

IMS: The New Generation of Internet-Protocol-Based Multimedia Services

An overview of the Internet protocol (IP)-based multimedia subsystem (IMS) infrastructure that relies on the session initiation protocol (SIP) is presented along with its services, applications, and future potential.

By ANTONIO SÁNCHEZ-ESGUEVILLAS, *Senior Member IEEE*, BELÉN CARRO, GONZALO CAMARILLO, YI-BING LIN, *Fellow IEEE*, MIGUEL A. GARCÍA-MARTÍN, AND LAJOS HANZO, *Fellow IEEE*

ABSTRACT | Legacy networks, both fixed and mobile, which were originally designed for voice communications, are progressively migrating to new infrastructures that promise to revolutionize the services offered. In this paper, we will cover this new generation of personal communication services, with an emphasis on the family of Internet protocol (IP)-based multimedia subsystem (IMS)-aided infrastructure that relies on the session initiation protocol (SIP). As a benefit, the end users will enjoy a new generation of personal communications services that are accessible anywhere and anytime. These services are directly related to the end users rather than to their diverse devices. It is anticipated that the new deployments of next-generation networks (all-IP based) will accelerate the adoption of the IMS technology.

KEYWORDS | Applications; Internet protocol (IP) multimedia subsystem (IMS); services; session initiation protocol (SIP); voice over IP

Manuscript received May 19, 2011; accepted October 22, 2012. Date of publication March 14, 2013; date of current version July 15, 2013. The work of Y.-B. Lin was supported in part by Academia Sinica AS-102-TP-A06 and the MoE ATU plan.

A. Sánchez-Esguevillas is with Telefónica, Madrid, Spain and also with the University of Valladolid, 47151 Valladolid, Spain (e-mail: a.sanchez-esguevillas@ieeee.org).

B. Carro is with the University of Valladolid, 47151 Valladolid, Spain (e-mail: belcar@tel.uva.es).

G. Camarillo is with Ericsson, 02420 Helsinki, Finland (e-mail: Gonzalo.Camarillo@ericsson.com).

Y.-B. Lin is with the National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: liny@csie.nctu.edu.tw).

M. A. García-Martín is with Ericsson, 28045 Madrid, Spain (e-mail: Miguel.A.Garcia@ericsson.com).

L. Hanzo is with the University of Southampton, SO17 1BJ Southampton, U.K. (e-mail: lh@ecs.soton.ac.uk).

Digital Object Identifier: 10.1109/JPROC.2012.2234431

ACRONYMS

3GPP	Third-generation partnership project.
AAA	Authentication, authorization, and accounting.
ARPANET	Advanced Research Projects Agency Network.
AS	Application server.
AT&T	American Telephone and Telegraph Company.
B2BUA	Back-to-back user agent.
BGCF	Breakout gateway control function.
BHSA	Busy hour session attempt.
BSC	Base station controller.
CAB	Converged address book.
CAPEX	Capital expenses.
CATV	Cable television.
CDF	Charging data function.
CDMA	Code-division multiple access.
CDR	Charging data record.
CGF	Charging gateway function.
CHT	Chung-Hwa Telecom.
CPM	Converged IP messaging.
CS	Circuit-switched.
CSCF	Call session control function.
CSFB	Circuit-switched fallback.
CTF	Charging trigger function.
CTI	Computer telephony integration.
DHT	Distributed Hash table.
DLNA	Digital living network alliance.
DNS	Domain name service.
DOCSIS	Data over cable service interface specifications.
DSL	Digital subscriber line.
ECC	Error checking and correcting code.
EDGE	Enhanced data rate for GSM evolution.
EVDO	Evolution-data optimized.
FB-DIMM	Fully buffer dual in-line memory module.

FMC	Fixed mobile convergence.	SCC-AS	Service consistency and continuity-application server.
FTTH	Fiber to the home.	S-CSCF	Serving CSCF.
GAN	Generic access network.	SDCC	Small device C compiler.
GANC	GAN controller.	SDOs	Standard Developing Organization.
GGSN	Gateway GPRS support node.	SDP	Service delivery platform, session description protocol.
GPRS	General packet radio service.	SIM	Subscriber identity module.
GSM	Global system for mobile communications.	SIP	Session initiation protocol.
GSMA	GSM Association.	SLF	Subscriber location function.
HLR	Home location register.	SMS	Short message service.
HSS	Home subscriber server.	SMTP	Simple mail transfer protocol.
HSPA	High-speed packet access.	SPAN	Services and protocols for advanced networks.
HTTP	Hypertext transfer protocol.	TBCP	Talk burst control protocol.
I-CSCF	Interrogating CSCF.	TEL URI	Telephone URI.
IETF	Internet Engineering Task Force.	TIPHON	Telecommunications and Internet protocol harmonization over networks.
IM	Instant messaging.	TISPAN	TIPHON SPAN.
IMS	IP multimedia subsystem.	UBB	Ultrabroadband.
IP	Internet protocol.	UE	User equipment.
ISDN	Integrated services digital network.	UMA	Unlicensed mobile access.
ITU	International Telecommunications Union.	UMTS	Universal mobile telecommunications system.
LTE	Long-term evolution.	UPB	Universal processor boards.
Mbps	Megabits per second.	UPnP	Universal plug and play.
MGCF	Media gateway control function.	URI	Universal resource identifier.
MGW	Media gateway function.	VCC	Voice call continuity.
MIME	Multipurpose Internet mail extensions.	VoIP	Voice over IP.
ML	Markup language.	VoLTE	Voice over LTE.
MMS	Multimedia messaging service.	WiMAX	Worldwide interoperability for microwave access.
MRF	Media resource function.	WLAN	Wireless local area network.
MRFC	Multimedia resource function controller.	WS-CSTA	Web services-computer supported telecommunications applications.
MRFP	Multimedia Resource Function Processor.	WWW	World wide web.
MSRP	Message session relay protocol.	XCAP	XML configuration access protocol.
MTA	Mail transport agent.	XDM	XML document management.
MTR	Mobile termination rate.	XDMS	XML document management server.
NAT	Network address translation.	XML	eXtensible markup language.
NGN	Next-generation network.		
O&M	Operation and maintenance.		
OCS	Online charging system.		
OMA	Open mobile alliance.		
OPEX	Operational expenses.		
P2P	Peer to peer.		
P2PSIP	Peer-to-peer SIP.		
PBX	Private branch exchange.		
PCM	Pulse code modulated.		
P-CSCF	Proxy CSCF.		
PDF	Policy decision function.		
PoC	Push to talk over cellular.		
PRACK	Provisional acknowledgment.		
PS	Packet-switched.		
PSTN	Public-switched telephone network.		
PTT	Push to talk.		
PTX	Push-to-X.		
RCS	Rich communication suite.		
RFC	Request for comments.		
RSTP	Real-time streaming protocol.		
RTCP	Real-time transport control protocol.		
RTP	Real-time transport protocol.		
SBLP	Session-based local policy.		

I. INTRODUCTION

A. Brief Telecommunications History

Since the conception of the first signaling system referred to as a semaphore, which took place in Europe in 1790, tremendous advances have been made, as exemplified by the invention of the telegraph by Morse in 1837, followed by the telephone invented by Meucci, Gray and, finally, Bell in 1876. In 1901, Marconi succeeded in transmitting a wireless message from Britain to the United States, which earned him the Nobel Prize in physics. The Advanced Research Projects Agency Network (ARPANET) was conceived in 1969, which may be deemed to be the predecessor of the Internet. The first-generation public mobile personal communications services commenced during the late 1970s and early 1980s.

At the time of writing traditional voice communications is still thriving, but it is no longer the main source of growth for the service providers. The market penetration of mobile voice services has reached and, in some cases, exceeded 100%, but it is still growing in emerging markets of the globe. Fixed-line-based broadband services still constitute a moderately growing market, with the expectation that UBB services, with downlink speeds in excess of 50–100 Mb/s mainly based on optical fiber, will constitute a growing market segment. Mobile broadband services have an even more rapid growth potential. Converged fixed-mobile services are also growing in popularity. From a societal perspective we are observing a transition from the generation of computer and Internet users to a new generation growing up with the wireless Internet.

The Internet has been rapidly growing since its inception, and it has become a mature data network with worldwide coverage. The system was initially rolled out across the academic and military communities in 1983 with the first commercial Internet services emerging in 1989 after the connection of the Internet to commercial e-mail services and by 1993 the WWW acquired wide commercial acceptance [1]. It is important to mention that Vinton Cerf, the father of the IP family, mentions also VoIP as the next key milestone in the Internet era: by 2003, VoIP services spread quite widely following its commercial introduction around 1995.

The evolution of voice telephony commenced over a century ago and led to public wireless mobile voice communications about 30 years ago, but still relied on plain old CS technology. With the advent of the Internet, the question is now no longer whether voice communications would also be embraced by the Internet [2], [3]—not even whether there would be other types of multimedia services carried over the Internet—but rather, how soon these advanced wireless multimedia services will reach the mass market.

B. Packet Switching: From VoIP Toward IMS

The era of the plain old CS voice services offered both by the fixed PSTN and by the wireless mobile cellular networks is gradually evolving toward PS environments. This new era was initiated with the advent of VoIP technologies, which were initially based on two popular solutions, namely on the ITU's H.323 protocol and on the IETF's SIP. Both of these relied on the RTP invoked for the sake of maintaining near-real-time interactive voice communications. At the time of writing, VoIP services are moving toward operator-grade fully fledged infrastructures based on the IMS protocol, which is the subject of this treatise. The VoIP was introduced as early as 1995 and, again, it relied on the H.323 protocol family derived from the ITU's ISDN, albeit the ISDN technology has now been essentially phased out. The H.323-based VoIP philosophy attracted large initial interests until about 2000. On the other hand, SIP, which relied on the Internet philosophy and

started from “scratch,” was initially published as an Internet Draft by the IETF in 1996, with the first RFC numbered as 2543 in 1999, which evolved to the RFC 3261 by 2002. SIP is a flexible signaling protocol that can support different system architectures [4], [5]. In the same year, the first IMS standard appeared as part of the 3GPP Release 5 [6], [7]. IMS uses SIP as its main signaling protocol for call control.

It is important to note that IMS can be combined with diverse transport technologies, such as the now ubiquitous DSL and its higher rate versions for copper infrastructure, and DOCSIS for CATV infrastructure. IMS can also be combined with both older legacy mobile data transmission technologies, such as GPRS, EDGE, EVDO, and UMTS, and with the more recent HSPA system, the emerging 3GPP LTE system [8]–[10], the WIMAX [11], [12], diverse WLANs [13] or even over satellite networks [14], [15], [31]. In a nutshell, IMS facilitates the transition from CS to PS communications and, hence, promises all the benefits of new powerful value-added services, as exemplified in [3], [16]–[30], [32], [78]–[83], and [85].

In order to support these services, IMS offers a series of building blocks referred to as service enablers, which include, in addition to classic voice services, presence [33], community/group management in the form of ML-based document management, the provision of location-based rich multimedia and messaging services, the support of a networked address book, user profile storage, the creation of gateways to legacy communication systems, etc. The OMA is one of the SDOs creating and standardizing service enablers relying on IMS. All in all, IMS can be considered as a versatile service provision infrastructure [34].

The first services that were launched based on IMS include PoC, VoIP, video sharing and even IP-TV. In general, VoIP-based services are typically developed first in the enterprise market. The first services offered on the consumer market were PoC-based services. IMS also plays a significant role in the process of FMC [35]–[88], where the user can be reachable through a single telephone number on either a fixed-line-based device, a computer, or a mobile device. The RCS initiative [37] represents an effort to speed up the development, testing, and introduction of rich commercial IMS-based communication services, which is fostered by GSMA. The RCS constitutes a collection of applications and services that will facilitate interoperable, enriched communications, including the provision of a sophisticated network-based phonebook, enhanced multimedia messaging, and enriched calls. Since 2012, consumers have had the option of enjoying RCS-aided (voice, video, presence, and instant messaging) communications. Furthermore, 2012 has seen the launch of mobile VoLTE based on IMS.

C. IMS Advantages and Disadvantages

IMS promises important advantages in terms of development cost reduction and time to market. But the question arises: Is this real or just vendor hype? VoIP and IMS

Table 1 IMS Advantages and Disadvantages

	Circuit switched voice	Voice over IP	
		Plain	IMS
Cheap tariffs	x->✓		✓
Infrastructure cost	✓		x
Operation costs	x		✓
Scalability&Reliability	✓	x	✓
Interoperability	✓	x	✓
Mobility	✓		✓✓
New services	x		✓
Video	x		✓
Audio quality	✓		x->✓
Security	✓		x->✓
Compatible with new access networks	x		✓✓

legend: x negative ✓ positive ✓✓ very positive x->✓ used to be negative but becoming positive

promise the aforementioned advantages, but critics may question, whether they are important enough to justify the associated migration, especially, because as time passes by, some of these competitive advantages of IMS might be eroded, since these enriched services might also be included in traditional systems. Let us analyze briefly the associated pros and cons, which can be summarized in Table 1.

- Economics: It might not be so lucrative after all in the short term.
 - Tariffs: VoIP-based systems offer reduced tariffs, especially for international calls. However, the classic telecommunication service providers also offer decreasing tariffs and new flat-fee plans. Indeed, regulators are drastically reducing the wholesale termination rates (e.g., MTR), claiming that they should be adjusted to costs (i.e., voice infrastructure investment payoffs over time), and this is being translated into retail tariffs. Of course, in general, tariffs depend more on pricing strategies and less on technology (which influences costs, as we can see below).
 - Operational cost (OPEX): VoIP equipment is more cost effective than traditional central office equipment, since it relies on a more lightweight distributed infrastructure (like the one used by Skype, based on P2P technologies). Moreover, the OPEX savings of main-

taining a single network (voice and data) are attractive, hence legacy data networks, such as X.25, ATM, etc., are being closed.

- Infrastructure costs (CAPEX): However, the associated new equipment requires a significant initial investment and investors have to carefully analyze how long the amortization period of the operational legacy equipment is and how little the migration costs are.
- Single network: The operating expenses of maintaining a single network seem attractive, but we cannot ignore that there are still many legacy networks (X.25, ATM, etc.) that have to be closed.
- End-user offerings: All experts tend to concur that this is certainly a strong motivation.
 - Scalability and reliability: There are more than seven billion phone subscriptions in the world, and there is no other service covering such a huge population. Moreover, it is well known that fixed (mobile has constraints in terms of coverage) telephony is extremely reliable since we can almost always make a phone call, except perhaps in extreme peak periods like New Year's first seconds). Therefore, scalability and reliability are often taken for granted in CS voice. IMS was also designed with scalability and reliability in mind, although it still has to prove it.

- Interoperability: Again classic voice telephony is completely interoperable (since we can call anyone anywhere). On the other hand, VoIP services have to be improved in this respect.
- Mobility: Personal mobility is undoubtedly one of the main advantages of decoupling the specific terminal from the particular service; however, mobile phones also provide mobility and terminal independence as a benefit of the SIM [38].
- New value added integrated services can be offered promptly, with a reduced time to market and at reduced development costs; however, they can also be provided by state-of-the-art intelligent traditional voice networks. As a matter of fact, it may be argued that VoIP/IMS does not enable any single new application that cannot be supported by evolved traditional systems. On the other hand, is there really a market for new services, such as crisp, high-resolution video telephony accompanied by high-fidelity audio, leading to the impression of flawless telepresence?
- Video: Support for video in CS networks is quite poor, given the conventional voice channels restrictions; high quality video can only be attained using many channels, which is quite expensive; therefore, its uptake has been marginal.
- Audio quality: The provision of poor voice quality has been closely associated with antiquated technology. This was mainly due to the employment of bandwidth-hungry, low-compression voice codecs, which were originally designed for high-quality fixed lines supporting the classic CS networks, rather than for error-prone PS networks, where packet loss events may be frequently encountered. More explicitly, in the early days, the 64 kb/s A-law and mu-law PCM codecs defined by the ITU's G.711 standard were used, which are sensitive to the loss of packets. However, the development of sophisticated voice codecs, which are capable of tolerating a packet loss ratio as high as 1%–3% has increased the perceptual voice quality. Congested IP networks operating without quality guarantees certainly degrade the user experience [39]–[41], [76], [85]. Using multipurpose terminals, such as, for example, PCs, which used to be equipped with modest-quality audio cards also used to erode the perceptual voice quality. With the advent of the recent developments, VoIP offers an increased grade of flexibility with the aid of improved multimedia codecs, despite their reduced bitrate.
- Security: IP networks are significantly more open than their traditional counterparts, where physical wiretapping was almost the only way to intercept communications in an unauthorized manner. Nonetheless, the introduction of sophisticated security measures is capable of supporting secure communications. On the other hand, security measures, such as firewalls and the provision of the so-called NAT techniques often made it a challenge to employ VoIP communications in the past. Fortunately, appropriate traversal mechanisms are now in place to avoid these limitations.
- Compatible with new access networks, hence it is future-proof: Switched circuits are no longer available in next-generation access networks, such as TTH and LTE, which are all-IP. As soon as these new networks are massively deployed, CS voice will progressively disappear. It is true, however, that this might take quite long.

At the time of writing, there are already commercial IMS-based services on the market, as detailed in the following sections. Major operators have evaluated their IMS-based pilot infrastructure, hence the launch of the related services is more of a business orientated, rather than technical decision. As seen in Table 1, the main limitation imposed on the widespread rollout of VoIP is the required investment required by the new infrastructure, whereas the main disadvantage of CS voice is the lack of compatibility with new access networks. These two factors are likely to determine the speed of the transition. On the other hand, the main theoretical differences between IMS and plain VoIP are related to scalability, reliability, and interoperability, as well as complexity. These features are typically associated with high-quality carrier-grade services, which are normally preferred by operators. The proliferation of successful VoLTE launches might be a good indication that IMS is finding favor in the market place, as discussed later.

D. Outline of the Paper

The paper is structured as follows. Section II covers the technology basics, describing the associated protocols, namely the SIP, the IMS core infrastructure architectures, and the IMS service support, relying on the so-called enablers. Section III portrays a range of end-user applications, commencing with the issues of fixed-mobile convergence support provided by IMS to exemplify the applications of the technology. Then, some of the recently launched services are described, such as PoC, VoIP, and video sharing. Apart from these individual services, both so-called RCS as well as mobile telephony based on VoLTE are detailed. We will demonstrate that the applicability of IMS is not limited to a single customer segment; indeed, it is capable of supporting both residential and business/corporation-level operations across diverse application sectors, as

exemplified by cutting-edge next-generation medical, educational, or intelligent transport scenarios. Finally, Section IV details the remaining technical challenges and a range of open research issues, before concluding.

II. TECHNOLOGY BASICS: PROTOCOLS, ARCHITECTURES, AND SERVICE SUPPORT

A. Call Control: SIP

As alluded to in our previous discourse, the SIP [42] is a signaling protocol designed to establish, manage, and tear down multimedia sessions on the Internet. Additionally, the capability of SIP also providing asynchronous notifications and announcements of events. It is also often used for fetching presence information from other users and establishing interpersonal communication sessions with them.

To elaborate a little further, the SIP is used for establishing multimedia sessions, but it does not act as a media transport protocol. The multimedia information associated with specific sessions established by the SIP is exchanged using a variety of different protocols. For example, the RTP [43] is used for transporting audio and video streams set up with the aid of the SIP. The MSRP [44] is used for transporting a sequence of arbitrary interactive multimedia objects, including instant messages, pictures, video clips, or, in fact, any other multimedia files. Rather than creating a monolithic architecture, a range of appropriate Internet multimedia protocols are combined in different ways in order to support different services.

The SIP's main role in establishing a session may be viewed as that of providing a rendezvous function, which ensures that callers can reach the called subscriber at his/her current location. Once a callee is located, the SIP provides a two-way message exchange for establishing the corresponding sessions, which may consist of multiple media streams of different types. This two-way message exchange is referred to as the offer/answer model [45]. Both messages involved in the offer/answer exchange, which is simply referred to as the offer and the answer, are encoded using a specific session description format. The most common session description format is the so-called SDP [46]. However, the SIP is independent of the particular session description format employed. Indeed, the SIP can be used in conjunction with session description formats other than the SDP.

To elaborate a little further, the SDP relies on a text-based session description format used by protocols such as the SIP and the RSTP. The SDP allows the endpoint of a session/call to describe the characteristics of the sessions to be established, which may include but are not limited to media types, the multimedia codecs used, the transport addresses, the bandwidth, etc. When two endpoints exchange their SDP identifiers as part of an offer/answer

exchange, they essentially agree about the parameters to be used for establishing the session.

The SIP infrastructure of a network handles the routing of SIP messages between the SIP endpoints, which are referred to as SIP user agents. The SIP defines the following logical entities: registrars, proxy servers, and redirect servers. The user agents register their current location with their domain's registrar so that all the incoming SIP messages can be successfully routed to their up-to-date location. The other servers, namely the proxy and redirect servers, use the data stored by the registrars for performing recursive and iterative routing of the SIP messages, respectively.

The routing procedures of the SIP follow a similar approach to routing in the SMTP and HTTP. The aforementioned proxy servers route the SIP messages in a similar way to MTAs, which route e-mail messages. Invoking a specific service in SIP is performed by routing the SIP messages to the relevant application servers, which act as user agents or proxy servers, depending on the specific service being provided. Some application servers act as so-called B2BUAs. A B2BUA is a SIP network entity that partitions the session and the media into two separate call legs. Therefore, B2BUAs provide some sort of isolation at both the signaling and the media level between the two endpoints; the B2BUA application effectively bridges the two call legs. B2BUAs have more freedom than proxy servers in terms of session manipulation, thus they are typically used for providing those particular services that proxy servers are unable to provide.

Before we explore the more intricate details of SIP routing, we have to understand what a SIP registration is and how the associated procedures are executed. Registration is the procedure whose result is the binding of a user's SIP URI—which is typically allocated to a user—to a contact address that is typically an IP address or a fully qualified domain name that effectively represents an IP address. The SIP registration procedure is carried out with the aid of the SIP REGISTER messages. When the procedure is successfully concluded, a SIP registrar—recall that this is a function of a SIP proxy server—stores the binding of the SIP URI to the contact address for a limited period of time. Registrations are refreshed before they expire.

There are two main ways in which SIP routing can be implemented. The first one is the traditional technique, which is based on the DNS. The SIP entities query the DNS in order to find the specific proxy server responsible for a given domain. When the proxy server that is responsible for a user receives a SIP request, it searches for the binding of the user's SIP URI to a specific IP address, and then routes the request to that particular IP address, which is allocated to the device utilized by the user. The second of the aforementioned two routing mechanisms is typically referred to as P2PSIP. In P2PSIP, the binding between the user's SIP URI and its contact address is not stored in a centralized proxy server, but instead, in an overlay [47],

which may be based on a DHT. Effectively, the overlay behaves like a distributed database containing the location information.

The SIP is a text-based protocol and its encoding was inspired by the HTTP, hence its request and response messages are very much like HTTP requests and responses. The SIP messages carry objects such as session descriptions in the same way as e-mail messages carry attachments, because the SIP uses MIME to transport objects.

It is vitally important to facilitate the flexible extension of the SIP. While most SIP implementations support a limited range of basic SIP functionalities, different implementations tend to support different extensions, depending on the specific services they were designed to support. The SIP also includes a range of mechanisms designed for extension negotiation. These mechanisms allow the diverse implementations to agree on the specific extensions to be applied to a particular session.

The SIP may be invoked in different system architectures. The IMS constitutes a particular example of a system architecture that uses SIP as its main signaling protocol. As for routing, the IMS uses the DNS-based traditional SIP.

B. IMS Context and Technology

1) *Standard Evolution—Access Networks*: The 3GPP initiated the standardization of IMS back in 1999 based on preliminary studies conducted by a group referred to as 3G.IP. The first IMS set of specifications was included in the 3GPP recommendation release 5, back in early 2002, and supported GPRS access to IMS. The 3GPP release 5 document also outlined the evolution of the GSM and WCDMA standards in addition to the new IMS network. The 3GPP release 6 document added the description of interworking with WLANs and that of a range of additional services. As a further advance, 3GPP release 7 incorporated the support of fixed networks [48], which relied on collaboration with the TIPHON as well as with the SPAN initiative. The TIPHON and SPAN groups were then referred to as the TISPAN initiative. The 3GPP release 8 inspired the new TISPAN release 2 specifications and, with the aid of the corresponding 3GPP2 efforts altogether, created the 3GPP IMS as the common IMS for all access types.

3GPP release 8 (2008) represents an important milestone with the introduction of all-IP networks as part of the LTE first release, which then evolved further to release 9, release 10 associated with LTE advanced (releases 11 as well as 12 are in progress; releases 11 and 12 might inductively be completed in June 2013 and June 2014, respectively); The evolved IMS standardization includes topics like service continuity handover of media streams between different accesses, emergency calls, etc.

As already mentioned, this all-IP paradigm might lead to dispensing with CS voice and its replacement by PS voice aka VoIP. With LTE becoming commercially deployed by an increasing cohort of operators, CS mobile

voice will be progressively phased out at latest, when second-generation (2G) and third-generation (3G) networks are switched off, as detailed later in the VoLTE section.

In summary, even though the IMS was initially conceived for mobile networks, at the time of writing, it also includes fixed access technologies. Again, the core functionality of the IMS was standardized by 3GPP. A number of NGN initiatives, such as the ETSI TISPAN [49], refer to the 3GPP core IMS and extend its scope of operation to different access technologies. The IMS plays a significant role in FMC, since a single network is capable of handling multiple different accesses for the same user.

2) *IMS Components*: The IMS is a system architecture designed for supporting multimedia services for transmission over different PS access technologies. Again, call control in the IMS is based on the SIP. Some of the IMS SIP entities are responsible for interfacing non-SIP IMS network nodes using different protocols designed for AAA, just to mention a few. Fig. 1 illustrates a simplified IMS architecture, where an IMS user is represented by UE (node *a* in Fig. 1).

In order to be able to support the IMS, the classic HLR of wireless systems was further developed for creating the HSS (node *b* in Fig. 1), which is the master database containing all user-related subscription information. The HSS consists of the user database, the HLR functionality required by the PS domain—as exemplified by the 3G GPRS HLR—and the CS domain—such as the 3G CS HLR—in order to provide support for all the call handling entities. To elaborate a little further, it keeps a master list of all the features and services associated with a specific user, such as their location and user profile information, including the corresponding user identities, the subscribed services, as well as the relevant numbering and addressing information. The SLF (node *c* in Fig. 1) is required for finding the specific HSS that stores the subscriber's data user profile.

The IMS defines a set of network entities termed as the CSCFs (nodes *d*, *e*, and *f* in Fig. 1) and behave as SIP proxy servers or as the previously defined B2BUAs. A CSCF communicates with the HSS for the sake of exchanging the relevant location information, which is required for handling the control-layer functions related to the application-level registration and to the SIP-based multimedia sessions. It is worth noting that if the home domain contains more than one HSS, the CSCF will communicate with the SLF and will find the appropriate HSS based on the relevant user profile. The CSCF processes all the call requests received from the other VoIP call control servers or terminals in IP-based multimedia networks. There are three main types of CSCFs, depending on their particular function, namely, those that are responsible for interrogating, those acting as a proxy, or those that are serving. The I-CSCF determines how to route mobile calls

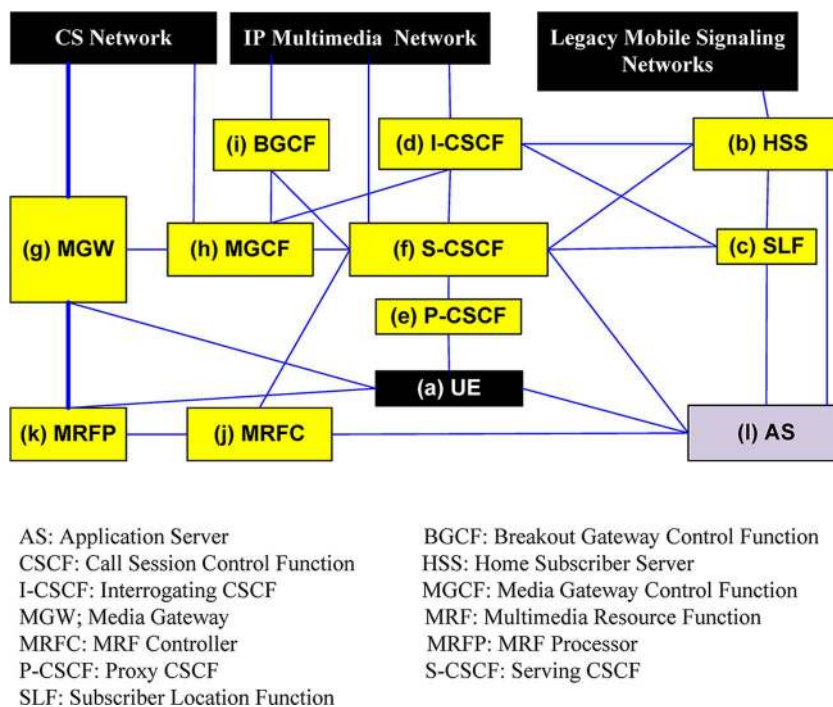


Fig. 1. The IMS architecture.

to the roaming destination UEs. In other words, the I-CSCF is the contact point for the home network of the destination UE, which may be used for concealing the configuration, capacity, and topology of the home network from the outside world. When a UE joins the network, a P-CSCF is assigned to the UE through the so-called Gm interface. The P-CSCF has only a limited subset of the CSCF functions, such as address translation, which may be used for forwarding the relevant request to the I-CSCF of the home network. Authorization of the bearer resources in the network is performed by a P-CSCF. Upon the completion of the application-level registration, an S-CSCF is assigned to serve both the UE during the call setup as well as the supplementary services control responsible, for example, for service requests and authentication.

The MGW (node g in Fig. 1) provides a gateway toward the conventional PSTN. It terminates the bearer channels arriving from the PSTN or legacy mobile networks as well as the media streams received from a packet network, such as an RTP-based stream arriving from an IP network. The MGW supports multimedia processing, including media conversion, bearer control, and payload processing, as exemplified by the appropriate choice of the multimedia codec, the echo canceller, and a conference bridge, just to name a few.

The MGCF (node h in Fig. 1) supports various call models and controls the connection of the media channels in an MGW. Through an H.248 interface, the MGCF interacts with the MGW in order to support flexible

connection handling. The MGCF communicates with an I-CSCF through the SIP for handling calls originated in the conventional CS domain. Specifically, the MGCF forwards the SIP call to an I-CSCF for handling incoming calls arriving from legacy networks. Then, the I-CSCF treats the call like any other multimedia call and forwards it to the specific S-CSCF that was allocated to the particular user that was called.

The BGCF (node i in Fig. 1) plays an important role in handling IMS-originated calls destined to the PSTN. The BGCF is responsible for selecting the appropriate PSTN breakout point based on the SIP request received from the S-CSCF. If the BGCF determines that the breakout is to occur in the same network, then the BGCF selects an MGCF, which is responsible for interworking with the PSTN. If, however, the breakout is in another network, then depending on the particular configuration, the BGCF forwards this SIP request to another BGCF or to an MGCF in the selected network.

The MRF supports multiparty calls, multimedia conferencing, tones, and diverse announcement functionalities. More specifically, the MRF has two main constituent parts, namely the MRFC (node j in Fig. 1) and the MRFP (node k in Fig. 1). The MRFC communicates with the S-CSCF for the sake of service validation during multiparty/multimedia sessions, while the actual media processing resources reside in the MRFP.

The AS (node l in Fig. 1) implements IMS services, such as PoC, the detection of subscriber presence, and so

on. It also interacts with the HSS, the S-CSCE, and the MRFC for the sake of maintaining service control and may directly communicate with the UE, for example, for the sake of service parameter setup and configuration. The MRF may also provide a resource management function in order to facilitate for the media resources to become common resources shared among multiple applications.

3) *IMS Call Flow*: The IMS session setup has several variants, mainly depending on whether the UE has already been assigned a bearer for sending the multimedia stream. Figs. 2 and 3 show the basic session setup procedures in IMS between two mobile UEs, neither of which has been assigned resources for transmitting multimedia streams prior to the session setup. According to Fig. 2, the first IMS terminal transmits the SIP INVITE request of Fig. 2 (1) containing the description of the desired media and multimedia codecs supported by this specific session. At this point in time only signaling information is sent. This INVITE request is then routed through the P-CSCF and S-CSCF allocated to the originating IMS terminal. The originating S-CSCF evaluates the set of initial filtering criteria, which is constituted by a set of rules that determines when the signaling information should be forwarded to one or more application servers. For the sake of simplicity, Fig. 2 assumes that no services are triggered, hence the signaling does not have to be forwarded to any application server. Then, the S-CSCF inspects the intended terminating network, finds an I-CSCF entry point to that particular terminating network with the aid of the DNS and forwards the INVITE request seen at stage Fig. 2 (5) to the corresponding I-CSCF. Then, the I-CSCF queries the HSS [Fig. 2 (7)] to find the address of the S-CSCF allocated to the terminating user and forwards the INVITE request [Fig. 2 (9)] to it. The terminating S-CSCF evaluates the terminating user's initial filtering criterion and potentially forwards the INVITE request to one or more application servers, which are not shown in the figure to avoid obfuscating details. Then, the S-CSCF looks up the terminating user's registration information, namely the address of the allocated P-CSCF and the address of the destination user's terminal, and then forwards the INVITE request to that P-CSCF [Fig. 2 (11)], which in turn forwards it [Fig. 2 (13)] to the destination IMS terminal. The destination IMS terminal may prealert the user at this moment in time; however, in the example considered, the IMS terminal remains unable to alert the user, because no resources are allocated resources for the transmission of the multimedia stream at this point in time.

The IMS terminal then sends the session progress SIP response [message 183 in Fig. 2 (15)] back to the originating IMS terminal through the same set of nodes that the INVITE request traversed. This session progress response also contains the supported media streams and codecs at the IMS terminal that was called.

When the originating IMS terminal receives the 183 response [Fig. 2 (20)], it activates its resource reservation for transmission of its multimedia streams, a process that may take a bit of time, depending on the specific type of the access network. Additionally, the originating IMS terminal generates an acknowledgement message in the format of a SIP PRACK request [Fig. 2 (21)], indicating that it has commenced its resource reservation process, but resources are not available as yet. When the terminating UE receives this PRACK request [Fig. 2 (25)], it starts its own resource reservation process.

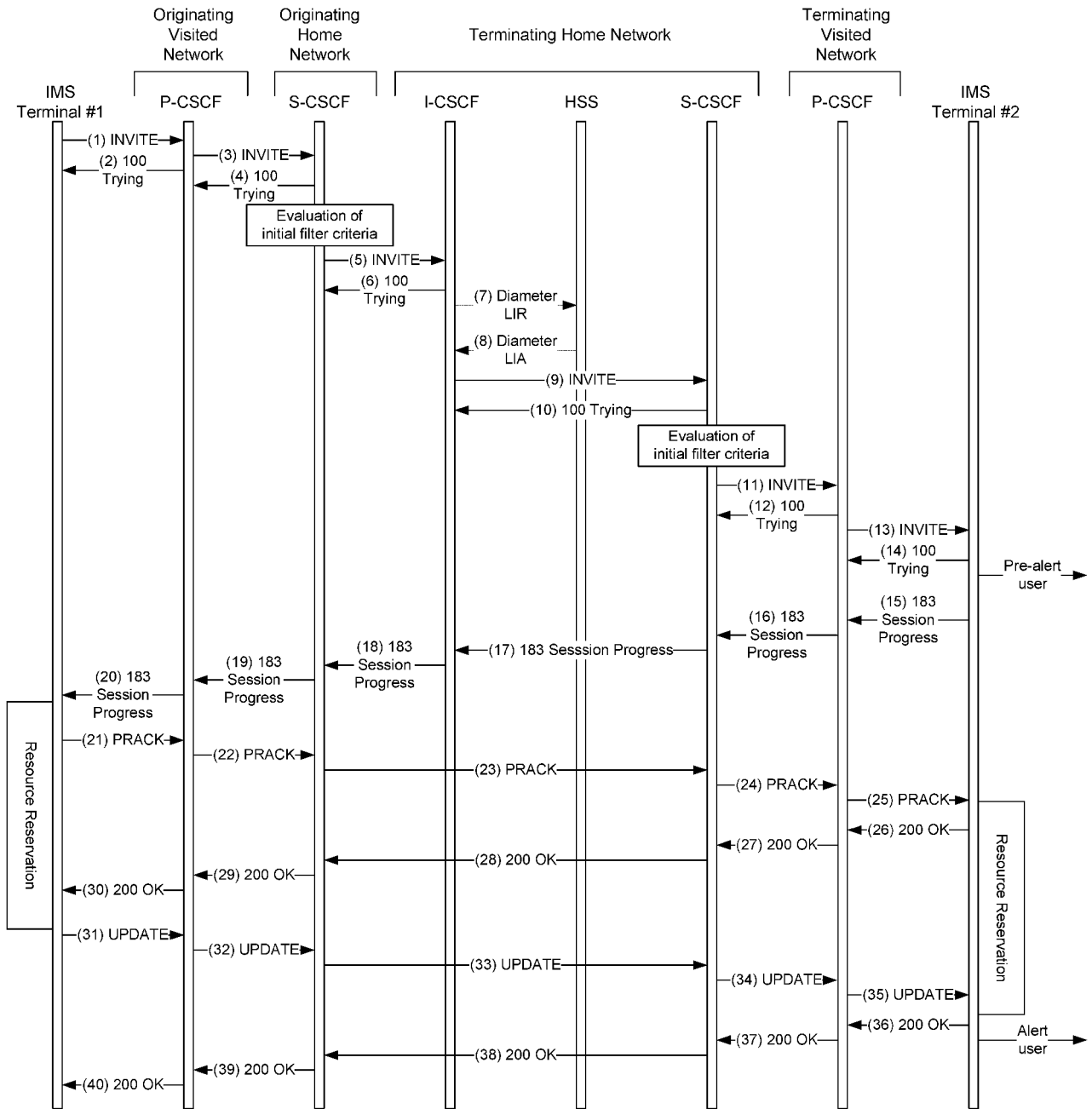
When the originating terminal completes its resource reservation process, the bearers become available for both sending and receiving the multimedia streams; hence, the session may now proceed from the originating terminal's point of view. This is communicated to the called terminal in an UPDATE request [Fig. 2 (31)].

When the destination terminal receives this UPDATE request [Fig. 2 (35)] and its own resource reservation was also completed, then the terminal rings the user and reports this to the originating side [Fig. 3 (41)]. The reception of the Ringing response [message 180 of Fig. 3 (46)] at the originating terminal is used for generating a local ring-back tone.

Eventually, the called party will accept the incoming session, which is signalled with the aid of the (OK) response [message 200 in Fig. 3 (57)]. This is then returned all the way to the originating terminal [Fig. 3 (62)], which generates an ACK request [Fig. 3 (63)] in order to confirm its reception. At this point in time, multimedia communications can be established between the two parties. The media packets are generally conveyed end to end, i.e., they do not traverse through the signaling nodes.

4) *Hardware Platform*: Fig. 4 illustrates an example hardware platform designed for the IMS nodes such as the HSS and CSCF. A total of 12 UPBs were installed in a subrack of this hardware platform. Every UPB relies on two Intel Xeon 5138 (2.13 GHz) low power-consumption dual-core processors. Each processor has 4-MB level-2 cache. The UPB supports a maximum of four FB-DIMMs. The storage capacity of a memory module may be 512 MB, 1 GB, 2 GB, or 4 GB. Therefore, the maximum memory capacity of the UPBs is 8 GB, and they are also capable of employing ECCs and an SDCC. A UPB has two Base interfaces (Ethernet 10/100/1000 M Base-T), two fabric interfaces (Ethernet 1000 M Base-BX), and one Update interface (Ethernet 1000 M Base-BX).

Based on this platform, an example of the HSS may be capable of supporting up to ten million IMS subscribers, which is sufficiently high for any realistic teletraffic scenario across the globe. The HSS guarantees a 99.9% message delivery success ratio. The transaction delay is less than 500 ms and the user registration time is less than 1 s. This CSCF product is capable of implementing I-, P-, and S-CSCFs.



Basic session setup, part 1

From *The 3G IP Multimedia Subsystem (IMS)*, Third Edition
 Gonzalo Camarillo and Miguel A. García-Martín
 © 2008 John Wiley and Sons, Ltd.

Fig. 2. IMS basic session setup, part 1.

5) *IMS and Legacy Networks*: As IMS is progressively introduced in the operational networks, there will be a need for supporting both new and existing services in either the conventional CS or in the IMS mode, or, in fact, both in the CS and IMS modes. In the long term, all

services are likely to be provided using the IMS. However, there will be a nonnegligible transitional period, where some services will be offered to a user over both CS and IMS. It is important, therefore, that during this transitional period the user experience remains favorable.

One of the goals of the IMS is to provide a unified way of handling communications not only for operators but also for users. The IMS users should be able to communicate with all users supported by legacy networks, such as the PSTN. The IMS architecture includes gateways toward other networks in order to provide its users with cross-network communications. For example, the gateways can convert VoIP sessions of the IMS to CS calls of the PSTN or SIP-based instant messages of the IMS to SMSs on a cellular network. Thus, the IMS enables the integration of different communication islands into a single universal communication network. It also allows users to communicate with other users in a universal way, regardless of the specifics of the network or of the communication community to which these users connect. In this way, all isolated personal communication islands may be con-

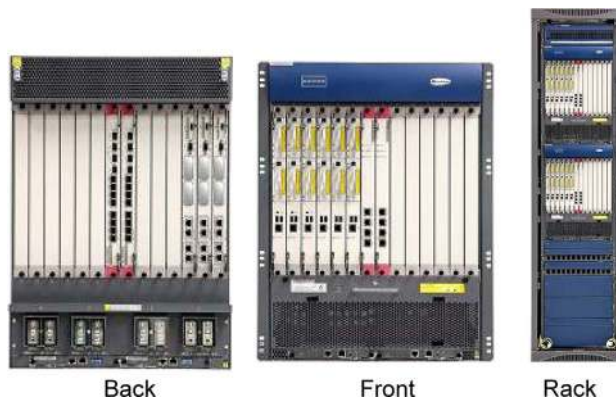
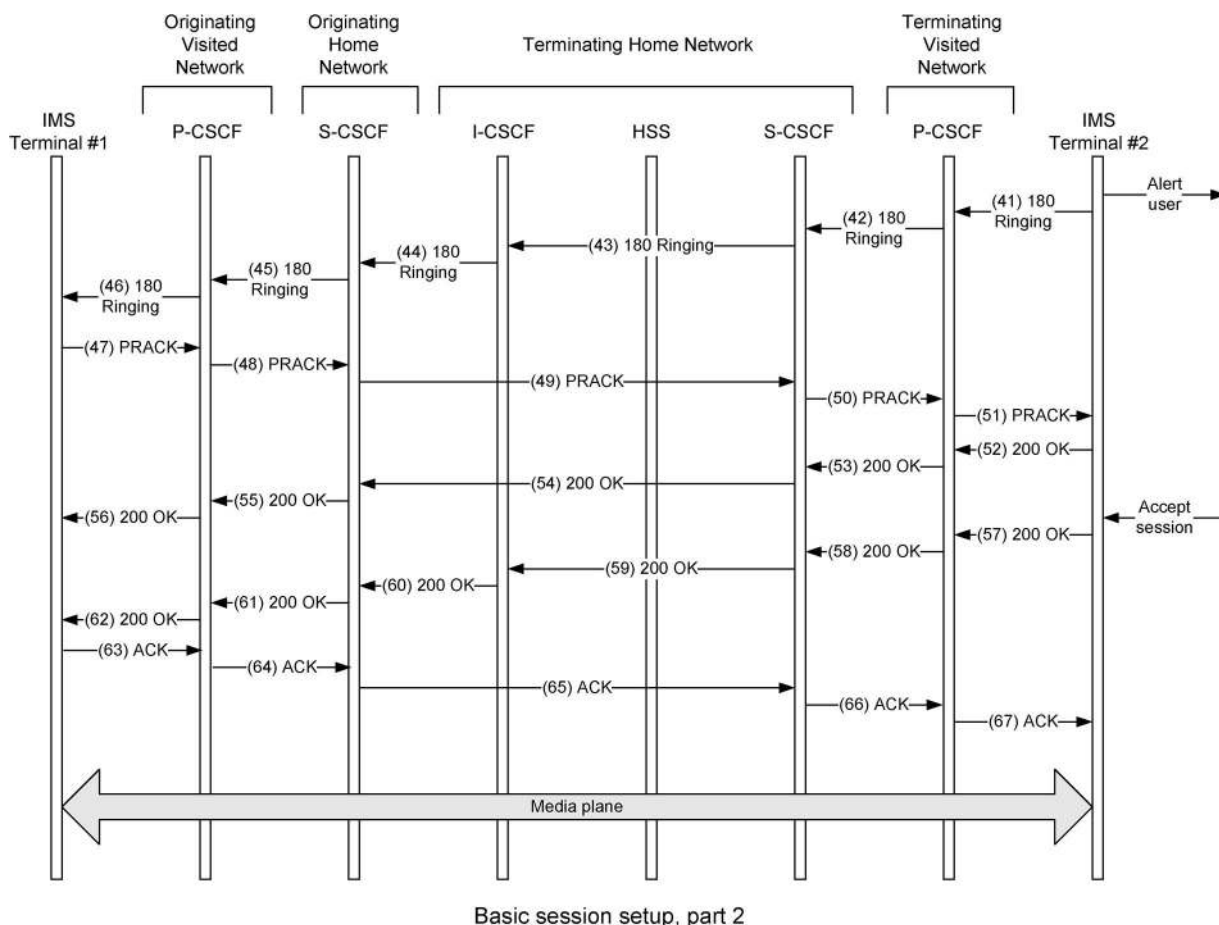


Fig. 4. Huawei's hardware platform for the IMS nodes (reproduced by permission of Huawei Technologies Co., Ltd.).



From *The 3G IP Multimedia Subsystem (IMS)*, Third Edition
 Gonzalo Camarillo and Miguel A. García-Martín
 © 2008 John Wiley and Sons, Ltd.

Fig. 3. IMS basic session setup, part 2.

nected in a unified fashion, regardless of the specific communications services requested.

6) *IMS and Service Delivery Platforms*: In the above discussions, we have shown that IMS may be considered as a multimedia SDP, which is capable of supporting arbitrary services provided by telecommunications operators or by third parties, such as operators providing services in several countries. This flexible deployment may be achieved without the need for any specific developments, which of course translates into a potentially wider service offer. The SDP concept is wider than that of the IMS, and it may be viewed as an additional call control layer, including additional higher layer components designed for service “orchestration” (composition and integration), Web services, etc. As a matter of fact, SDP architectures might consider the IMS as one of its underlying networks. However, these issues are beyond the scope of this paper.

C. IMS Service Support

The main idea behind the IMS is to facilitate innovative service creation and delivery. In order to achieve this ambitious goal, the IMS implements functionalities that are common to most services. In this way, this common functionality does not have to be repeatedly implemented for each different service. In the IMS, the services are horizontally integrated, as opposed to being vertically integrated. Therefore, the services relying on the IMS are capable of taking advantage of the functionality provided by the IMS. This functionality includes SIP-based reachability, service capability negotiations, registration, authentication, integrity protection, confidentiality, service invocation, service routing, charging, O&M, user management, and group management.

1) *Enablers*: As already mentioned above, the OMA is one of the standard development organizations conceiving and ratifying service enablers based on the IMS. More specifically, the OMA IMS (v1.0 release) defines the architecture for all OMA exploitation of the IMS environment in the OMA specifications. In particular, the following *Enabler* specifications have been released (either approved or in the form of Candidate release); in some cases a second version is specified, incorporating new features, such as the new media types supported, the performance enhancements introduced, etc.), complemented by so-called *Reference* (set of rules and guidelines) releases.

- Presence-SIMPLE v2.0, and its data extensions v1.3 reference: It defines user presence and availability attributes for a user.
- PoC v2.1: It defines a mechanism for a group of mobile subscribers to use their mobile phones as half-duplex radios, while relying on the concept of floor control for one-to-many communications. The latest version is also known as PTX, which may be viewed as the multimedia-based evolution capa-

ble of supporting both audio and video streaming, discrete content exchange, including images, text, video, and audio clips, as well as sessions with multiple groups.

- SIMPLE-IM v1.0 (v2.0 candidate): Instant messaging using IMS.
- XDM v2.1 (v2.2 candidate) (and XDM V1.1 and presence SIMPLE V1.1 implementation guidelines reference): Group management of users and other resources.
- CPM v1.0: It aims for combining all messaging techniques relying on IMS: short messages (text), multimedia messaging, instant messaging, PoC, mobile e-mail, etc.
- The CAB v1.1 candidate (simplified CAB v1.0 candidate, RESTful Network API for CAB v1.0 candidate): This facility allows a single address book to be used by all the messaging techniques running under an IMS-based environment.
- SIP Push v1.0: It is an updated version of the push enabler designed for the IMS environment, which allows for server-initiated messages to be sent to the mobile after authentication and authorization checks.
- Location in SIP/IP core v1.0 (LOCSIP): It constitutes an updated version of the mobile location service enabler conceived for IMS, which operates in conjunction with the presence enabler for the sake of providing location information either when requested, or periodically, or when the user enters a specific area.

It is also worth mentioning that a few of the OMA groups now consider the support of IMS in the context of both current and new feature developments, since IMS is now firmly recognized as a valuable infrastructure that will remain important for the industry. Therefore, all of the OMA enablers have to be based on IMS. By the same token, external groups now refer to the OMA enablers as part of their specification development, including the aforementioned TISPAN, which has adopted Presence SIMPLE, XDM, and PoC. In this sense, OMA continues to promote its mission of being bearer independent, hence encompassing more and more diverse communications environments.

2) *Presence Enabler*: The subscriber presence detection allows an IMS user to access another user’s presence information, including the user’s status, his/her specific activities (e.g., working, playing, etc.), the e-mail/phone addresses, the user device’s capabilities, such as whether it can receive a video call or a picture message, and so on. The IMS subscriber’s presence is considered as a service enabler, which implies that other services may use it to facilitate advanced services, and the mobile devices will be enabled to act as personal-networking devices.

Fig. 5 illustrates a simplified presence service architecture, where an UE—the presence service user—interacts

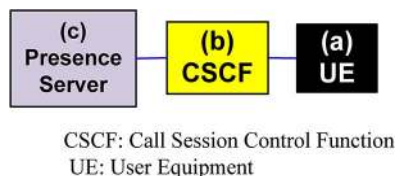


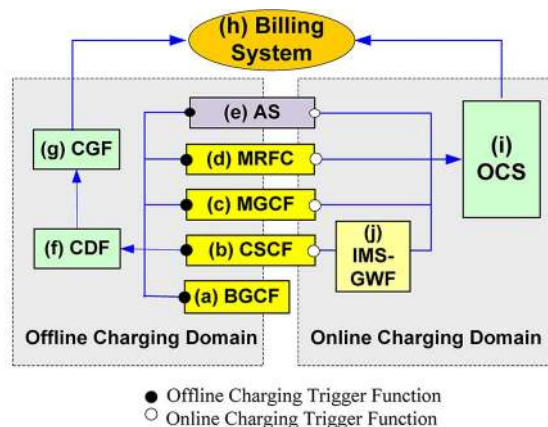
Fig. 5. Simplified presence service architecture.

with the presence server with the aid of the CSCF. The presence procedures include subscription, publication, and notification [50], [51], which are implemented by utilizing the SIP.

A user who accesses the presence information of other users is referred to in parlance as a “watcher,” while the subscriber, who provides the presence information, is the “presentity.” In order to subscribe to the presence information, the watcher carries out the subscription procedure by communicating with the presentity’s presence server with the aid of exchanging the SUBSCRIBE and the 200 OK message pairs, which were introduced in the context of Fig. 2. In order to change the presence information—for example, by changing the activity from “playing” to “working”—a presentity carries out the publication procedure by interacting with the presence server upon exchanging the PUBLISH and the 200 OK message pair highlighted in the context of Fig. 2. The presence server forwards the subscribed presence information to an authorized watcher with the aid of the notification procedure by exchanging the NOTIFY and 200 OK message pair.

The provision of the presence information is a potential service in its own right, depending on how the service is introduced and what the handset capabilities will be. In a practical scenario, the provision of presence information facilitates other services, such as instant messaging, the availability of a presence-aware address book, PoC, push to view, and so on. The role of the presence service in PoC is further elaborated on in the context of Section III, related to the available end-user services and applications. Since the presence server maintains real-time presence information, the IMS network may experience a high SIP traffic. This issue is detailed in [52].

3) *Charging*: The IMS charging functionality provides a single, cost-effective solution for consolidating all charging capabilities, so that the telecom operators may readily launch new services, pricing plans, and loyalty programs. Fig. 6 illustrates the IMS charging architecture, where the IMS nodes involved include the BGCF, the CSCF, the MGCF, and the MRFC of Fig. 6(a)–(d). The AS [Fig. 6(e)], which offers value-added IP-based multimedia services, resides either in the user’s home network or in a third-party location [53], [54]. The IMS supports both offline and online charging.



AS: Application Server
CDF: Charging Data Function
CSCF: Call Session Control Function
IMS: IP Multimedia core network Subsystem
MRFC: Media Resource Function Controller
BGCF: Breakout Gateway Control Function
CGF: Charging Gateway Function
IMS-GWF: IMS Gateway Function
MGCF: Media Gateway Control Function
OCS: Online Charging System

Fig. 6. IMS charging architecture.

The offline CTF represented by the blocks marked by bullets in Fig. 6 is a mandatory function integrated in all network nodes [Fig. 6(a)–(e)]. The CTF provides metrics that identify both the users as well as the network resources occupied by them, and generates chargeable events from these metrics. It then forwards the offline charging information to the CDF of Fig. 6(f) through the interfaces provided by a so-called accounting protocol. The CDF processes the charging information and constructs it in a well-defined CDR format. The CDR is then transferred to the CGF of Fig. 6(g). The CGF performs CDR preprocessing, including functions, such as the validation, consolidation, reformatting, error handling, persistent CDR storage, CDR routing and filtering, as well as CDR file management. Then, it passes the consolidated offline charging data to the billing system block of Fig. 6(h) using a common, standard file transfer protocol, such as FTP or SFTP [55], [56]. The CDR files may be transferred in either push or pull modes.

In contrast to the offline charging, the IMS online charging is carried out by the OCS of Fig. 6(i), which handles the subscriber account balance and the charging transaction control [57], [58]. The operator invokes the OCS for ensuring that the relevant credit limits are enforced and the resources requested are authorized on a per-transaction basis. Online charging for IMS is achieved by interacting with the OCS through interfaces provided by the credit control protocol. The IMS-GWF block of Fig. 6(j) represents a specific SIP application server, which provides protocol translation between the IMS service control and the credit control protocol.

The OCS activities of the IMS nodes are initiated by the online CTF. This CTF provides similar functionalities as the offline CTF described above, complemented by several enhancements that support online charging. These

enhancements include requesting, granting, and managing resource usage permissions according to what the OCS grants or denies. In other words, the online CTF is capable of interrupting a service, when the user runs out of credit. The online CTF delays the actual service usage until the permission is granted by the OCS. If the permission is not granted, the CTF denies, blocks, or cuts the service usage. Such a decision is made based on the pricing mechanisms in the OCS. Further intricate details of the IMS charging protocols can be found in [59] and [84].

III. END-USER SERVICES AND APPLICATIONS

The first commercially launched IMS services included PoC and IP Centrex. In general, VoIP-based services are being initially developed in the enterprise market and virtually all major operators have deployed their pilot schemes of the IMS infrastructure.

A. Fixed-Mobile Convergence

An emerging trend in the telecommunications industry is the provision of the same set of services over both fixed and mobile access networks. Before highlighting a range of new services, it is worth mentioning that IMS supports fixed-mobile network convergence in order to achieve cost savings by operating a single core IP network. Furthermore, the system is also capable of integrating potentially complex fixed legacy networks, as well as separate CS and PS networks into a single IP network for the sake of supporting rich multimedia services.

However, even before IMS networks reach their full potential, we can identify beneficial hybrid scenarios for fixed-mobile convergence, such as, for example, the provision of seamless handover of an active call between different networks. To elaborate a little further, the so-called GAN—which is also referred to as the UMA initiative—and the VCC techniques pave the way for fixed-mobile convergence. The UMA technique was created in 2004 and it was then adopted as the GAN solution by 3GPP release 6 in 2005. The UMA basically allows WLAN-enabled phones to access mobile voice networks with the aid of standard mobile voice signaling protocols. The core network remains unaltered and only a single new element, namely the GANC, is introduced at the same level of hierarchy as the BSC, so that the network handles the call as if the user were connected through another base station. It is also possible to manage call switching between a fixed and mobile access during a phone call, while guaranteeing the quality of service. This action may also be referred to as a handover. Since UMA is integrated into the mobile network, it handles this handover as transparently as a change of base station during a conventional BS handover.

However, UMA did not consider interoperability with IMS until the completion of the 3GPP release 8 recommendation, where all access technologies connect to an all-

IP network also for voice communications. An alternative approach is based on the more recent concept of femtocells, which are typically illuminated by radio waves emanating from broadband access routers [89]. The creation of femtocells eliminates the need for WLAN support, which is typical only in high-end smartphones that usually require an increased power.

On the other hand, the aforementioned VCC concept emerged later, which supports handovers between CS and PS networks. The IMS was conceived for delivering calls over PS networks. However, the associated implementation details are beyond the scope of this paper, especially in the context of CS networks requiring intelligent signaling. From the user's perspective, it is quite beneficial to exploit the fixed-line-based cost savings without sacrificing the mobility that may be relied upon even during a call. It is also interesting to note that PS voice networks are capable of seamlessly interworking with existing mobile networks, even during a specific conversation, which supports seamless migration. Finally, it is worth mentioning that the VCC philosophy is gradually evolving into a wider concept that would extend beyond the scope of voice services under the IMS service continuity concept incorporated in the 3GPP release 8 document.

B. First Services Launched: PoC, VoIP, and Video Sharing

As already mentioned, IMS might become the next-generation multimedia service-creation environment relying on an IP network, enabling operators to deliver advanced services and to improve the service quality guaranteed. It provides a session-control layer between IP terminals/transport networks and the applications/services. Since full IMS deployments require an all-IP network, some new operators have built an all-IP network from the outset. However, most wire-line and wireless carriers have a large collection of legacy network elements and protocols, such as, for example, CS voice systems, special-purpose data service platforms, and pre-IMS service delivery systems, which they cannot afford discarding. To provide an adequate return on their investment in full-scale IMS deployments, an attractive range of successful new services is needed. In order to avoid a high risk, an evolutionary migration relying on replacing the systems that became outdated is preferred, rather than opting for their radical replacement.

IMS promises important advantages in terms of attractive cost reductions and reduced time to market for radically new services. In general, VoIP-based services are being developed first in the enterprise market. Of the many possible services that IMS could enable, some are receiving particular attention as short-term revenue-earning opportunities, including PTT, video sharing and IM, just to name a few.

The first services deployed in the consumer market are PoC-based services, which offer half-duplex voice

communications, where a single party can speak at a time by pushing a button on the mobile device, which is then received by many mobiles. This service is usually reserved for short bursts of voice messages, typically lasting up to 30 s. This basic service can also be enhanced with the aid of presence information, group-call capability, and IM. While PTT services can potentially be delivered using the legacy GPRS, they may suffer from the limited capacity of the entire system as well as from undesirable latency. On the other hand, traditional PTT services use simple terminals, endowed with a dedicated button to activate the PTT mode. Standard cellular handsets do not have this facility. A more detailed discourse will follow in the detailed subsection dedicated to the PoC service.

Video sharing is another IMS service that has been gaining popularity, which allows mobile users to establish a regular CS call and complement it, at a later time, with a stream of video originated at one of the user's end. Hence, the pair of users having an audio call also benefit from having an additional single video stream, where both speakers can watch the same video and may curtail video sharing without ending the voice call. The introduction of video sharing during a voice call may be a key evolutionary step toward fully fledged video services.

The IMS enables mobile operators to introduce successful new wireless IM services, such as chat rooms and group messaging, which have become so successful for existing fixed-line-based networked PCs. This creates non-real-time user-to-user services capable of extending the classic mobile text messaging to allow three or more users to engage in shared messaging sessions. The basic services can be enhanced with the aid of presence information and can be integrated with other applications, such as mobile gaming. However, previous attempts to deliver mobile IM or group messaging services have failed because of terminal constraints, owing to their difficulty of use and lack of interworking between networks.

Finally, moving to the enterprise segment, the most prominent application is IP Centrex, which is an enhanced IP-based form of the traditional circuit-based Centrex, whereby a company's internal telecommunications are hosted by a fixed operator. This provides a cost-effective alternative to an inhouse PBX supervised by internal staff. The IMS potentially enables fixed operators to host a complete set of personal and group services, with the added benefit of advanced features such as video communication, conferencing, collaboration, presence management, IM, and e-mail servers' integration.

1) *Commercial IMS Services*: The IMS infrastructure is being rolled out at the time of writing by major operators, such as Telefonica in Spain, extending progressively to other countries such as Germany [77], the Czech Republic, Taiwan, Mexico, Colombia, and Argentina. Telefonica has announced that it will be deployed in all 20 countries where it operates by 2013. As anticipated, LTE deploy-

ments will foster IMS deployments, which are likely to become mainstream in the near future. Commercial services based on IMS, especially in the context of PoC and video sharing have already found favor by users.

PoC services became commercially available as early as 2003 in the United States [60] and a few months later in Europe. Initially, they were limited to customers equipped with PoC-enabled phones, who were able to talk to anyone communicating to people on the same carrier, but not mapped to other carriers. Indeed, PoC subscribers may initiate a PoC session with any landline or wireless-phone user; in other words, they are not limited to PoC sessions addressed to PoC subscribers. Naturally, the subscribers have to use a PoC-enabled wireless handset operating on the same wireless network in order to connect with multiple users, send a voice message to anyone's e-mail—worldwide and in an instant—as well as pictures, even while being on a call. Chung-Hwa Telecom in Taiwan was the first operator launching this service in 2006 using 2.5G technology. Major subscribers include business corporations and government organizations such as the National Security Bureau. Surprisingly, some users who subscribed to PoC services in Taiwan predominantly operate in a one-to-one mode, not in a one-to-many mode. One of the key issues to guarantee the successful spreading of PoC services is the provision of handsets, especially when people replace their handsets frequently. In its plan to launch 3G PoC, NTT Docomo requested the Japanese manufacturers to include PoC as a basic feature of new handset models.

Video sharing as a commercial service has been on offer in Asia since 2005, which enables users to share live video and stored content through an IP connection in real time, while also participating in an ongoing CS call.

C. PoC

Again, the PoC service provides a walkie-talkie-like service within the cellular communication infrastructure [61] with the aid of prompt and simple call setup for one-to-one and one-to-many group calls using the same mobile devices as traditional voice calls. In this service, several predefined PoC group members participate in the PoC sessions. Since the PoC session is half-duplex, only one group member speaks at a time, while the others listen. PoC also supports features such as the provision of presence information, the support of dynamic groups, do-not-disturb indication, and so on. The early PoC protocols were proprietary, but at the time of writing the vendor-independent OMA PoC is supported by all major handset vendors.

The PoC architecture is illustrated in Fig. 7, where the SIP is utilized for supporting the PoC service, where a PoC group includes a predefined set of members, and the SIP universal resource identifier of each PoC group member is maintained in the associated group member list. The PoC group is identified by a Telephone URI, such as for example, tel: +88635131350) or a SIP URI, such as

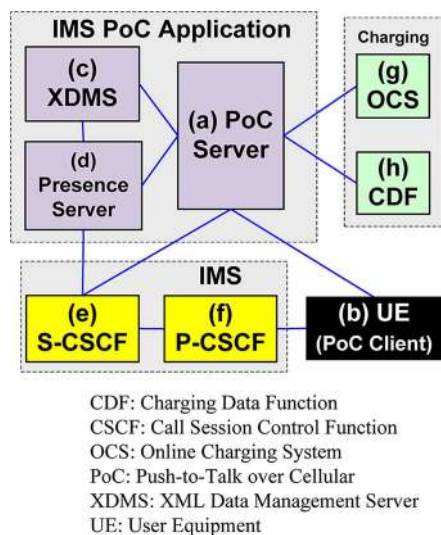


Fig. 7. IMS PoC network architecture.

sip:PoCGroup1@pcs1.csie.nctu.edu.tw. The PoC server of Fig. 7(a) includes both signaling and media processing functions. The signaling is responsible for the provision of the subscriber service and management, while the media processing subfunction is responsible for session setup, floor control, and audio distribution.

The PoC server stores the PoC subscription information, including the access control rules for a user and manages the PoC services related, for example, to the creation and closure of user groups. It also interacts with the CDF block of Fig. 7(h) for the sake of offline charging and with the OCS block of Fig. 7(g) for online charging, as described in Section II-C3. Apart from its basic features, the OMA PoC specifications also define further features such as prearranged and chat group sessions, group advertisements, incoming session and instant personal alert barring, conference state notifications, and the support of user anonymity.

Within a PoC session, a user has to get permission to speak before his/her message is reproduced for the rest of the participants in the PoC session. In other words, PoC activates the floor control mechanism in order to allow only a single user talk at any given moment in a specific PoC session. Typically, a user presses a button in his PoC UE to request the floor. If the PoC server grants the floor to this particular user, the user is notified with an audible signal. Floor control is realized with the aid of the TBCP, which is not a new protocol, but rather a set of extensions to the RTCP. The aforementioned TBCP contains messages conceived for requesting, granting, denying, and revoking a permission to speak, which effectively implies the denial of the floor.

The PoC client of Fig. 7(b) uses the SIP relying on the P-CSCF and on the S-CSCF of Fig. 7(e) and (f) to transmit the session management requests to the PoC server [see

Fig. 7(a)], although the floor control requests are provided with the aid of the media plane using the TBCP/RTCP. Once the PoC session was established, each PoC group member creates a unicast RTP session, which links it to the PoC sever. If a PoC group member obtains the floor, his/her voice is sent to the PoC server with the aid of the RTP session. The PoC server then forwards the voice packets to each of the group members directly through the GGSN or packet data gateway in WiFi/WiMAX. In other words, the RTP packets do not have to be delivered through the MGW in IMS.

When a PoC client joins the PoC service or when the PoC server handles a call invitation, they obtain the relevant group-formation information from the XDMS of Fig. 7(c), which is stored in the group member list. The XDMS is responsible for the PoC list and group management, which is achieved by accessing and manipulating the XML-format PoC information stored in networked document repositories. The IETF XCAP allows a PoC client or the PoC server to read, write, and modify the relevant application configuration data stored in the XDMS [62].

The presence server of Fig. 7(d) is a SIP-based IMS application server that collects, stores, and distributes the status of the PoC clients, such as their mood, activity, location, and service capabilities [63], [64]. The related information may originate from several different devices, such as the PoC clients and other application servers. The presence server merges all the information gleaned and forms a complete view of each PoC client's presence information. Then, it sends the combined presence data to all watchers who have subscribed to the presence of the particular PoC client.

In the PoC implementation, the efficiency of the talk burst control procedure significantly affects the overall performance of the service [65]. The number of clients involved in a PoC session also has an effect on the performance of IMS. In Chunghwa Telecom's PoC service based on GPRS, the maximum number of clients in a PoC session is 20. The cost of PoC calls is about 30%–50% lower than that of the regular cellular phone calls. Like Chunghwa Telecom, NTT DoCoMo also supports up to 20 members in a PoC session. Based on 3G IMS, more than 26 handset models support the IMS PoC software and have a dedicated hard-key for facilitating DoCoMo's PoC service. The GPRS-based PoC service has not been successful due to the relatively long PoC call setup time of 13 s and the high handoff time of 2.5–3 s required for reconnecting a PoC client, when it moves from one base station to another. This problem has been resolved by the IMS-based PoC services established on the 3G networks. For example, the 3G PoC call setup time is less than 6 s. Another key factor of providing a successful PoC service is that the vendors should support fully compliant OMA PoC handsets. OMA has specified standard testing procedures for PoC and other OMA services. An example of the OMA PoC test can be found in [66].

D. The Rich Communications Suite

Contemporary mobile users utilize a range of potentially independent/disparate services, which may result in a disconnected communications experience. The RCS [37] aims for circumventing this problem by providing enhanced integrated services that can be customized. The ability to exploit the core components of the RCS for enhancing other applications and services is compelling. The RCS bundles a range of applications as well as services and, hence, it is capable of filling the gap among their services with the aid of a common, well-defined set of features of the standardization bodies, such as 3GPP (in general ETSI) and OMA. The corresponding standards cover a broad range of services and with great latitude in what functions may be included from each of them. The ultimate goal is to provide interoperable, enriched communications, benefitting all the end users with a new range of exciting, feature-rich services, including the provision of enhanced address book, enhanced messaging, content sharing, and file transfer services. The applications and services in the RCS are being developed as a result of a collaborative effort of about 60 industry players working together in the GSMA's RCS project. Various RCS implementations have been showcased at the GSMA's Mobile World Congress as early as February 2009. The project seeks to speed the development, testing, and introduction of commercial IMS-rich communication services and to develop a stable framework for the definition of business models, interworking guidelines as well as technical reference implementations. The focus is upon enriching the user experience with the aid of advanced features, but these efforts do not aim for improving the user interface or device implementation. Instead, the implementation guidelines for aiming for an improved user experience include key issues such as service discovery, the continued support of traditional contacts already stored in the phonebook, and seamless handset configuration.

The RCS releases (release 3 published in December 2009, together with updates to previous releases) entail a number of work items for the sake of ensuring that vendors develop standard, interoperable implementations, which initially targeted mobile phones. Later they were extended to other endpoints, such as PCs and fixed devices, as well as to future client software development kits relying on key RCS enabler building blocks. The RCS releases include the following.

- **Enhanced address book:** It provides presence information as an evolved "OMA SIMPLE Presence" feature and communications capability indications, which enable the users to initiate communications sessions, including voice calls, video calls, file transfers, or messaging. This feature also allows users to integrate a range of multimedia sources, such as the photos of contacts. Furthermore, it allows the selection of the most suitable commu-

nication mode, which may be selected from the phonebook by choosing a scenario-specific communication type and/or medium.

- **Content sharing:** It enables users to exchange different types of content, such as video or photos during a call. The specific forms of multimedia sharing available at a given time for the communicating parties are explicitly presented to the call participants for their decision, in order to circumvent potentially unpleasant situations occurring, when one of the communicating parties does not want to or is unable to opt for a specific communications medium. This may be exemplified by the unavailability of 3G coverage, or by subscribers who are not equipped with a screen, do not want to pay for a higher tariff, etc.
- **Rich messaging:** It expands the traditional IM by unifying multiple media and, hence, provides an enhanced user experience. This is achieved by facilitating the reception and initiation of all types of communications, including classic voice calls, SMS, MMS, instant messaging, while allowing the user to view his/her communications history in a manner, which is reminiscent of chat history in IM.
- **File transfer** allows the users to exchange different types of files during an ongoing multimedia session or even without having an ongoing session.

RCS also includes several other services such as the feature to centrally backup and restore the contacts in the local phonebook to a safe network repository, hence minimizing the impact of lost, broken, or stolen devices. It is also possible to access network-based directories, such as white/yellow pages, intranet directories, service specific directories, and to search for as well as to retrieve contacts from the local phonebook for storing in the mobile device.

Existing services, including video, voice, MMS, and SMS, as well as a variety of future services can be integrated into the RCS. In addition, as new services are added or when users upgrade their devices, the new service capabilities can be broadcast to contacts via the presence service.

The latest RCS specification (v5.1) has been published in August 2012 (v4.0 was released in December 2010 including Chat Interworking with legacy SMS/MMS, extension of RCS to full fixed access and LTE). An enhanced version of RCS, known as RCS-e (which is a simple interoperable extension of the voice and text services), was published in 2011 (v1.1 April 2011).

Commercial deployments have been launched in 2012 under the *Joyn* brand, starting with Spain (Vodafone, Telefonica; Orange announced) and Germany (Vodafone; Deutsche Telekom and Telefonica announced), guaranteeing cross-operator interoperability. Further near-future markets include Italy, France, and South Korea.

E. VoLTE

The previously mentioned VoLTE has been launched commercially in 2012 in South Korea (nationwide by *SK Telecom*, *LGU Plus*) as well as in the United States (*MetroPCS* in one city), all three relying on Ericsson equipment delivered by CDMA operators.

VoLTE enables the decommissioning of legacy (2G/3G) networks (GSM/UMTS, CDMA), allowing to eventually reform their spectrum for use in LTE, given the fact that legacy voice networks may be phased out. Initial LTE deployments for smartphones have used the so-called CSFB, which means that voice is carried over legacy networks, which has to be maintained in parallel when full LTE coverage becomes a reality; it will be no longer economical to maintain two or even three networks. Moreover, the legacy networks occupy valuable spectrum that could be reassigned to LTE which is more spectrum efficient.

F. Potential Applications Based on IMS

Personal communications relying on voice, video, and messaging as well as personal mobility—including device-independence—are intrinsic features of IMS. As a benefit of the enabling techniques discussed earlier (presence, groups, location, networked address book, converged messaging, user profile, PTT, pushing messages, etc.), the introduction of sophisticated communications services is considerably simplified. This is a very strong advantage of the IMS ecosystem, conveniently fueling service innovation.

Indeed, innovating beyond the realms of these basic technologies offers new opportunities for the application sectors. In this section, we may speculate a little further with the aid of a few examples. In these intelligent applications, specific servers are used for integrating the services with the public IMS infrastructure, while the associated mobility management is capable of supporting both user and device mobility. Below we consider a few specific applications, without being exhaustive.

Digital residential services: IMS may be employed in the home as a multimedia gateway between the external network and the internal devices for content exchange, including multimedia streaming, downloading, storing, remote rendering, registration notification, etc. The IMS also holds the promise of seamless communications and interworking with home-based technologies, such as those conceived by the DLNA as well as with residential femtocells with the new generation of digital enhanced cordless telephones.

- Corporate applications: Historically it has been the corporate sector, which was the advocate of VoIP technologies—for example, in the context of PBX applications. The fixed-mobile convergence capability is also quite relevant in the corporate sector, as is the integration of telephony and PCs, leading to a whole host of potent computer-supported telecommunications applications. Given the wide availability corporate PC-based unified commu-

nications endpoints, the integration with the IMS public infrastructure may render the corporate networks obsolete.

- More specifically, thinking about vertical industries, we have the following.
 - Health: The healthcare and telemedicine sector is a prominent example of new applications, facilitating a radically new approach to the provision of remote healthcare. To elaborate a little further, high-definition stereoscopic or holographic videoconferencing among doctors and their patients may avoid unnecessary hospitalization or may assist junior doctors during their practical training, etc.
 - Automotive: IMS enabled multimedia services are attractive in vehicular environments for commuters, in haulage, in maritime and other transport scenarios.

Last but not least, it is worth mentioning machine-to-machine (M2M) communications, which is a promising emerging business for operators, that can be applied in different industries including the smart metering in the connected car, etc. There is consensus among analysts that, in the midterm, the number of mobile devices connected to cellular networks will surpass the number of users. It is natural, therefore, to consider integrating M2M services with the IMS for inspiring new innovative applications. Further examples include video surveillance for security, remote operation of devices enhanced with the aid of video, eHealth applications, etc.

IV. TECHNICAL CHALLENGES, OPEN ISSUES, AND FUTURE RESEARCH

Despite its relative maturity, numerous issues remain to be solved in IMS, which delay its wide-ranging applications and limit the number of full IMS deployments.

Operators recognize the need to migrate toward next-generation IP-based networks, but they prudently try to avoid considerable investments required. One of the impediments delaying a wide-scale IMS deployment in wireless networks has been the relatively modest performance of the existing 2G radio access networks, especially in the context of real-time communications. The 3G radio networks improved the attainable real-time performance, but they still might fall short of the required quality. By contrast, numerous IMS deployments have been introduced in the area of fixed networks, where there is a less stringent limit on the available bandwidth, apart from residential WLAN access points. It is expected that the emerging fourth-generation (4G) technologies, such as the LTE system, will finally eliminate this problem, since it has a substantially increased overall system bandwidth and minimum latency, hence, it may be expected to support major strategic initiatives by both mobile and fixed-line service providers and operators.

The related research activities are intensive at the time of writing [67]–[71]. The creation of an all-IP IMS-centric network constitutes a major effort in the context of next-generation telecommunications research and development. With the advent of IMS, mobile operators expect radical service-related innovations to generate extra revenues. Apart from PoC and location/presence services, the mobile operators are targeting innovative IMS-based services, which are expected to support communications, access to information, and entertainment in the home, such as IPTV, triple play, and so on. The support of emergency calls is also of high significance [72], [87]. Furthermore, the mobile operators are in the process of providing multiaccess extensions for IMS by integrating existing wireless technologies. Therefore, the feasibility of voice call continuity across standards has been intensively investigated [73].

Online charging: At the time of writing the “consumption” of network resources varies considerably among subscribers, who are generally charged the same flat rate for Internet access. However, mobile operators are unable to guarantee the same QoS for all subscribers. Therefore, it is important that a mobile operator has the ability to provide different, but quantifiable QoS guarantees for subscribers who are charged different rates and the charging mechanism should be sufficiently prompt and efficient to support near-real-time or online charging. It should also be mentioned that the application of the OCS protocols may generate significant IMS network traffic, hence further improved mechanisms are required for supporting real-time charging. Several research issues regarding the associated IMS charging overhead can be found in [74] and [75].

Performance Optimization: When employing IMS application servers, selecting the appropriate parameters for optimizing the related operations is essential. For example, the employment of different talk-burst control policies will affect the attainable performance of PoC [65]. As another example, the appropriate adjustment of the presence information update and notification rate has to be based on a tradeoff between the IMS network traffic and the accuracy of the user-presence information in the presence service [52].

REFERENCES

- [1] M. Maresca, N. Zingirian, and P. Baglietto, “Internet protocol support for telephony,” *Proc. IEEE*, vol. 92, no. 9, pp. 1463–1477, Sep. 2004.
- [2] B. Goode, P. Vidales, and H. Einsiedler, “Voice over Internet protocol (VoIP),” *Proc. IEEE*, vol. 90, no. 9, pp. 1495–1517, Sep. 2002.
- [3] A. Cuevas, J. I. Moreno, P. Vidales, and H. Einsiedler, “The IMS service platform: A solution for next-generation network operators to be more than bit pipes,” *IEEE Commun. Mag.*, vol. 44, no. 8, pp. 75–81, Aug. 2006.
- [4] S. Zaghoul, W. Bziuk, and A. Jukan, “Signaling and handoff rates at the policy control function (PCF) in IP multimedia subsystem (IMS),” *IEEE Internet Comput.*, vol. 12, no. 7, pp. 526–528, Jul. 2008.
- [5] M. A. Melnyk, A. Jukan, and C. D. Polychronopoulos, “A cross-layer analysis of session setup delay in IP multimedia subsystem (IMS) with EV-DO wireless transmission,” *IEEE Trans. Multimedia*, vol. 9, no. 4, pp. 869–881, Jun. 2007.
- [6] G. Camarillo and M. A. García-Martín, *The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds*, 3rd ed. London, U.K.: Wiley, 2008.
- [7] M. Toy, H. Stuttgart, and M. Ulema, “IP multimedia systems infrastructure and services,” *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 112–113, Mar. 2007.
- [8] D. Vingarzan, P. Weik, and F. Fokus, “IMS signaling over current wireless networks: Experiments using the open IMS core,” *IEEE Veh. Technol. Mag.*, vol. 2, no. 1, pp. 28–34, Mar. 2007.
- [9] A. Udugama, K. Kuladinithi, C. Gorg, F. Pittmann, and L. Tionardi, “NetCAPE: Enabling seamless IMS service delivery across heterogeneous mobile networks,” *IEEE Commun. Mag.*, vol. 45, no. 7, pp. 84–91, Jul. 2007.
- [10] K. Munasinghe and A. Jamalipour, “Interworking of WLAN-UMTS networks: An IMS-based platform for session mobility,” *IEEE Commun. Mag.*, vol. 46, no. 9, pp. 184–191, Sep. 2008.
- [11] X. Fangmin, Z. Luyong, and Z. Zheng, “Interworking of Wimax and 3GPP networks

V. CONCLUSION

An overview of the IMS infrastructure, services, and applications has been provided. Based on the all-IP concept the underlying call control protocol (SIP), the associated IMS architecture and the related services enablers were discussed.

Albeit the concept of VoIP dates back to the mid-1990s, the IMS technology relaying on carrier-grade VoIP has only appeared during the late 2000s and has been evolving ever since. We have seen that there are still some technical challenges, but all in all the IMS technology is deemed mature enough to become the mainstream multimedia communications enabler.

Nonetheless, sceptics argue that there are insufficient radically new applications that can be offered exclusively by IMS. Another camp of sceptics claims that having a fully centralized infrastructure following the traditional operator philosophy has a limited future. Nonetheless, at the time of writing there is a reliable telecommunications infrastructure in place or indeed, being rolled out in form of 2G mobile telephony in emerging markets accounting for billions of users. The question is: How long can the legacy systems be exploited? Furthermore, how much investment is required to create the all-IP infrastructure (even for a progressive migration)?

On the other hand, there are a number of important trends, which suggest a bright future for IMS. First, the FITH concept suggests that the classic copper wire is likely to disappear after a transitory period and, with it, the local loop designed for traditional voice telephony is likely to become antiquated, making VoIP imperative. Similarly, the recently standardized 4Gwireless system is all-IP based, eliminating the circuit switching concept in the radio access. Naturally, it will take time until both the FTTH and 4G technologies will reach predominant penetration, but it is only a matter of time. Moreover, the OPEX may be expected to reduce owing to the convergence, since a single network operator will be able to integrate both fixed and mobile telephony into the IP network. It is also anticipated that IMS will accelerate the conception of new services and applications—all in all, this is an exciting era for communications engineering. ■

- based on IMS IP multimedia systems (IMS) infrastructure and services," *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 144–150, Mar. 2007.
- [12] S. Cherian, P. Feder, B. Sadeghi, and R. Wisenocker, "Integration of the IMS/PCC framework into the mobile WiMAX network," *IEEE Commun. Mag.*, vol. 46, no. 10, pp. 66–73, Oct. 2008.
- [13] F. G. Marquez, M. G. Rodriguez, T. R. Valladares, and L. A. Galindo, "Interworking of IP multimedia core networks between 3GPP and WLAN," *IEEE Wireless Commun.*, vol. 12, no. 3, pp. 58–65, Jun. 2005.
- [14] H. Cruickshank, Z. Sun, A. Sanchez, and F. Carducci, "Analysis of IP voice conferencing over EuroSkyWay satellite systems," *Inst. Electr. Eng.—Commun.*, vol. 148, pp. 202–206, Aug. 2001.
- [15] Z. Sun, D. He, H. Cruickshank, L. Liang, A. Sanchez, C. Miguel, V. Schena, C. Tocci, and B. Carro, "Scalable architecture and evaluation for multiparty conferencing over satellite links," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 3, pp. 594–605, Apr. 2004.
- [16] G. Camarillo, T. Kauppinen, M. Kuparinen, and I. M. Ivars, "Towards an innovation oriented IP multimedia subsystem IP multimedia systems (IMS) infrastructure and services," *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 130–136, Mar. 2007.
- [17] H. Montes, G. Gomez, R. Cuny, and J. F. Paris, "Deployment of IP multimedia streaming services in third-generation mobile networks," *IEEE Wireless Commun.*, vol. 9, no. 5, pp. 84–92, Oct. 2002.
- [18] K. D. Wong and V. K. Varma, "Supporting real-time IP multimedia services in UMTS," *IEEE Commun. Mag.*, vol. 41, no. 11, pp. 148–155, Nov. 2003.
- [19] A. Buono, S. Loreto, L. Miniero, and S. P. Romano, "A distributed IMS enabled conferencing architecture on top of a standard centralized conferencing framework IP multimedia systems (IMS) infrastructure and services," *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 152–159, Mar. 2007.
- [20] S. Q. Khan, R. Gaglianella, and M. Luna, "Experiences with blending HTTP, RTP, and IMS IP multimedia systems (IMS) infrastructure and services," *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 122–128, Mar. 2007.
- [21] E. Mikocz, 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.
- [22] R. Levenshteyn and I. Fikouras, "Mobile services interworking for IMS and XML webservices," *IEEE Commun. Mag.*, vol. 44, no. 9, pp. 80–87, Sep. 2006.
- [23] H. Khelifi and J. C. Gregoire, "IMS application servers: Roles, requirements, and implementation technologies," *IEEE Internet Comput.*, vol. 12, no. 3, pp. 40–51, May 2008.
- [24] C. Gourraud, "Using IMS as a service framework," *IEEE Veh. Technol. Mag.*, vol. 2, no. 1, pp. 4–11, Mar. 2007.
- [25] J. O'Connell, "Service delivery within an IMS environment," *IEEE Veh. Technol. Mag.*, vol. 2, no. 1, pp. 12–19, Mar. 2007.
- [26] M. Gomez and T. P. de Miguel, "Advanced IMS multipoint conference management using web services," *IEEE Commun. Mag.*, vol. 45, no. 7, pp. 51–57, Jul. 2007.
- [27] I. Mas, V. Berggren, R. Jana, J. Murray, and C. W. Rice, "IMS-TV: An IMS-based architecture for interactive, personalized IPTV," *IEEE Commun. Mag.*, vol. 46, no. 11, pp. 156–163, Nov. 2008.
- [28] D. Tolle, A. Plankl, and G. Butscher, "IMS-controlled ethernet domains for networked services delivery," *IEEE Commun. Mag.*, vol. 46, no. 5, pp. 162–171, May 2008.
- [29] V. K. Varma, T. Magedanz, and K. C. Chua, "IMS as service delivery platform for converged networks: Architecture, protocols, and applications," *IEEE Technol. Mag.*, vol. 2, no. 1, pp. 2–55, Mar. 2007.
- [30] Y. B. Lin and M. H. Tsai, "Caching in I-CSCF of IP multimedia subsystem," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 186–192, Jan. 2006.
- [31] Y.-B. Lin, Y.-C. Sung, H. F. Rashvand, C.-L. Liu, and Y.-J. Liao, "Mobile-Taiwan experience in voice over IP-worldwide interoperability for microwave access trial," *IET Proc. Commun.*, vol. 4, no. 9, pp. 1130–1141, 2010.
- [32] A. Sánchez, C. Baladrón, J. Aguiar, B. Carro, L.-W. Goix, J. Sienel, R. Trapero, and J. L. Cantanero, "User-centric service creation and execution," *At Your Service: An Overview of Results of Projects in the Field of Service Engineering of the IST Programme*, Cambridge, MA, USA: MIT Press, 2007.
- [33] P. Bellavista, A. Corradi, and L. Foschini, "IMS-based presence service with enhanced scalability and guaranteed QoS for inter-domain enterprise mobility," *IEEE Wireless Commun. Mag.*, vol. 16, *Special Issue on Enterprise Mobility Services*, no. 3, pp. 16–23, Jun. 2009.
- [34] C. J. Pavlovski, "Service delivery platforms in practice," *IEEE Commun. Mag.*, vol. 45, no. 3, pp. 114–121, Mar. 2007.
- [35] D. Geerl, "Building converged networks with IMS technology," *Computer*, vol. 38, pp. 14–16, Nov. 2005.
- [36] A. Salkintzis, M. Hammer, I. Tanaka, and C. Wong, "Voice call handover mechanisms in next-generation 3GPP systems," *IEEE Commun. Mag.*, vol. 47, no. 2, pp. 46–56, Feb. 2009.
- [37] GSM Association, "GSM rich communication suite release 1 specification," Tech. Specification, 2008.
- [38] Y.-B. Lin, M. F. Chang, M. T. Hsu, and L. Y. Wu, "One-pass GPRS and IMS authentication procedure for UMTS," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 6, pp. 1233–1239, Jun. 2005.
- [39] L. S. Kapov, M. Mosmondor, E. N. Tesla, O. Dobrijevic, and M. Matijasevic, "Application-level QoS negotiation and signaling for advanced multimedia services in the IMS," *IEEE Commun. Mag.*, vol. 45, no. 7, pp. 108–116, Jul. 2007.
- [40] W. Zhuang, S. G. Yung, J. L. Kok, and C. C. Kee, "Policy-based QoS architecture in the IP multimedia subsystem of UMTS," *IEEE Commun. Mag.*, vol. 17, no. 5, pp. 51–57, May 2005.
- [41] S. R. Yang and W. T. Chen, "SIP multicast-based mobile quality-of-service support over heterogeneous IP multimedia subsystems," *IEEE Trans. Mobile Comput.*, vol. 7, no. 11, pp. 1297–1310, Nov. 2008.
- [42] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, *SIP: Session initiation protocol*, IETF RFC 3261, 2002.
- [43] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," IETF RFC 3550, 2003.
- [44] B. Campbell, R. Mahy, and C. Jennings, "The message session relay protocol (MSRP)," IETF RFC 4975, 2007.
- [45] J. Rosenberg and H. Schulzrinne, "An offer/answer model with the session description protocol (SDP)," IETF RFC 3264, 2002.
- [46] M. Handley, V. Jacobson, and C. Perkins, "SDP: Session description protocol," IETF RFC 4566, 2006.
- [47] W. E. Chen, Y.-B. Lin, and A. C. Pang, "An IPv4-IPv6 translation mechanism for SIP overlay network in UMTS All-IP environment," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 11, pp. 2152–2160, Nov. 2005.
- [48] 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.
- [49] B. Chatras, "Business communications standardization in ETSI," *IEEE Wireless Commun. Mag.*, vol. 16, no. 3, pp. 8–14, Jun. 2009.
- [50] Third Generation Partnership Project (3GPP), "Technical specification group services and system aspects; presence service; architecture and functional description," Tech. Specification 3G TS 23.141 version 8.1.0, 2008.
- [51] Third Generation Partnership Project (3GPP), "Technical specification group core network and terminals; presence service using the IP multimedia (IM) core network (CN) subsystem; stage 3," Tech. Specification 3G TS 24.141 version 8.1.0, 2008.
- [52] W. E. Chen, Y.-B. Lin, and R. H. Liu, "Weakly consistent scheme for IMS presence service," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3815–3821, Jul. 2009.
- [53] Third Generation Partnership Project (3GPP), "Technical specification group core network; IP multimedia subsystem (IMS); stage 2 (release 5)," Tech. Specification 3G TS 23.228 version 5.15.0, 2006.
- [54] 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.
- [55] Third Generation Partnership Project (3GPP), "Technical specification group service and system aspects; telecommunication management; charging management; charging data record (CDR) file format and transfer (release 6)," Tech. Specification 3G TS 32.297 version 6.2.0, 2006.
- [56] Third Generation Partnership Project (3GPP), "Technical specification group services and system aspects; telecommunication management; charging management; charging architecture and principles (release 6)," Tech. Specification 3G TS 32.240 version 6.4.0, 2006.
- [57] Third Generation Partnership Project (3GPP), "Technical specification group service and system aspects; telecommunication management; charging management; online charging system (OCS): Applications and interfaces (release 6)," Tech. Specification 3G TS 32.296 version 6.3.0, 2006.
- [58] Third Generation Partnership Project (3GPP), "Technical specification group service and system aspects; telecommunication management; charging management; diameter charging applications (release 6)," Tech. Specification 3G TS 32.299 version 6.12.0, 2007.

- [59] Y.-B. Lin and S. I. Sou, *Charging for Mobile All-IP Telecommunications*. Chichester, U.K.: Wiley, 2008.
- [60] A. Sánchez, D. Moro, and L. E. Martín, "PoC as VoIP service in mobile networks," in *Telecommunications and Mobility in the Information Society*, A. Becerra, Ed. Madrid, Spain: AHCJET, 2005, pp. 243–58.
- [61] Third Generation Partnership Project (3GPP), "Technical specification group services and system aspects; 3GPP enablers for open mobile alliance (OMA); push-to-talk over cellular (PoC) services; stage 2 (release 8)," 3GPP TR 23.979 version 8.0.0, 2008.
- [62] J. Rosenberg, "The extensible markup language (XML) configuration access protocol (XCAP)," IETF RFC 4825, 2007.
- [63] H. Sugano, S. Fujimoto, G. Klyne, A. Bateman, W. Carr, and J. Peterson, "Presence information data format (PIDF)," IETF RFC 3863, 2004.
- [64] J. Rosenberg, "A data model for presence," IETF RFC 4479, 2006.
- [65] M. H. Tsai and Y.-B. Lin, "Talk burst control for push-to-talk over cellular," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2612–2618, Jul. 2008.
- [66] Y.-B. Lin, C. F. Liang, K. H. Chen, and H. Y. Liao, "NTP-SIOT: A test tool for advanced mobile services," *IEEE Network*, vol. 21, no. 1, pp. 21–26, Jan./Feb. 2007.
- [67] A. Sanchez, B. Carro, P. Romo, and C. Pinart, "Telecom operator services for end customers within Spanish programmes," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 24–28, Jun. 2008.
- [68] A. Sanchez, B. Carro, and S. Wesner, "Telecom operator services for end customers: European perspective," *IEEE Commun. Mag.*, vol. 46, no. 2, pp. 14–18, Feb. 2008.
- [69] A. S. Esguevillas, B. Carro, B. Khasnabish, and A. Gupta, "Applications and support technologies for mobility and enterprise services," *IEEE Wireless Commun. Mag.*, vol. 16, no. 3, pp. 6–7, Jun. 2009.
- [70] A. Sanchez, B. Carro, B. Wei, and W. Feng, "Digital home services," *IEEE Network*, vol. 23, pp. 6–7, Nov. 2009.
- [71] V. G. Cerf, "On the evolution of Internet technologies," *Proc. IEEE*, vol. 92, no. 9, pp. 1360–1370, Sep. 2004.
- [72] M.-H. Tsai, Y.-B. Lin, and H.-H. Wang, "Active location reporting for emergency call in UMTS IP multimedia subsystem," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5837–5843, Dec. 2009.
- [73] Y.-B. Lin, M.-H. Tsai, H.-W. Dai, and Y.-K. Chen, "Bearer reservation with preemption for voice call continuity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2716–2725, May 2009.
- [74] S. I. Sou, Y.-B. Lin, Q. Wu, and J. Y. Jeng, "Modeling prepaid application server of VoIP and messaging services for UMTS," *IEEE Trans. Veh. Technol.*, vol. 56, no. 3, pp. 1434–1441, May 2007.
- [75] S. I. Sou, Y.-B. Lin, and J. Y. Jeng, "Reducing credit re-authorization cost in UMTS online charging system," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3629–3635, Sep. 2008.
- [76] Third Generation Partnership Project (3GPP), "Technical specification group services and system aspects; end-to-end quality of service (QoS) concept and architecture (release 5)," 3G TS 23.207 ver. 5.10.0, 2005.
- [77] A. Cuevas, W. Nicoll, and K. Schröder, "The challenges of IMS deployment at Telefónica Germany," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 120–127, Aug. 2012.
- [78] C. Baladron, J. M. Aguiar, B. Carro, L. Calavia, A. Cadenas, and A. Sanchez-Esguevillas, "Framework for intelligent service adaptation to user's context in next generation networks," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 18–25, Mar. 2012.
- [79] A. Cadenas, A. Sanchez-Esguevillas, and B. Carro, "Deployment of contextual corporate telco services based on protocol adaptation in the NGN environment," *IEEE Commun. Mag.*, vol. 48, no. 4, pp. 34–40, Apr. 2010.
- [80] J. Governado, C. Baladron, J. M. Aguiar et al., "Management of service sessions in an NGN-SOA execution environment," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 103–109, Aug. 2010.
- [81] C. Baladron, A. Cadenas, J. Aguiar, B. Carro, and A. Sánchez, "Multi-level context management and inference framework for smart telecommunication services," *J. Universal Comput. Sci.*, vol. 16, no. 15, pp. 1973–1991, 2010.
- [82] A. Sanchez-Esguevillas, B. Carro, and V. Poosala, "Future convergent telecommunications services: Creation, context, P2P, QoS and charging," *IEEE Commun. Mag.*, vol. 49, no. 1, pp. 58–59, Jan. 2011.
- [83] C. Baladron, J. M. Aguiar, A. Cadenas, L. Calavia, B. Carro, and A. Sanchez-Esguevillas, "User oriented environment for management of convergent services," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 142–149, Nov. 2012.
- [84] H.-Y. Lee and Y.-B. Lin, "Credit pre-reservation mechanism for UMTS prepaid service," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, pp. 1867–1873, Jun. 2010.
- [85] W.-E. Chen, P.-J. Lin, and Y.-B. Lin, "Real-time VoIP quality measurement for mobile devices," *IEEE Syst. J.*, vol. 5, no. 4, pp. 538–544, Dec. 2010.
- [86] Y. C. Ho, Y.-B. Lin, R. H. Liou, and Y.-K. Tu, "Implementing value added applications in next generation networks," *cFuture Internet*, vol. 2, no. 3, pp. 282–294, 2010.
- [87] Y.-B. Lin, M.-H. Tsai, and Y.-K. Tu, "IMS emergency services: A preliminary study," *IEEE Wireless Commun.*, vol. 18, no. 1, pp. 6–14, Feb. 2011.
- [88] G.-D. Fan, C.-C. Huang, Y.-B. Lin, C.-S. Tang, C.-Y. Twu, and Y.-H. Wen, "Enhanced video phone services for NGN/IMS," *Wireless Commun. Mobile Comput. J.*, vol. 12, no. 7, pp. 580–587, 2012.
- [89] H.-Y. Lee and Y.-B. Lin, "A cache scheme for femtocell reselection," *IEEE Commun. Lett.*, vol. 14, no. 1, pp. 27–29, Jan. 2010.

ABOUT THE AUTHORS

Antonio Sánchez-Esguevillas (Senior Member, IEEE) received the Ph.D. degree (with honors) in the field of QoS for real time multimedia services over IP networks from the University of Valladolid, Valladolid, Spain, in 2004.

He has been managing innovation at Telefonica (both at Telefonica I+D-Services line and at Telefonica Corporation), Madrid, Spain. He has also been Adjunct Professor and Honorary Collaborator at the University of Valladolid, supervising several Ph.D. students. He has coordinated very large (in excess of €100 million) international R&D projects in the field of personal communication services, particularly related to voice over IP (VoIP) and Internet protocol (IP) multimedia subsystem (IMS). He has more than 50 international publications and several patents. His current research interests are in the area of digital services.

Prof. Sánchez-Esguevillas is a member of the Editorial Board of *IEEE COMMUNICATIONS MAGAZINE* among others.



Belén Carro received the Ph.D. degree in the field of broadband access networks from the University of Valladolid, Valladolid, Spain, in 2001.

She is a Professor at the Department of Signal Theory and Communications and Telematics Engineering, University of Valladolid. She is the Director of the Communications Systems and Networks (SRC) laboratory, working as a Technical Researcher and Research Manager in several Europe-wide projects in the areas of service engineering and SOA systems, IP broadband communications, NGN/IMS and voice over IP and quality of service. Recently, she has been Technical Director of the OPUCE (Open Platform for User-centric service Creation and Execution) a €10 million project based on Internet protocol (IP) multimedia subsystem (IMS). She has supervised several Ph.D. students on topics related to personal communications, IMS, and session initiation protocol (SIP). She has extensive research publications experience, as an author, a reviewer, and an editor.



Gonzalo Camarillo received the Ph.D. degree in electrical engineering from Aalto University, Helsinki, Finland, in 2011.

He works for Ericsson Research, Helsinki, Finland, as Principal Researcher in the area of services and software. In the past, he funded and headed the Advanced Multimedia Research Laboratory also at Ericsson Research Finland. His research interests include signaling, multimedia applications, transport protocols, and networking architectures. He is an active participant in the Internet Engineering Task Force (IETF), where he has authored and coauthored several specifications used in the Internet protocol (IP) multimedia subsystem (IMS). In particular, he is a coauthor of the main SIP specification, RFC 3261. He has also authored books, patents, and scientific papers on related topics. He has served on the Internet Architecture Board (IAB) and has chaired a number of IETF working groups. Since 2010, he has been the Director of the Real-time Applications and Infrastructures (RAI) area at the IETF. He is also the IETF liaison manager to the Third-Generation Partnership Project (3GPP). He is the Ericsson representative in the SIP Forum and is a regular speaker at different industry conferences. He is a coauthor of the reference book for IMS (bestseller): *The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds* (New York, NY, USA: Wiley, three editions: 2004, 2006, and 2008).



Miguel A. García-Martín received the B.Eng. degree in telecommunications engineering from Universidad de Valladolid, Valladolid, Spain, in 1992.

He is an Expert Systems Engineer in the Product Development Unit User Databases, Ericsson, Madrid, Spain. In the past, he was a Senior Standardization Specialist in the Industry Environment unit of Nokia Siemens Networks, Espoo, Finland and a Principal Research Engineer in the Networking Technologies Laboratory of Nokia Research Center, Helsinki, Finland. Before joining Nokia, he held several positions within Ericsson Finland and Ericsson Spain related to the development of the Internet protocol (IP) multimedia subsystem (IMS). He has been an active participant of the Internet Engineering Task Force (IETF), and for a number of years has been a key contributor to the Third-Generation Partnership Project (3GPP). For some time he has also been participating in the specification of the next-generation network (NGN) in the European Telecommunications Standards Institute (ETSI). In the IETF, he has authored and coauthored several specifications related to the IMS. In 3GPP, he has been a key contributor to the development of the IMS standard. He is also a regular speaker at different industry conferences. He is a coauthor of the reference book for IMS (bestseller): *The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds* (New York, NY, USA: Wiley, three editions: 2004, 2006 and 2008).



Yi-Bing Lin (Fellow, IEEE) received the Ph.D. degree from University of Washington, Seattle, USA in 1990.

He is a Senior Vice President and Lifetime Chair Professor of the National Chiao Tung University (NCTU), Hsinchu, Taiwan. He is the author of the books *Wireless and Mobile Network Architecture* (New York, NY, USA: Wiley, 2001), *Wireless and Mobile All-IP Networks* (New York, NY, USA: Wiley, 2005), and *Charging for Mobile All-IP Telecommunications* (New York, NY, USA: Wiley, 2008).



Prof. Lin is a Fellow of the American Association for the Advancement of Science (AAAS), the Association for Computing Machinery (ACM), and the Institution of Engineering and Technology (IET). He serves on the Editorial Board of the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. He is the General or Program Chair for prestigious conferences, including ACM MobiCom 2002. He is Guest Editor for several journals, including the IEEE TRANSACTIONS ON COMPUTERS. He received numerous research awards, including 2005 NSC Distinguished Researcher, 2006 Academic Award of Ministry of Education, 2008 Award for Outstanding contributions in Science and Technology, Executive Yuen, 2011 National Chair Award, and 2011 TWAS Prize in Engineering Sciences (The Academy of Sciences for the Developing World). He is on the advisory boards or the review boards of various government organizations, including Ministry of Economic Affairs, Ministry of Education, Ministry of Transportation and Communications, and National Science Council. He is a member of Board of Directors, Chunghwa Telecom.

Lajos Hanzo (Fellow, IEEE) received a degree in electronics and the Ph.D. degree from the Technical University of Budapest, Budapest, Hungary, in 1976 and 1983, respectively.

During his 35-year career in telecommunications he has held various research and academic posts in Hungary, Germany, and the United Kingdom. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K., where he holds the Chair in Telecommunications. Currently, he is directing a 100-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the U.K. Engineering and Physical Sciences Research Council (EPSRC), the European IST Programme and the Mobile Virtual Centre of Excellence (VCE), U.K. Since 2009, he has been a Chaired Professor also at Tsinghua University, Beijing, China. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He has successfully supervised 80 Ph.D. students, and coauthored 20 Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, and published 1300+ research entries at IEEE Xplore.



Prof. Hanzo is a Fellow of the Royal Academy of Engineering, the European Association for Signal Processing (EURASIP) and the Institution of Engineering and Technology (IET). In 2009, he was awarded the honorary doctorate “Doctor Honoris Causa” by the Technical University of Budapest. He acted both as TPC and General Chair of IEEE conferences, presented keynote lectures, and has been awarded a number of distinctions. He is also a Governor of the IEEE Vehicle Technology Society. During 2008–2012, he was the Editor-in-Chief of the IEEE Press. In 2012, he was awarded the European Research Council’s Advanced Fellow grant.