Multimedia telephony for IMS – Interoperable VoIP with multimedia support

Daniel Enström, Anders Nohlgren, Håkan Olofsson, Janne Peisa and Per Synnergren

ETSI and 3GPP have standardized a new multimedia telephony service. The objective is to give operators the tools with which to build a converged fixed and mobile VoIP service offering with multimedia support. Multimedia telephony combines the quality, interoperability, reliability, efficiency, and regulatory and supplementary services of traditional telephony with the rich, flexible media capabilities of internet community-based services. And because the service is built on IMS, it is access-agnostic – that is, it can run over a variety of access networks.

The authors describe the service in terms of its media-handling capabilities and its set of supplementary services. They also show that WCDMA with HSPA is well suited to transport multimedia telephony.

Introduction

A new, rich communication culture is emerging as a consequence of community interaction on the internet. Communication has evolved into and revolves around the sharing of everyday experiences anywhere and at any time. Operators who choose to ride this wave to maintain their position in the value chain can employ IMS and benefit from rapid and innovative application development.

Global interconnect agreements for services (such as telephony and SMS) between operators constitute a key pillar of the telecommunications industry. Indeed, successful mass-market uptake of new services requires interoperability between operators, networks and devices. Being able to make a telephone call or to send an SMS to anyone, and knowing that the communication will work even across multiple operators and service providers, is a powerful concept that can only be achieved through global standards.

Therefore, in addition to the emergence of IMS as the service engine for next-generation IP networks, the industry has identified a need for detailed standardization of key IMS-based services, in particular for person-to-person communication. Standardizing such services facilitates multi-vendor interoperability and operator interconnect, and thus creates a potential for truly global, mass-market uptake.

Standardized services on IMS

Efforts within the Open Mobile Alliance (OMA) have resulted in standards for messaging and push-to-talk over cellular (PoC) services. These standards, in turn, have been complemented by service enablers, such as presence, group, and list management.

With the aim of organizing the architecture behind IMS and how it handles these services, OMA has introduced *communication services* into 3GPP standards. Accordingly, services use service-specific identities that allow clients and IMS network nodes to identify a communication service and route associated signaling to the correct client software and application server. The concept also allows the application development community to combine communication services in order to build powerful end-user applications.

OMA messaging and PoC are highly capable of handling IP-based services, but neither of them enables full conversational, real-time multimedia communication. As a consequence, ETSI TISPAN took the initiative to standardize what has come to be known as the *multimedia telephony communication service*.

Overview of multimedia telephony

To address a lack of standards for IP telephony communication, ETSI and its TISPAN working group initiated work in 2004 to specify a standard for fixed access. TISPAN endorses IMS as the service engine, and thus refers to 3GPP specifications for session control and basic communication. TISPAN also identified a need to extend standards to include the supplementary telephony-type services that were not part of IMS. These services, referred to as *PSTN/ISDN simulation services*, were specified by year's end 2005.[1]

A parallel activity, initiated in mid-2005, was undertaken to ensure that these same services can be used over 3GPP accesses. The 3GPP maintains its own service definition[2], which refers to applicable TISPAN supplementary services. In addition, to fulfill requirements for 3GPP accesses, the 3GPP standards specify media capabilities and handling.[2-3]

Multimedia telephony, the new global service standard for VoIP with multimedia support, provides a foundation for products that can offer fixed-mobile convergence of services, networks and devices. The service combines the quality, interoperability, reliability, efficiency and renowned supplementary services of traditional telephony, with the rich media and dynamics of internet community-based communication. As such, it can be positioned both as a new service for driving operator revenue and as a replacement for circuit-switched telephony.

TERMS AND ABBREVIATIONS

3GPP	Third Generation Partnership Project
CSICS	Circuit-switched IMS combinational services
ETSI	European Telecommunication Standards Institute
E-UL	Enhanced uplink
HSDPA	High-speed downlink packet access
HSPA	High-speed packet access
IETF	Internet Engineering Task Force
IMS	IP Multimedia Subsystem
IP	Internet protocol
ISDN	Integrated services digital network
MIMO	Multiple input/multiple output
OMA	Open Mobile Alliance
PCRF	Policy and charging rules function
PDP	Packet data protocol
PLMN	Public land mobile network
PoC	Push-to-talk over cellular
PS-CN	Packet-switched core network
PSTN	Public switched telephone network
QoS	Quality of service
RAB	Radio access bearer
RLC	Radio link control
ROHC	Robus header compression
SDP	Session description protocol
SIP	Session initiation protocol
SMS	Short message service
TAS	Telephony application server
VoIP	Voice over IP
WCDMA	Wideband code-division multiple access

Basic communication part – media capabilities

The basic communication part of multimedia telephony allows a single SIP session to control media transfer. Two or more users can communicate in real time using different media components, such as voice, video and text. The basic communication part also supports file transfer and the sharing of video clips, pictures, and audio clips.

IP transport and SIP/SDP session control guarantee service flexibility. In other words, users may add and remove different types of media during a session without having to stop and restart it.

Multimedia telephony is a telephony-grade service, which is to say operators who provide it must meet the following basic service requirements:

- Users need not be required to rely on backup communication solutions.
- The service must provide the audio quality users have come to expect from traditional telephony service.
- Users need not know which operator or kind of phone the called party has.
- Transport and access networks must be built and optimized to handle large volumes of calls in a reliable and efficient manner.

For end-to-end service, several nodes in the transport chain – in particular, terminals – contribute to the fulfillment of service criteria. The real-time requirements of multimedia telephony service put great demands on the media transport processing that takes place almost exclusively in end-user terminals and gateway nodes. The standardization community understands this. The 3GPP, for example, is finalizing specifications that will govern media handling and interaction.[3] The very nature of IP transport introduces new challenges, which if unattended, can degrade performance. Apart from the selection of suitable, highly efficient media codecs for real-time media (which was done in previous 3GPP releases), new additions in Release 7 include media layer adaptation, jitter management, coding, and packetization guidelines for optimizing transport efficiency, PLMN/PSTN interworking, and media-related session negotiation.

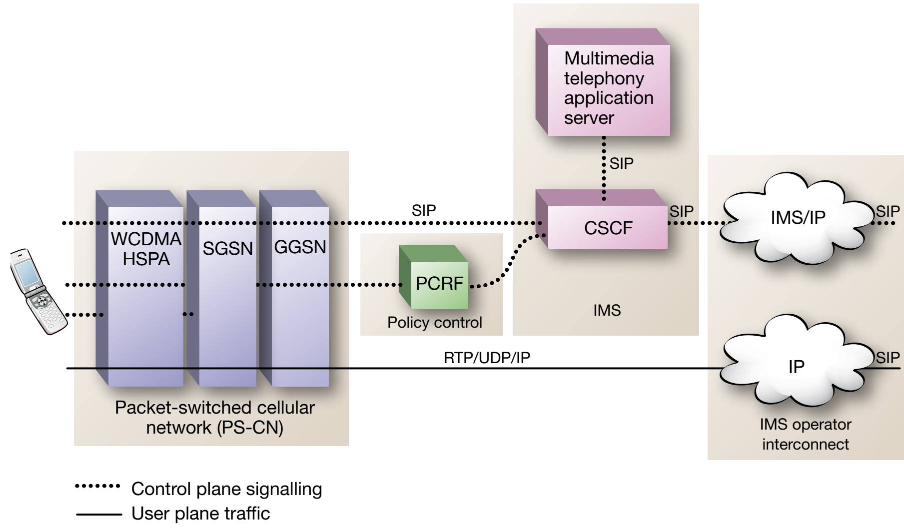

Figure 1
High level network architecture for one of the two IMS+HSPA networks involved in an inter-operator end-to-end Multimedia Telephony call.

Supplementary services

Certainly, attractive and common supplementary services have an essential role in creating a familiar telephony-type user experience. When specifying the supplementary service set, TISPAN considered usefulness, familiarity, and even foreseeable regulatory requirements.

Besides media transfer, users of multimedia telephony expect to be able to forward calls, identify the calling party before answering a call, and so forth. Most of the services are more or less exact replicas of PSTN/ISDN services but put into the context of the entire communication session, including its many active media components. The services have been fully standardized to ensure interoperability of SIP signaling, including operator interconnect.[4]

Interworking aspects

When launched, the multimedia telephony service will have to interwork with existing mass-market services. Successful interworking between terminals that use different services is dependent on

- user plane transpackaging and (sometimes) transcoding;
- translation of session control signaling; and
- translation of signaling for supplementary services.

Interworking between the multimedia telephony voice component and PSTN/ISDN/PLMN circuit-switched voice is crucial. Standardization in this area was initiated in 2005, and today the 3GPP and TISPAN interworking specifications provide a stable framework. Furthermore, circuit-switched video telephony is a widely deployed third-generation mobile service that will be used even more frequently once multimedia telephony has been launched. Therefore, 3GPP is working to standardize interworking between multimedia telephony voice+video and circuit-switched video telephony.

Other services that will gain market presence when multimedia telephony is launched are media-sharing applications based on 3GPP circuit-switched IMS combinational services (CSICS). These applications enable two users to simultaneously maintain a (circuit-switched) voice call and an IMS session to share media of different kinds. Because some use cases are identical to those of multimedia telephony, 3GPP is also working to standardize interworking between multimedia telephony and CSICS. Multimedia telephony uses only one signaling session for all its components, whereas CSICS uses one signaling session for voice and one additional session for the IMS-based media. Therefore, a full interworking solution involves specifying an architecture that splits and combines signaling sessions.

Accesses for multimedia telephony

Multimedia telephony is agnostic to access networks thanks to the existing multiaccess features in IMS. In principle, it can be deployed over any access network that uses

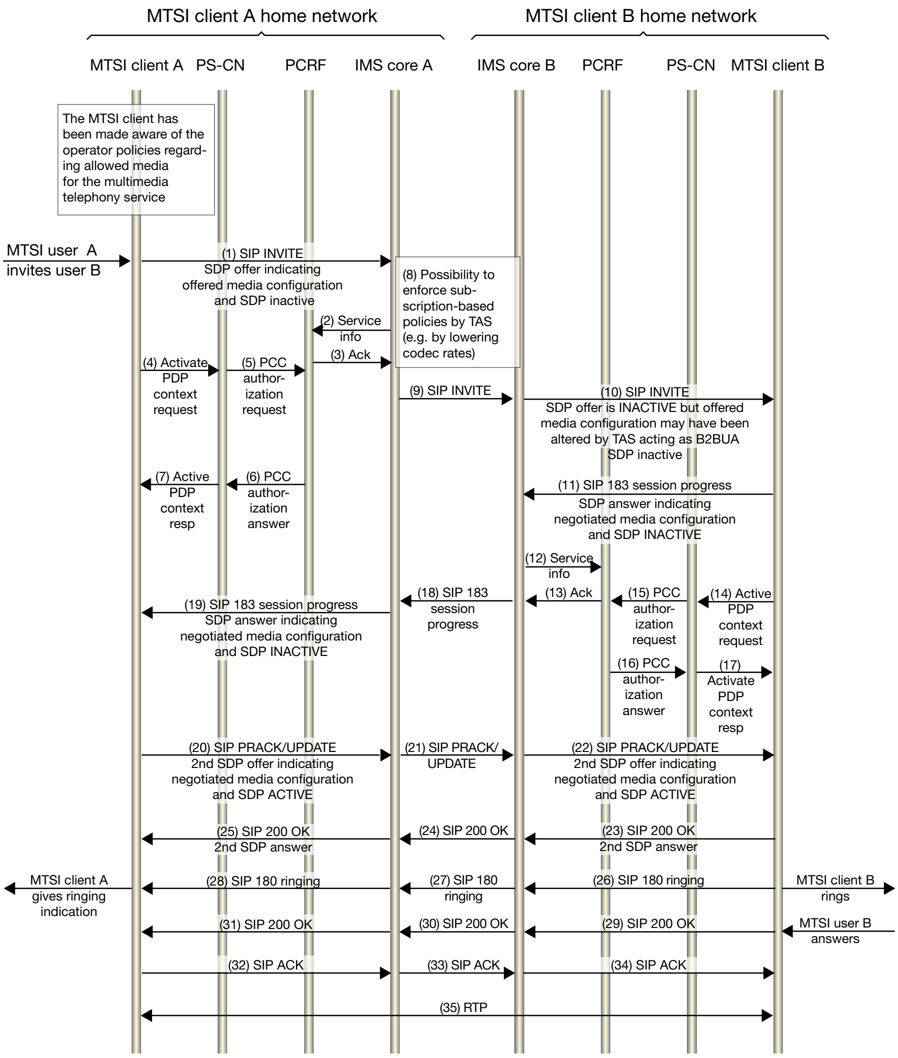

Figure 2
Example of high-layer signaling flow of a multimedia telephony session at setup.

IMS as its service engine. Given its broad applicability, multimedia telephony is an important enabler of true, converged VoIP and multimedia communication.

Realization example: multimedia telephony over WCDMA HSPA

High-speed packet access (HSPA) is a combination of two major WCDMA evolution packages: high-speed downlink packet access (HSDPA) and enhanced uplink (E-UL).[5-6]

To demonstrate how multimedia telephony works throughout the protocol layers, we describe the service as realized over a WCDMA HSPA access network. In the example below, we assume a call case between two users

- in different networks; and
- with HSPA-capable mobile terminals equipped with multimedia telephony clients.

In addition, we assume that the two terminals have registered with their respective IMS networks, and that radio access bearers (RAB) have been established to carry signaling between IMS and the terminals (Figure 1).

The signaling needed to set up a multimedia telephony communication session over HSPA access involves IMS, policy control, the packet core network, and radio access.

Figure 2 shows the signaling flow of a multimedia telephony session at setup. Note: The focus in this example is on higher layers. What is more, we have assumed the use of the network-initiated QoS concept, which is currently under development.

The originating terminal (UE#A) sends a SIP INVITE (1) to indicate that a multimedia telephony session is to be set up. The invitation also indicates

- the media configuration user *A* wants to use; and
- that resources are not available for media transfer (SDP inactive).

The SIP INVITE might also trigger subscription-based policies in the telephony application server (TAS), for instance, the use of a lower coding rate than originally offered (2).

The invited terminal (UE#B) responds to the invitation with SIP 183 (3), which indicates the negotiated media configuration. The response also indicates that resources are not available for media transfer at UE#B. In addition, it triggers *policy control in the policy and charging rules function* (PCRF). This, in turn, causes the packet-switched core network (PS-CN) to set up a bearer through the 3GPP network (PDP context). The negotiation (5-12) also involves setting up a radio access bearer with matching quality-of-service characteristics and high efficiency.

Once resources have been successfully reserved on the terminating side, application signaling continues to the originating side where corresponding resource-reservation procedures are executed (13-22). When both sides have reserved resources, the originating terminal (UE#A) sends a second SDP offer (23-25) to UE#B, which alerts its user.

Call setup is not complete and media cannot start to flow (38) until signaling is sent to indicate that UE#B is ringing (29-31) and user *B* has answered the call (32-37).

The flow in Figure 2 is merely an example. The 3GPP and IETF SIP standards are flexible in terms of the order in which events are executed during the setup phase. For example, according to 3GPP release '99, a terminal (and not the network) can initiate QoS, requesting a bearer and associated QoS.

WCDMA offers four different classes of radio access bearers for data transfer:

- interactive;
- background;

- streaming; and
- conversational.

As long as the terminal is registered with IMS, the SIP signaling between the terminal and IMS typically requires access to a high-priority bearer (interactive). The services information provided in steps (6) and (16) and the operator policies implemented in PCRF determine the type or combination of RABs that will be activated for media.

For multimedia telephony, one conversational RAB should be set up for the voice component. If a video component is to be used, a second conversational RAB should also be set up. The conversational RAB for video is treated with lower priority than the voice RAB, thereby ensuring that voice communication can progress unimpeded should the mobile terminal move out of range (coverage) of the video telephony service. A second interactive RAB (typically with lower priority than for signaling) is used for non-real-time media components, such as text, file transfer, and image sharing (Figure 3).

A conversational RAB over HSPA is mainly characterized by a predefined, guaranteed bit rate and low transfer delay. This is achieved by operating the radio link control (RLC) protocol without retransmissions but using fast retransmissions of small packets between the base station and the terminal. QoS-aware scheduling in the base station prioritizes conversational RAB data ahead of less time-critical data carried, for example, over the interactive RAB used for file transfer. The same HSPA principles are used for interactive RABs, but RLC retransmissions are also enabled at the expense of increased transfer delay to guarantee error-free data streams during bad radio conditions.

To optimize efficiency – in particular, for the transfer of small IP packets that contain voice – WCDMA with HSPA includes a robust header compression (ROHC) scheme that operates over the radio link. When operating in steady state, ROHC typically compresses an IPv6 header from 60 to just 3 or 4 bytes. Greater efficiency is currently being standardized in 3GPP, for example, multiple-input/multiple-output (MIMO) antenna solutions, higher-order modulation, and layer-1 and layer-2 protocol optimizations.

The use of standardized WCDMA HSPA features together with advanced implementations of scheduling functionality make WCDMA HSPA networks a highly suitable access for multimedia telephony. Likewise, excellent flexibility permits dynamic catering to various media and codec rates. HSPA technology has the potential to match or exceed the performance of current circuit-switched communication.[7]

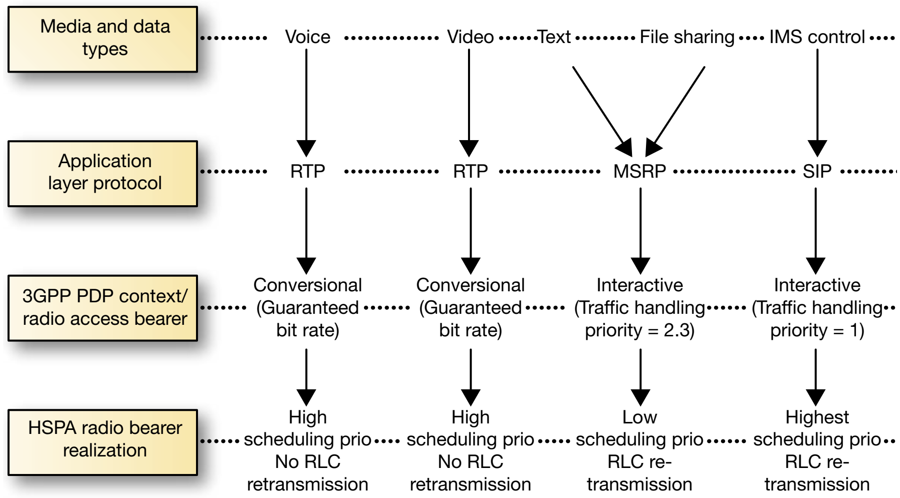

Figure 3
Examples of mapping from media and data types to protocol and bearer realizations over WCDMA HSPA access.

Conclusion

The new multimedia telephony service is set to have a key role in the IMS multiservice ecosystem. For reasons of interoperability and industry acceptance, only standardized solutions can drive a mass market with a multitude of interconnected operators and service providers. In this context, multimedia telephony is not just another IMS service. Rather, it is the global standard for VoIP and general conversational multimedia which will eventually replace today's fixed and mobile circuit-switched telephony services. Multimedia telephony combines the quality, interoperability, reliability, efficiency, regulatory and supplementary services of traditional telephony, with the rich media and dynamics of internet community-based communication.

Multimedia telephony can be launched over any access network that makes use of the IMS service engine. The service is designed to cater to any constraint brought on by mobile access. Consequently, it will run over mobile and fixed accesses alike. Real-time capabilities and inherent flexibility make WCDMA with HSPA a highly suitable access for multimedia telephony.

REFERENCES

1. ETSI TS 181 002, "Multimedia Telephony with PSTN/ISDN simulation services"
2. 3GPP TS 22.173, "Multimedia Telephony Service and supplementary services; Stage 1"
3. 3GPP TS 26.114, "Multimedia telephony; Media handling and interaction"
4. 3GPP TS 24.173, "IMS Multimedia telephony service and supplementary services; Stage 3"
5. A. Furuskär et al, "Performance of WCDMA High Speed Packet Data", in Proceedings of the 55th semi-annual IEEE Vehicular Technology Conference, (VTC'02 / Spring), 2002
6. S. Parkvall et al, "WCDMA Enhanced Uplink – Principles and Basic Operation", in Proceedings of the 61st semi-annual IEEE Vehicular Technology Conference, (VTC'05 / Spring), 2005
7. S. Wänstedt et al," Realization and performance evaluation of IMS MultiMedia telephony for HSPA", in Proceedings of the 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'06), September 2006
8. S. Chakraborty, J. Peisa, T. Frankkila and P. Synnergren, "IMS Multimedia Telephony over Cellular Systems: VoIP Evolution in a Converged Telecommunication World", John Wiley & Sons, Ltd