Chapter 1

Introduction

Shyam Chakraborty, Tomas Frankkila

The innovation of voice communication with electrical means by Alexander Graham Bell initiated the era of Public Switched Telephone Networks (PSTN), and revolutionized the way people in modern civilization would communicate. The important elements of PSTN are full-duplex conversational service, narrowband speech and Circuit Switching (CS). The success of PSTN provided a strong impetus for seeking improvements and modifications to the basic conversational service to provide a richer experience, better convenience and wider availability to the user, in a variety of ways. The evolution of digital signal processing allowed conversion of analog signals to digital format to take advantage of digital communication and switching techniques. The development of the Integrated Service Digital Network (ISDN) provisioned different services, speech telephony, video telephony and data, over the same interface of a single network. Intelligent networking provided enhancement to basic services and cellular systems added the freedom from desktop telephony. Needless to say, these innovations have also been strong stimulants for industrial growth.

The scope of non-conversational data communications that was catering primarily for telegraphy and facsimile expanded very much with the proliferation of computers. The information generated by computers is highly bursty in nature and communicated either one way (simplex) or interactively (half-duplex). This necessitated a totally new switching paradigm, known as packet switching, which essentially stemmed from a US Department of Defense (DoD) research network ARPANET. Computers proliferated from research laboratories to work places as daily working tools, and ultimately to homes as consumer goods providing information with web (World Wide Web, WWW) browsing, communication with emails and entertainment with a variety of games, music, videos, etc., over a single user interface. This growth is actually aided by the evolution of data networks from the ARPANET to the Internet, a huge collection of heterogeneous networks glued together by the TCP/IP protocol suite, and reaching practically all the offices and homes of a modern society.

Despite the provision of mobility, the first generation analog systems, for example, NMT, TACS and AMPS, met with limited market success, mainly because of limited coverage, bulky terminals and high cost. The potentials of mobile communications were better exploited with the digital second generation Global System for Mobile communication (GSM) that provides circuit switched connections to carry low bit rate coded speech or data. The GSM systems have enjoyed enormous market response due to the highly portable terminals, thanks

IMS Multimedia Telephony over Cellular Systems S. Chakraborty, T. Frankkila, J. Peisa and P. Synnergren

to component miniaturization and improved battery technology, a voice quality that matches closely the quality of PSTN, reduced cost due to scale-factor cost benefits and rapidly expanding coverage. These attributes have transformed cellular mobile systems from business tools to consumer items with a much bigger market potential. An important dimension added to conversational telephony was the introduction of Short Message Service (SMS). Akin to email, the SMS is a much simpler, informal and user friendly service. Though SMS is not a truly conversational service, its high interactive feature is extensively used in near-real-time chatting, and often complements a voice conversation with such additional text information as exchange of address, telephone number, jokes etc. Not only has this enhanced interpersonal communications to a new height, but also it has been a major revenue churner for the operators. With advanced display devices, processors and networking technology, cellular systems are increasingly equipped with data services, Multimedia Message Service (MMS), video telephony and other non-communication facilities, such as camera, MP3 player, etc. An up-to-date mobile terminal is thus a highly sophisticated user equipment with multi-service capabilities.

1.1 Convergence of Networking Paradigms

The concept of an integrated network that provides the user with a variety of services over a single interface has a number of advantages for both the users and the operators. This is well reflected in the developments of both circuit switched narrowband ISDN (N-ISDN) and further packet switched broadband ISDN (B-ISDN) networks. However, a homogeneous, 'end-to-end Asynchronous Transfer Mode (ATM)' concept of a B-ISDN had some obvious difficulties in taking off, and the Internet as a heterogeneous collection of subnetworks has become increasingly dominant. The Internet provides an abundance of bandwidth owing to its core optical fiber network. This factor makes a good case for an integrated network. However, the inherent design principles of the Internet provide only a 'best effort' service that is typically not optimized for conversational services. To provide a variety of services over the Internet, a few improvements in traffic management and handling within the network are beneficial. The first of them is to define Quality of Services (QoS) requirements for the different services. The second is to develop means within the Internet to provide 'better than best-effort' services that would suit the different QoS requirements of different conversational, streaming and interactive classes of services. Examples of such improvements are provisioning of Diffserv (differentiated services) and Multi-Protocol Label Switching (MPLS) that handles the different QoS classes with a statistical guarantee rather than an absolute guarantee. Equipped with these techniques, the Internet can provide the basis of an efficient packet switched voice, commonly known as Voice over IP (VoIP).

The explosion of the Internet into the consumer market has also influenced the cellular systems. While GSM provided data communications over fixed rate circuit switched links at 9.6 kbps, packet services with higher data rate capabilities are increasingly offered. This is reflected in such enhancements of the second generation GSM systems as General Packet Radio Services (GPRS) and Enhanced Data rates for GSM Evolution (EDGE), and the third generation system Universal Mobile Telecommunication Systems (UMTS). These systems offer increasingly higher speed packet access to the Internet, in parallel to the circuit switched links to be provided for narrowband voice telephony. One can therefore truthfully conclude that fixed and mobile networks have converged for data communication.

The fixed–mobile convergence of real-time services has however not been realized so far. Real-time services may have bit rate requirements that are significantly lower than high-speed data services, but the real-time services have the special requirement that the transport of the packets must be constant. For data services, interruptions of for example 100–200 ms or even up to 500 ms are insignificant, even if they occur quite frequently. Interruptions of 100–200 ms, in addition to the delay jitter, can only be allowed for real-time services if the interruptions are very rare. And longer interruptions can never be allowed. The reason is that the receiving client cannot buffer enough frames to survive such interruptions without causing interruptions in the produced sound, because long buffering times increase the response time from the users, which reduces the conversational quality for the users that are involved in the conversation. The alternative, to maintain a short buffering time and accept silence periods in the produced sound, is even worse since this reduces the listening quality even faster. Another fundamental requirement for successful convergence between real-time and data services is the fact that the capacity for the real-time services must be as good for the packet switched system as for the corresponding circuit switched system.

IMS is the system that enables fixed–mobile convergence, as shown in Figure 1.1, and bridges the gap between these two environments. The IMS system will control the session and route the media stream regardless of the access types that are used and regardless of what operators are involved. A key benefit of the IMS system is that it is also backwards compatible since signaling and media gateways will be allocated when at least one of the users is using a legacy circuit switched system.

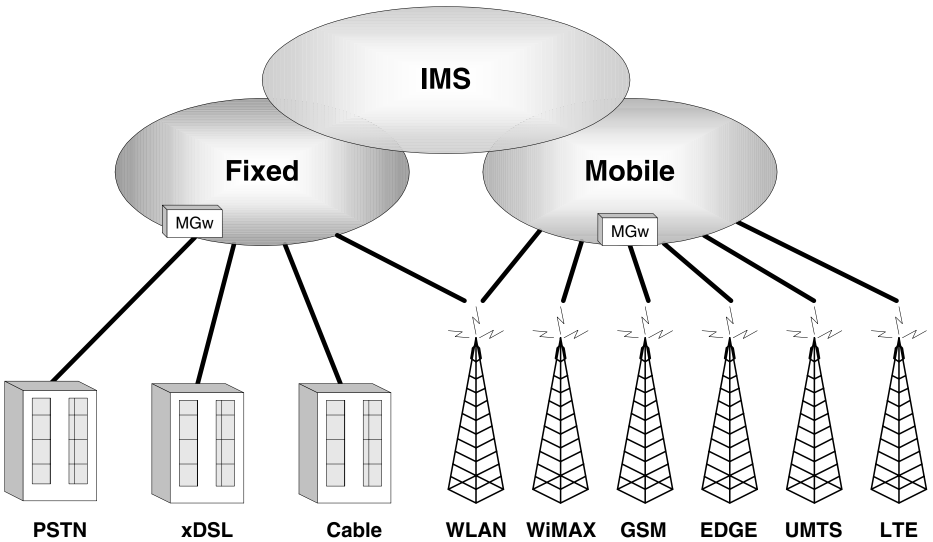

Figure 1.1: Fixed–mobile convergence with IMS as the platform that enables communication regardless of access type and regardless of whether the users have subscriptions with the same or different operators.

1.2 IMS and the IMS Multimedia Telephony Service

The IMS Multimedia Telephony service is seen as the next step in telecommunications and the services which would finally harmonize telephony with data communications. The service will offer enriched communication with real-time speech, video and text communication. In addition, it also offers file sharing and media sharing capabilities that allow users to send, for example, images and video clips to other users. The development of the Multimedia Telephony service is driven by the fact that end-users desire to communicate in new ways, while still requiring a telecom grade service for the traditional real-time voice and video telephony service components, i.e. the same quality, reliability and security.

For enterprise users, Multimedia Telephony will also offer integration of email, support for remote workers, conferencing and collaboration features. Another important property of Multimedia Telephony is personal mobility. Professional users want to use the telephony and data communication services in the same way while traveling as when being in the office. A third important feature for enterprise users is the possibility to control what calls and sessions should be allowed at any given point in time. For example, when attending a business meeting or a conference, the users may want to allow only the most important calls or sessions to reach the receiver. Other, less important, calls may be routed to an answering machine where the sound is recorded and attached in an e-mail to the subscriber's e-mail address.

IMS and the IMS Multimedia Telephony service are interesting also for the operators for several reasons. The IMS system incorporates the control mechanisms that are required in order to ensure that they can meet the users' expectations on quality of the service, reliability and security. The generic architecture and flexibility of the IMS platform also offers simple development and deployment of new services that can be added to the existing services, which gives the operators new revenue opportunities. In addition, the operator also requires an efficiency that is on a par with existing circuit switched systems and also interoperability with legacy systems. IMS also supports developing standardized services, such as IMS Multimedia Telephony. The standardization is required to make the core set of service components behave the same for all operators since a user visiting another country still expects the same service behavior as in his or her home network.

Another attractive property of the IMS is that it allows for operating a single all-IP network, thereby removing the need for maintaining a circuit switched system in parallel with the IP network. This reduces the need for capital expenditure (CAPEX) as well as the operational expenditure (OPEX).

The Multimedia Telephony service is also transport agnostic. This means that service providers only need to implement one version of the service, which significantly reduces the implementation, testing and verification efforts and allows for faster time to market.

1.3 Requirements and Challenges

In all cellular telephony systems up to 3G, voice communication is the service that has defined the toughest requirements on the system. These requirements are related to: capacity, quality, error rates, end-to-end delay and consistent delivered bit rate. Before the introduction of GPRS, voice communication even defined the bit rate requirements.

This changed with the introduction of GPRS, EDGE and data services in 3G. For future systems, it is also the data services that will put the highest requirements on the system regarding error rates (packet loss rates) and end-to-end delay. This is because low packet loss rates and short round-trip times are required if rate control in TCP is to be able to adapt to data rates of several megabits per second. Thereby, it is no longer voice communication that defines the requirements for throughput and delay, but rather data communication.

The requirements that IMS Multimedia Telephony has to fulfill are still challenging. To be able to replace legacy circuit switched services, the capacity and the coverage need to be at least as good. The capacity and coverage requirements must be fulfilled while still delivering the same quality, or better, even when the system is loaded to the capacity limit.

In addition, the system design must also be more flexible than the existing circuit switched systems. This is needed in order to allow for developing and deploying different service variants and to have a future-proof solution that allows for simple introduction of new service components.

During the development of the IMS Multimedia Telephony service, one important property has been that the quality experienced should be consistent. The service should deliver similar performance, in terms of both capacity and quality, regardless of network, operating condition, equipment, etc.

Another requirement is that the IMS Multimedia Telephony service must show consistent behavior for different networks. Equipment from different vendors must also give comparable performance. This is needed because the end-users want the freedom to purchase phones from different vendors and they expect that these will show consistent behavior. Inconsistent performance would also give problems for the operators, if the inconsistencies had an impact on the cell planning, since the operator then might not know if his cell planning gives sufficient coverage or not. This would, for example, be the case if one phone gives adequate performance while another phones gives too much frame erasures, for the same delay jitter.

The IMS Multimedia Telephony service must therefore be standardized in sufficient detail so that the vendors can know that their products will fulfill the performance requirements. This requirement is a real challenge since different transport networks may have quite different characteristics and capabilities.

1.4 Outline of this Book

A prerequisite for IMS Multimedia Telephony to become successful is, as described above, that the basic voice service must be able to replace the traditional circuit switched voice communication in existing cellular systems. To fulfill this requirement, the VoIP in Multimedia Telephony must deliver the same quality as CS voice, while still matching the capacity of the legacy systems. In this book we describe how this requirement can be fulfilled for HSPA systems. We try to cover all areas needed to understand how the IMS Multimedia Telephony service works and how the performance requirements can be fulfilled.

In Chapter 2, the IMS Multimedia Telephony service is discussed in more detail. A few show examples how the service may be used. Requirements are also briefly discussed.

Chapter 3 describes the architecture of the IMS system and the service realization for the Multimedia Telephony. Interworking with legacy system is also an important aspect, since one can expect that the legacy systems will not be replaced overnight. Interconnecting to legacy systems is therefore also briefly described.

Session control is an important part of IMS services and is discussed in Chapter 4. The session setup obviously sets up the session between the users, but it also defines the service components that can be used in the session. One can expect that different users have different subscriptions, which will allow different service components. Session control is also used for routing (finding the other user), setting up radio bearers, QoS, charging, policing, etc.

Chapter 5 describes the media flow between the users. This includes: how the media is encoded, the protocols that are needed for the media, the transmission impairments that may occur in packet switched networks, and how these transmission impairments are managed in the receiving client. This chapter furthermore outlines solutions for interworking with legacy systems. Most of these things are not unique for HSPA or other cellular systems but rather generic for all IP systems. The last section in this chapter therefore presents the special considerations that have been made during the development of the IMS Multimedia Telephony service.

Chapter 6 provides an overview of the security components and mechanisms defined for IMS. The access domain and IMS domain security solutions are discussed, and the outlook of the currently discussed extensions is presented.

In Chapter 7, the performance of the IMS Multimedia Telephony system over HSPA is addressed. The chapter shows that the capacity for voice service is as good, or even better than for the circuit switched service with similar quality requirements. The possibilities to enhance quality are discussed. In addition to the capacity, the coverage for Multimedia Telephony service is addressed, and the network characteristics are presented. Finally the call setup delays are briefly examined.

Chapter 8 discusses services that are related to IMS Multimedia Telephony. These services are: the Circuit Switched IMS Combinational Service (CSICS), Push-to-talk over Cellular (PoC), weShare, Instant Messaging (IM), presence and group list management.

The summary of the book is found in Chapter 9.

The focus of this book is IMS Multimedia Telephony over HSPA, and especially the real-time speech, video and text media included in this service. There is nothing that prohibits using the IMS Multimedia Telephony service on other access methods like EDGE, TISPAN (land-line IP) networks, WLAN or WiMAX, but these access methods are outside the scope of this book.

A few related services, for example PoC, instant messaging and presence, are briefly discussed. Supplementary services like Message Waiting Indication (MWI), Originating Identification Presentation (OIP), Communication Hold (HOLD), etc., are also described.