the process of the development of newer versions of mobile communications systems, on the one hand, and moving towards a new concept of NGN or Next Generation Networks, on the other hand, it turned out that however important the session control can be, it is just one of the functionalities needed for supporting the development, implementation and provisioning of multimedia services over packet switched networks. Therefore, in the standardization body of the mobile world, 3GPP - Third Generation Partnership Project, a more complex new element of the network architecture, incorporating also SIP, called IMS - IP Multimedia Subsystem – was specified in their Release 6 [28].

The need for such a functionality has also emerged in the telecommunication world within the context of NGN standardization in ETSI TISPAN. (ETSI TISPAN – Telecommunications and Internet converged Services and Protocols for Advanced Networking - has been the key standardization body in creating the NGN - Next Generation Networks specifications.) NGN represents a paradigm shift from the classic telecom service model of independent, vertically integrated networks to a new architecture that comprises a variety of access networks and has a new horizontal layer or platform that supports service provisioning with important functionalities such as call control, quality of service provisioning, media gateways, authentication, authorization, and accounting (AAA) and the like. This new architecture allows telecom companies to successfully compete with Internet-based services, and in general, supports the convergence of the Internet, telecommunication and media industries. On the other hand, the horizontal separation of functionalities in telecom networks allow third parties to come and put their services on top of the network infrastructure of network service providers, see e.g. [29]. IMS was standardized both in the mobile world and telecom world, another big step in the process of their convergence.

IMS entities and key functionalities include:
- Session management and routing, based on SIP – the Session Initiation Protocol.
- Databases (like HLR – Home Location Register - in cellular mobile systems),
- Interworking elements (e.g. media gateways).
- Application servers and services, e.g. AAA – Authentication, Authorization and Accounting - based on Diameter protocol.

The IMS architecture is illustrated in Fig. 6. As it is shown on the top of the block diagram, the IMS system supports interfacing with legacy mobile call control systems as well as interworking with non-IP networks such as with the circuit-switched PSTN. To go a bit more into the IMS system, let us briefly mention the functionalities of its building blocks. The three CSCF – Call Session Control Function – nodes implement the SIP session control protocol. P-CSCF or Proxy-CSCF is the closest to the user agent and functions as SIP proxy server. I-CSCF or Interrogating CSCF determines the route of a call to the called UA, while S-CSCF or Serving CSCF serves the UA. These units communicate with HSS or Home Subscriber Server (identical to Location Server in SIP) and with SLF or Subscriber Location Function. BGCF or Breakout Gateway Control Function handles calls originated by the IMS and destined to PSTN. MGCF or Media Gateway Control Function takes care of the interworking process, while
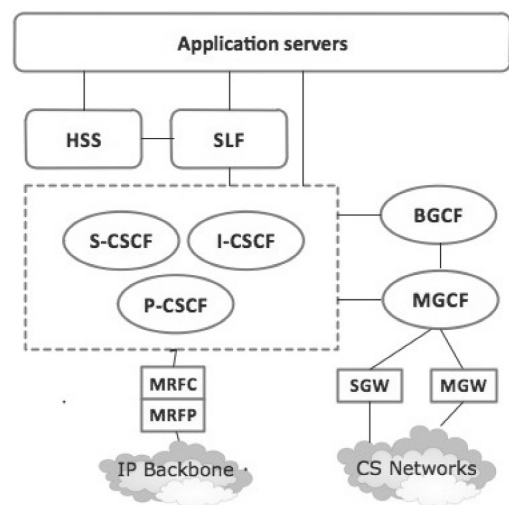
HSS - Home Subscriber Server, SLF - Subscriber Location Function
S/I/P-CSCF - Serving/Interrogating/Proxy Call Session Control Functions
BGCF - Breakout Gateway Control Function
MGCF - Media Gateway Control Function
SGW - Signalling Gateway, MGW - Media Gateway
MRFC/MRFP - Multimedia Resource Function Controller/Processor

Figure 6 High-level architecture of IMS

MDW or Media Gateway carries out the necessary media conversions. Finally, the units MRF or Multimedia Resource Function and MRFC, MRF Controller, and MRFP, MRF Processor handle multiparty calls such as for multiparty multimedia conferencing.

Additional services, provided by IMS and could not be illustrated in Fig. 6, include the already mentioned AAA or Authentication, Authorization and Accounting, as an example. For an extensive treatment of IMS architecture and services, see [30], [31], [32].

Based on the services provided, IMS can also be considered as a multimedia SDP or Service Delivery Platform, offering the necessary support for multimedia services to be provided by telecom operators or third parties. SDP however is a term more often used in a broader sense, and denotes facilitation of service composition and integration, so that IMS can be considered as an additional layer on top of SDP.

In spite of the relative maturity of IMS and the potential advantages it can offer, its penetration has been so far slower than one might have expected. At the beginning the majority of significant telecom operators have purchased IMS systems from leading vendors such as Ericsson, Nokia or Huawei, primarily for testing purposes. There are several reasons why the commercial deployment has been not so fast. One of them might be that mobile operators are going pretty well without it and are reluctant to make a significant investment. It looks like that the advantage of IMS we mentioned above, i.e. that it is an unified platform for developing, deploying and providing multimedia services over IP networks, and that the operator can do it more efficiently using the "toolset" IMS provides, have not been transformed into specific business benefits so far. Also operators seem to be not too enthusiastic attracting third parties to bring their services and putting them on the operators network. The driving force will apparently be some new services that only IMS offers and that can immediately generate revenues. These services include push-to-talk, presence, multimedia sharing, emergency calls etc. Without them, the future of IMS will be unclear.

IV. MOVING FROM DEDICATED NETWORKS TO THE INTERNET

A move from dedicated and managed IP-based networks to the public Internet seems to be simple, since the communication protocols of the TCP/IP stack are common, but it is a huge step at least in two aspects. One, delivering broadcasting content over the public Internet represents challenges in terms of ensuring access bandwidth, reliability, quality of service and like. Two, specific distribution and consumption models, including business models, arise. In this section, the networking aspects will be addressed.

A. Networking and access issues

1) On the architecture of the Future Internet

Will there be a radically new architecture? "Clean slate" or "evolutionary" design shall be followed? What shall be the design requirements and principles of the Future Internet, in particular of the Future Media Internet? How will this new architecture relate to the already existing and standardized in 3GPP and ETSI TISPAN NGN architecture?

These and similar questions have been posed and answers sought by several projects and working groups, labeled by the term "Future Internet" or FI, around the world, supported in particular by NSF in the USA and EU research framework programs in Europe. NSF launched its FIA – Future Internet Architecture program in 2010 and funded four projects [33], [34], then launched the second round of in 2013. In EU, the "Future Media Internet - Think Tank (FMIA-TT) supported by the nextMEDIA project aimed at working out a reference architecture model of the "Future Media Internet", "covering the delivery, in the network adaptation/enrichment and consumption of media over the Future Internet ecosystem" [35]. According to the leading professionals teamed together in this project, the existing Internet architecture should be replaced by a new three-layer one. In this hierarchical FMI architecture, the lowest layer is the Service/Network Provider Infrastructure Overlay. This is where the users who are both Content Producers and Consumers (therefore called "Prosumers") are located. They are connected through the infrastructure of the ISPs and network service providers. The nodes of this infrastructure have limited functionality and intelligence. The second layer is the Distributed Content/Services Aware Overlay, contains content-aware network nodes which are more intelligent as compared with the infrastructure nodes and are capable of identifying and qualifying content and services and reporting to the third layer of the architecture (Content/Services Information Overlay). It consists of intelligent nodes or servers that have a distributed knowledge of the locations and cashing of the content and of the conditions in the network. Based on this information, decisions can be made e.g. on the optimal delivery of content to the subscribers. We should note, however, that while introducing content aware network nodes and layers is certainly a good approach to the building of the "Future Media Internet", it somewhat contradicts to the network neutrality principle currently required from the ISPs and network service providers.

By now it has become clear that there will not be a radically new FI architecture. However, new approaches, design and improvements are needed in areas including:

- New networking protocols, in particular cross-layer solutions.
- Efficient methods to handle multimedia traffic which is already dominant and continues to grow.
- Ensuring throughput, Quality of Service, Quality of Experience.
- Providing access from anywhere, from any device, with the desired quality to users who are prosumers, that is consumers of media as well as creators of content.
- Ensuring seamless mobility, between arbitrary network technologies and systems.
- Adaptivity to the capabilities of user devices and network, ensuring the desired quality.
- Meeting the requirements of the Internet of Things towards the network, e.g. wireless (multimedia) sensor networks, with self-organizing capabilities.

Multimedia Communications:
Technologies, Services, Perspectives

Techniques and solutions outlined in the sub-sections to follow address some of these challenges.

2) Challenges of providing ubiquitous Internet access

When discussing multimedia services, it is often assumed that access to the public Internet is available everywhere with the desired speed („bandwidth") and quality of service. In a NSF study, we can read: "Historical infrastructures – the automobile/gasoline/roadway system, electrical grids, railways, telephony, and most recently the Internet – become ubiquitous, accessible, reliable and transparent as they mature." [36] While it is true for some historical infrastructures, ubiquitous access and reliability certainly cannot be taken for granted in the case of telecommunication networks and the Internet. And we are not talking about developing countries only and their under-developed regions, where providing just basic telecom access presents a huge problem. Ensuring broadband access to everyone and everywhere is also a challenge in developed countries because relying merely on market economy cannot solve this problem. Telecom and Internet companies operate according to their specific business models, which do not allow expanding their infrastructures to sparsely populated and/or geographically challenged areas, therefore, these areas remain underserved. This is one of the manifestations of the so-called "digital divide", a gap between those having proper Internet access and those who do not. Therefore, providing broadband access to citizens, communities, public institutions and developing businesses has become a strategic objective for state and local governments worldwide. A large number of initiatives, under the collecting name "community networks" or "municipal wireless" have been launched in North America as well as in Europe (see [37], [38], [39], [40]). By creating telecom infrastructure in underserved regions, local governments can prevent remote communities from digital divide, and are able to create a healthy climate for economic development, can help startups grow, and bring new businesses into the region. Often cited examples include the municipal network pioneer city of Corpus Christi, TX in the USA or the more recently deployed municipal network in Barcelona, Spain and the large scale network of the Province of Trento in Italy [41]. Solving the digital divide issue by building and operating city-wide or regional network infrastructures, local administrations create possibilities for advanced multimedia services such as telemedicine, e-learning applications, portals for tourists, regional TV channels, surveillance systems and the like, thus bringing additional benefits to the citizens and businesses as well as making these networks sustainable.

B. Enabling technologies for the implementation and provision of multimedia services

1) Streaming techniques

Audio/video streaming or multimedia streaming has been around for quite a long time and is today perhaps the most important technology component in networked multimedia applications and services. Its history started in 1995 when Real Networks launched RealAudio then RealVideo in 1997.

In 1998 Apple announced QuickTime Streaming. A decade later, in 2007, Hulu launched its streaming service, offering ad-supported streaming video of TV shows and movies from many networks and studios. Today there are several thousands of TV stations available online on the Internet. In 2013 YouTube reached one billion monthly users with 4 billion views per day. Today, the most commonly used streaming technologies are Microsoft's Windows Media [42], RealNetwork's RealPlayer [43], and Apple's QuickTime [44].

Multimedia streaming is a technology that enables clients to download audio/video files from servers and to start viewing them immediately without waiting for complete download, and continue viewing without interruption. In addition, the user is provided with some DVD-like functions such as pause, resume, fast forward, rewind, etc. Key elements of the streaming system are playout buffer on the client side, protocols ensuring or supporting quality of service and specific protocols for streaming applications.

There are three classes of streaming applications: (i) stored media streaming, (ii) uni-directional real-time (live) streaming such as TV stations, and (iii) bi-directional real-time (live) streaming e.g. video conferencing. The technology and protocols used are essentially the same for all three classes. A playout buffer is used on the client side to compensate the fluctuations in the transmission delay and handle lost or out-of-order packets. The three classes significantly differ in the required quality of service parameters in particular delay, jitter and packet loss. For example, unidirectional live streaming requires less than 10 ms initial delay, less than 2 ms delay variance and < 2% packet loss. For interactive streaming applications, the end-to-end delay shall be around 150 ms, the delay variation < 1 ms and packet loss < 1%. In addition to the QoS parameters that are measurable in an objective way, QoE or Quality of Experience plays an important role, too. QoE is a subjective measure of the user experience which is influenced by many factors.
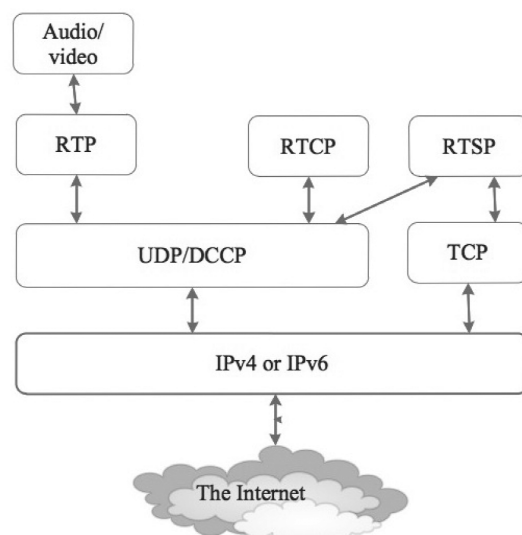

Figure 7 Streaming protocol architecture

Figure 7 shows the protocol architecture and the protocols commonly used in streaming applications. Going from bottom upwards along the architecture, the network protocol is obviously IP. At the transport layer, UDP is generally used for media transmission. Its limited functionality (only multi-plexing/de-multiplexing, no error control via retransmissions, no congestion control) makes it robust and suitable for media transmission since it introduces almost no delay. TCP is used for control purposes. The transport protocols are not media-specific, therefore we need additional ones that support media transmission, such as RTP or Real-Time Transmission Protocol which uses sequence numbers and timestamps to help reconstruct the media stream on the receiver side. Its companion protocol, RTCP or Real-Time Transmission Control Protocol provides measurement information on the quality of transmission to the sender and receiver. Finally SIP or Session Initiation Protocol is used for session establishment and control, and RTSP - Real-Time Streaming Protocol is an application level protocol to provide the user with some DVD-like control functions during the streaming session.

More recent streaming technology is HTTP streaming. As the name suggests, it uses the HTTP protocol, and media is transmitted, using HTTP, in the form of successive short pieces (short files called chunks) and the client reconstructs the media stream from these independent chunks. HTTP streaming was first introduced by Apple for its QuickTime software. It is called HTTP Live Streaming or HLS. Its relatives are Microsoft's IIS Smooth Streaming, Adobe's Flash Dynamic Streaming and DASH, Dynamic Adaptive Streaming over HTTP.

HTTP Live Streaming is an adaptive protocol. At the sender side, multiple files are created for distribution to the player, which can switch between streams in an adaptive way to optimize the playback experience. The media stream at the source is encoded into multiple files at different data rates and is divided into short chunks of 5-10 seconds long. These are loaded onto an HTTP server along with a text-based manifest file that directs the player to additional manifest files for each of the encoded streams.

HTTP-based streaming has several advantages; no streaming server is required and the download of the media chunks should use HTTP caching servers located at different places of the networks of service providers, cellular providers, resulting in improved video quality for clients served from these caches. An important advantage is that content via the HTTP protocol can pass through most firewalls and proxy servers which is not the case with RTP over UDP.

HLS is currently being standardized in IETF and at the time of writing (beginning of 2014) its specification is an Internet Draft [45].

2) Content Delivery Networks

As the sharing and consumption of multimedia content on the Internet has been growing rapidly, it has become obvious that many web servers hosting content and applications are unable to handle this demand, not to speak about bandwidth and quality of service requirements. The concept of CDN or Content Delivery Network has emerged to cope with the exponentially growing demand for exchange of multimedia information on the public Internet, to ensure scalability of multimedia networks and to enhance quality of experience of users.

To put it simply, CDN is a set of web servers, collaborating with each other, and hosting multiple copies of the same content to accomplish more efficient delivery of the desired content to the end users. The CDN concept is not entirely new as caching has been used to deliver general web content for many years, however, moving to delivery of on-demand or pre-recorded video and even of live video required new architectures and protocols. According to [46], CDNs have evolved from their first generation that delivered general static and dynamic content through 2nd generation that supported video-on-demand, streaming media and also mobile media applications during late 2000s to their 3rd generation, the community-based CDNs at the beginning of 2010s.

Main functional elements of a CDN architecture are: (i) origin servers where the content is put by the content owner and stored, (ii) edge servers or surrogate servers (caches) servers where copies of the multimedia content are distributed to and stored, (iii) distribution network which delivers content requests to the optimal location, (iv) redirector or request routing system that identifies the optimal (closest, not only in geographical sense) edge server for each user, and (v) some accounting mechanism for the origin server. Fig. 8 serves as an illustration. User request for the desired content is re-directed to the optimally closest edge server (1). The latter then searches for the content on its storage facility and if not available, checks other edge servers in its proximity (2, 3). If content is not found in the proximity of the end user, the request is sent to the origin server (4) which then delivers the content to the edge server and the latter delivers it to the end user (5).
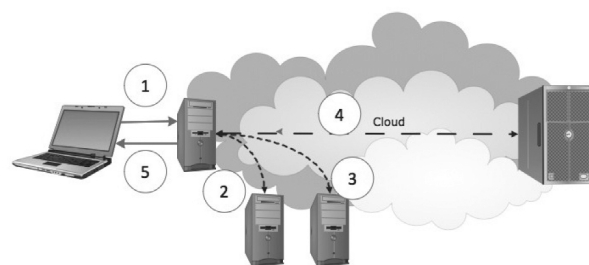
1: User agent is directed to the closest edge cache server
2-3: Edge cache checks other edge caches since the content is not cashed there or it is old
4: Content is not cashed in edge servers or it is old so edge cache request it from the origin server
5:Content is delivered to user

Figure 8 The simplified scheme of obtaining content via a CDN

The largest CDN service providers include Akamai, the market leader [47] and Limelight Networks [48]. Akamai's market share is estimated to be over 80%, it operates 12000+ servers in 60+ countries.

A recent direction of CDN development is to support collaborative media streaming services using the Hierarchical Cooperative Control Protocol (HCOOP) [46].

The role of CDNs in multimedia communications is already significant and will continue to grow. According to [49], Content Delivery Networks (CDNs) will carry over half of Internet traffic in 2017, up from 34 percent in 2012, and the share of video traffic delivered over CDNs will be over two-thirds of total video traffic by 2017.

REFERENCES

[1]   S. Weinstein and A. Gelman, "Networked Multimedia: Issues and Perspectives", IEEE Communications Magazine, Vol. ..., No. 6, pp. 138-143, June 2003.

[2]   Charles N. Judice, „Digital Storytelling: The Next Killer App ", Keynote speech at „Multimedia Services Access Networks" conference, Orlando, Florida, USA, June 13-15, 2005, http://msan.org.

[3]   Michael L. Brodie, "The end of Computing Era", 2nd IEEE Conference on Digital Ecosystems Technologies (DEST), Phitsanulok, Thailand, February 2008.

[4]   Sheau Ng, "A Brief History of Entertainment Technologies", Vol. 100, May 13, 2012.

[5]   Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services. ETSI European Standard EN 300 421.

[6]   Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for cable systems. ETSI EN 300 429 European Standard.

[7]   Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television. ETSI EN 300 744 European Standard.

[8]   ETSI Standards for 2nd generation DVB systems.

[9]   El-Hajjar – L. Hanzo, "A Survey of Digital Television Broadcast Transmission Techniques", IEEE Comm. Survey&Tutorials, 2013.

[10] J.-N. Hwang: "Multimedia Networking – from Theory to Practice", Cambridge University Press, 2009.

[11] Amitabh Kumar, "Mobile TV: DVB-H, DMB, 3G Systems and Rich Media Applications", Elsevier, 2007.

[12] Parameter values for ultra-high definition television systems for production and international programme exchange. Recommendation ITU-R BT.2020 (08/2012).

[13] http://www.bbc.co.uk/news/entertainment-arts-23195479, 5 July 2013, BBC News, downloaded January 19, 2014.

[14] http://www.telegraph.co.uk/technology/news/10334043/Broadcasters-failing-to-keep-up-with-3D-TV-demand.html, The Telegraph, downloaded January 19, 2014.

[15] http://www.etsi.org/technologies-clusters/technologies/broadcast/dab

[16] W. Hoeg and Thomas Lauterbach. "Digital Audio Broadcasting: Principles and Applications of Digital Radio". John Wiley, 2003.

[17] www.worlddab.org

[18] OECD Communications Outlook 2013. Digital audio broadcasting in OECD countries. DOI:10.1787/comms_outlook-2013-table119-en

[19] WorldDBM seminar", Riva del Garda, Italy, 14 April 2013.

[20] P. Hannon, "Digital radio in Europe", World DAB Workhsop, Budapest, January 28, 2014, downloadable from http://www.worlddab.org/events/detail/335#informations

[21] http://www.drm.org

[22] G. O. Driscoll: Next generation IPTV services and technologies. Wiley Interscience, 2009.

[23] E. Mikoczy, D. Sivchenko, X. Bangnan, and J. I. Moreno, "IPTV systems, standards and architectures: Part II. IPTV services over IMS: Architecture and standardization," IEEE Commun. Mag., vol. 46, no. 5, pp. 128–135, May 2008.

[24] Naeem Ramzan and Ebroul Izquierdo, "Scalable and Adaptable Media Coding Techniques for Future Internet", J. Domingue et al. (Eds.): Future Internet Assembly, LNCS 6656, pp. 381–389, 2011.

[25] László Bokor, László Lois, Csaba A. Szabó, Sándor Szabó, "Testbed of a Novel Media Streaming Architecture for Hetero-geneous Wireless Environment," Proc. Tridentcom07, the 3rd Int'l IEEE Conference on Testbeds and Research Infrastructures, Orlando, USA, May 21-23, 2007, 10 pages. Available in IEEE XPlore, http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4444715.

[26] http://www.itu.int/rec/T-REC-H.323-200912-I/en

[27] http://www.ietf.org/rfc/rfc3261.txt

[28] Third Generation Partnership Project (3GPP), "Technical specification group service and system aspects; telecommunication management; charging management; IP multimedia subsystem (IMS) charging (release 6)," Tech. Specification 3G TS 32.260 version 6.8.0, 2007.

[29] K. Knightson, N. Morita, and T. Towle, "NGN architecture: Generic principles, functional architecture, and implementation," IEEE Commun. Mag., vol. 43, no. 10, pp. 49–56, Oct. 2005.

[30] Gonzalo Camarillo, Miguel-Angel García-Martín: The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds, 3rd Edition, John Wiley & Sons, 2008.

[31] Miikka Poikselkä, Georg Mayer: The IMS: IP Multimedia Concepts and Services, 3rd Edition, John Wiley & Sons, 2009.

[32] A. Sanches-Esguevillas, B. Caro, G. Camarillo, Y. B. Lin, M. A. Garcia-Martin, and L. Hanzo, "IMS: The New Generation of Internet-Protocol-Based Multimedia Services", Proc. IEEE, Vol. 101, No. 8, August 2013, pp. 1860-1881.

[33] http://www.nets-fia.net

[34] J. Pan, S. Paul, and R. Jain, "A Survey of the Research on Future Internet Architectures", IEEE Commun. Mag. July 2011, pp. 26-36.

[35] Future Media Internet Architecture Reference Model - White Paper: FMIA-TT_Reference Model_Ver1 0 _20110301

[36] Edwards et al, Understanding Infrastructure: Dynamics, Tensions and Design, An NSF Report, Jan. 2007. Available on: http://cohesion.rice.edu/Conferences/Hewlett/emplibrary/UI_Final_Report.pdf

[37] C. Szabo, "European broadband initiatives with public participation," in: I. Chlamtac, A. Gumaste and C. Szabo (Eds.): Broadband Services: Business Models and Technologies for Community Networks, John Wiley&Sons, 2005.

[38] Csaba A. Szabó and Károly Farkas,, „Planning wireless municipal networks based on Wi-Fi/WiMAX mesh networks – applications, technologies and business models", tutorial lecture, ICC2010, Cape Town, South Africa, May 2010.

[39] C. Szabo, "Services to Meet Society-related Needs", in: I. Chlamtac, A. Gumaste and C. Szabo (Eds.): Broadband Services: Business Models and Technologies for Community Networks, John Wiley&Sons, 2005.

[40] C. A. Szabo, "Planning Wireless Cities and Regions To Support Telemedicine Applications", in Malina Jordanova (ed.): "Global Telemedicine and eHealth Updates: Knowledge Resources", Vol. 1, pp. 169-173. Publ. Luxexpo, Luxembourg, 2008, ISSN 1998-5509.

[41] Károly Farkas, Csaba A. Szabó, Zoltán Horváth, „Motivations, Technologies and Sustainability Models of Wireless Municipal Networks", IEEE Communications Magazine, Vol. 47, No. 12, pp. 76-83, December 2009.

[42] Microsoft Windows Media Technologies, http://www.microsoft.com/windows/windowsmedia/default.mspx.

[43] RealNetworks RealPlayer, http://www.realplayer.com/.

[44] Apple QuickTime, http://www.apple.com/quicktime/.

[45] http://tools.ietf.org/html/draft-pantos-http-live-streaming-12

[46] R. Buyya, M. Pathan and A. Vakali (Eds.), Content Delivery Networks. Lecture Notes in Electrical Engineering 9. Springer-Verlag Berlin-Heidelberg, 2008.

[47] www.akamai.com

[48] www.limelight.com

[49] Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. Cisco White Paper, May 29, 2013. Downloaded from http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html, on February 24, 2014.

Leonardo Chiariglione obtained his Ph. D. degree from the University of Tokyo in 1973. During his career he launched or participated in several initiatives. Among these are MPEG and the Digital Media Project (DMP). He is currently CEO of CEDEO.net a company developing its own portfolio of technologies, products and services in the area of digital media, such as WimTV (http://wim.tv/).
Leonardo is the recipient of several awards: among these the Eduard Rhein Foundation Award, the IBC John Tucker Award, the IEEE Masaru Ibuka Consumer Electronics Award and the Kilby Foundation Award.

Csaba A. Szabó obtained his MSc from St. Petersburg State University of Telecommunications, Dr. Univ., Ph. D. and Dr. Habil. from the Budapest University of Technology and Economics (BME) and the title Doctor of Technical Science from Hungarian Academy of Sciences. He is a Professor and Head of Laboratory of Multimedia Networks and Services at the Dept. of Networked Systems and Services of BME.
Prof. Szabo has long-term experience in academia, R&D and telecommunication business. He co-founded a leading network system integrator and service provider company that was subsequently integrated into Hungarian Telecom. Prof. Szabo has been a member of editorial boards and EiC of several international journals. He is Senior Member of IEEE.