The Fundamentals of IMS

Abstract

IP Multimedia Subsystem (IMS) is a solution that connects multiple access networks over one IP based core network. IMS is considered to be the corner stone of session management in many next generation networks deployed in telecommunication. Functionality vital for telecommunication operators, such as charging, guaranteed Quality of Service (QoS) and ease of service development is well defined in IMS. Open, well known and thoroughly tested standards such as Internet Protocol (IP), Session Initiation Protocol (SIP) and Real-time Transport Protocol (RTP) are used in IMS. This ensures a stable and future proof solution with broad support from many of the major standardization organizations.

The first part of this white paper discusses the business aspects of IMS and its major advantages. Thereafter follows a technical walk-through describing the IMS architecture and the protocols involved. The content is suitable for anyone interested in getting an overview of IMS and the concepts involved.

Content

1 Introduction

IMS is regarded as the next generation telecommunication system. For example it has been selected to be the media session management system of choice for Long Term Evolution (LTE) networks. IMS offers a variety of possibilities to supply not only the telecom services currently available, but a variety of new ones, cheaper and more efficient than today.

This paper discusses benefits with IMS, both from the operator's perspective and the user's point of view. The paper also gives a technical introduction to IMS and its included parts and concepts.

1.1 Why IMS?

At a first glance, there are lots of benefits with IMS for the service provider and the operator. The operator gets a relatively cheap and convenient way of distributing different services to the end user but can still easily keep track of the transactions. This simplifies end user control for the service provider and thereby gives a new possibility to offer a variety of services.

Different services and applications, formerly only accessible on the Internet, can now be used in new contexts with IMS. This will give the end user, the consumer, the possibility to use all kinds of services similar to those found on the Internet in a more mobile and flexible way than today. The same services can be accessed and used both from the user's fixed terminal at home or mobile device when the user is on the move. This, combined with a higher quality of service than usually delivered with mobile best effort broadband, makes IMS services an attractive option.

IMS is mainly a system in line with the interests of the telecom suppliers and driven by such motives as cost reduction and ease of deployment of new services. This will in the end benefit the consumer by lower prices, guaranteed QoS and a broader spectrum of available services in the telecommunication networks.

Compared to other IP based systems for telecommunication, which are usually proprietary or based on a loose collection of different standards, IMS has some benefits. IMS is an umbrella standard defining most interfaces within and between different networks. This allows telecom operators to buy equipment from different vendors depending on who has the best offer. Operations such as roaming and charging are also covered by the IMS standard meaning that interwork between different networks becomes easier. This allows the service provider to focus on providing high value services to the end customer instead of struggling with low level interwork and integration problems.

Consequently IMS provides a possibility for service providers to offer not only services existing today, but also new ones to the user, and to a significantly lower price than today.

1.2 Advantages for the Operators and Service Providers

The IMS standard will mainly affect the operators. IMS offers advantages for the operators in various areas compared to older technologies. Some of the most central ones are briefly described below.

1.2.1 Cheaper Equipment

IMS is based on IP technology and therefore standard hardware, such as Ethernet switches and cables, can be used. Specialized legacy telecom equipment like Asynchronous Transfer Mode (ATM) switches and Time-Division Multiplexing (TDM) based transmission is often expensive compared to IP-based data communication equipment. The use of one converged IP network for all traffic reduces both the initial and maintenance costs, and thereby substantially lowers the operator's expenses.

1.2.2 Traffic Supervision

IMS provides good possibilities to control the network data transmissions compared to regular IP networks. From an operator's point of view, this increases the possibilities to charge consumers for what they actually use. It also provides them with new opportunities to create new charging models for new and old services. It also allows the operator to prioritize premium services in a standardized way.

1.2.3 Power Saving

Since the IMS architecture is centralized and the network is packet based, the power consumption is reduced compared to traditional circuit switched networks such as GSM. With IMS only one switching architecture is needed for all of the operator's networks. This can be compared to many older architectures where each separate network has its own switching solution. This can lower the power consumption and thus the environmental impact and costs remarkably.

1.3 Advantages for the End Users

IMS allows the operator to provide better services at a better price. This will affect the end users mainly in the ways described below.

1.3.1 More Services

A wider range of services can be offered to the end user when new Internet-like services enter the telecom networks. Further, it will be easier to provide new services to a reasonable price.

1.3.2 Service Provider Competition

The increased possibilities to easily provide services increase the service provider competition, with reduced prices on services as a result.

1.4 Advantages for the Service Developers

The benefits with IMS will affect the service developers as well. Standardized interfaces and platform independent development facilitates the application development and lowers the costs. Technologies that are frequently used and well tested in other areas, like Java EE, can be utilized in much the same way in IMS as in their original context. Because of the experience already available from these technologies the initial start-up time can be reduced considerably. Hence, it is easier to get established on the IMS service market compared to the service market that surrounds legacy telecom systems.

1.5 Involved Organizations

IMS was originally designed by the 3rd Generation Partnership Project (3GPP), which is an organization that mainly develops standards for wireless communication. Their vision was to evolve mobile networks beyond GSM, and IMS is part of that effort.

Later, the Telecoms and Internet converged Services and Protocols for Advanced Networks (TISPAN) began to work together with 3GPP, mainly to extend the IMS architecture with wire-line access support. TISPAN was originally formed 2003 by a merge of the organizations Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) and Services and Protocols for Advanced Networks (SPAN). The standards are constantly evolving and new features are constantly introduced.

Another organization involved in the standardization of IMS is the Internet Engineering Task Force (IETF), mainly focused on Internet standards. IMS is intended to standardize the access of multimedia and voice applications from different terminals, both wired and wireless. To facilitate this, existing Internet protocols are used when possible.

1.6 Relation to other Telecommunication Systems

The IMS system is not intended to fully replace neither existing telecommunication systems, nor the Internet service providers. Rather, it is assumed to be a complement.

Media session switching architectures are not part of, for example, 3GPP mobile phone standard LTE or other recent communication standards. These standards usually focus on access technologies and mobility handling rather than media session management. Here IMS could provide a standardized and comparatively cheap packet based switching alternative.

Another possibility is that the existing switching technology in present telecom systems will be outdated for different reasons. In that case, IMS is a strong candidate to replace the old technology.

2 High Level Architecture

This chapter describes the high-level architecture of IMS.

2.1 Basic Concepts

Below some basic concepts that will be used throughout the paper are presented.

2.1.1 Service

The *service* is what is being delivered to the end user. The simplest service is a normal voice phone call, but it could be anything that gives value to the end user when delivered. A service doesn't necessarily contain only one type of *content*, but could be a combination of several *media streams* that together create the service, e.g. a video conference.

2.1.2 Session

A *session* is an abstract term for the combination of the *content* as well as the necessary *signaling* required to be able to deliver the *service*. The lifetime of a session extends for as long as the service is being delivered.

2.1.3 Signaling

The *signaling* is the traffic that makes all the parts that are required to deliver the *service* work together. It involves the set up of resources and the quality adjustment depending on the conditions for the delivery. Signaling is also required to manage encryption of the *content* as well as the signaling and to be able to save information on the service delivery so that the provider can charge the user for the service.

2.1.4 Content

The *content* is at the core of the *service*. It could be the voice (and video) of a phone call, a movie that is viewed or the content of any other service that is delivered. The content could be delivered as one *media stream* or as several depending on the type of service and the number of different media that are involved. Normal phone calls only carry one type of media, audio, while a movie requires at least two media streams, video and audio. There are a few administrative services that do not contain any content, e.g. redirect services.

2.1.5 Media Stream

A *media stream* is the actual delivery of the *content*. It is the stream of bits that, when put together on the receiving end, is the *content*. Different contents have different characteristics. Voice calls have short bursts of speech while video calls have a constant transmission of the moving images. These different characteristics will put different requirements on the *core network*.

2.1.6 Core Network

The *core network* is the central part of the operator's network. This is where much of the vital functionality such as *signaling* and transportation of *media streams* take place.

2.2 Legacy Requirements

This section describes a number of important legacy use cases that have affected the evolution of the IMS architecture and contributed to a number of requirements. These use cases are all supported by today's telecommunication systems and support for these uses cases in future telecommunication systems is crucial. The use cases are briefly described here together with some common concepts.

2.2.1 Phone Call

The most important use case is to be able to call another person and have a conversation. This use case requires signaling to find the requested user in the databases, given that user's identity, as well as the set up of the media streams necessary for transportation of the voices.

2.2.2 Phone Call to another Operator

However important the first use case is, it will sooner or later be impractical if it is not possible to call a user that is a user at another operator. This adds additional signaling to be able to find another operator's network and establish a call between users in different networks.

2.2.3 Phone Call to another Telecommunication Technology

In the globalized world it is not certain that every operator uses the same technology for delivery of phone calls and other services. However people will still want to be able to call each other. Therefore it is necessary to introduce functionality that translates between the different technologies.

2.2.4 Phone Call when Visiting another Country

Another aspect of the globalized world is the fact that people travel between countries and cannot be expected to buy a new subscription in every country they visit. People want to be able to call their loved ones as well as accessing their services even when they are traveling. To be able to do that it is necessary to strengthen the connection between different operators even more and provide the service of roaming, i.e. make it possible to use your phone as if you where in your own operator's network.

2.3 Emerging Requirements

In addition to the use cases presented in the previous section a number of factors not supported well enough by today's telecommunication systems have formed the IMS architecture.

2.3.1 Easier Integration with Today's Internet Services

The explosion of internet services, for example social networking services, has proven that the "Internet way" is a successful model for development of new services. The ease of development and the short time-to-market are attractive characteristics of these services that can hardly be said to exist in the legacy telecommunication systems of today. An important factor that contributes to this success is the ability to create new services through the collection and compilation of information from many different sources. These kinds of mashups using open APIs has proven to create high value services for the end user. To apply this pattern for service creation and deployment also within telecommunications and to merge the Internet and telecommunications worlds could provide the base for new and innovative services.

2.3.2 Cost Reduction

The competition between service providers leads to reduced margins. To compensate for this service providers are always on search for ways to cut their costs. One way to achieve this is to unify the session administration and service delivery for all different kinds of access networks to reduce the number of different technologies used in the networks.

2.3.3 Delivery of Demanding Multimedia Services with High Quality

As services get more and more rich with streaming video and advanced graphics the networks must be able to cope with large amounts of data being transferred. It must therefore also be easy and cheap to expand the network and add more capacity whenever needed. Furthermore it´s not enough to provide high bandwidth, the transmission characteristics such as delay and jitter must also be very good. Bad network conditions are very harmful to the end users perception of services such as real time communication.

2.4 Architecture

The architecture of IMS can be divided into three parts to better understand it. There is no exact definition of these parts, but here it is described with the basic concepts given earlier. Figure 1 illustrates the different parts and their relation to each other.

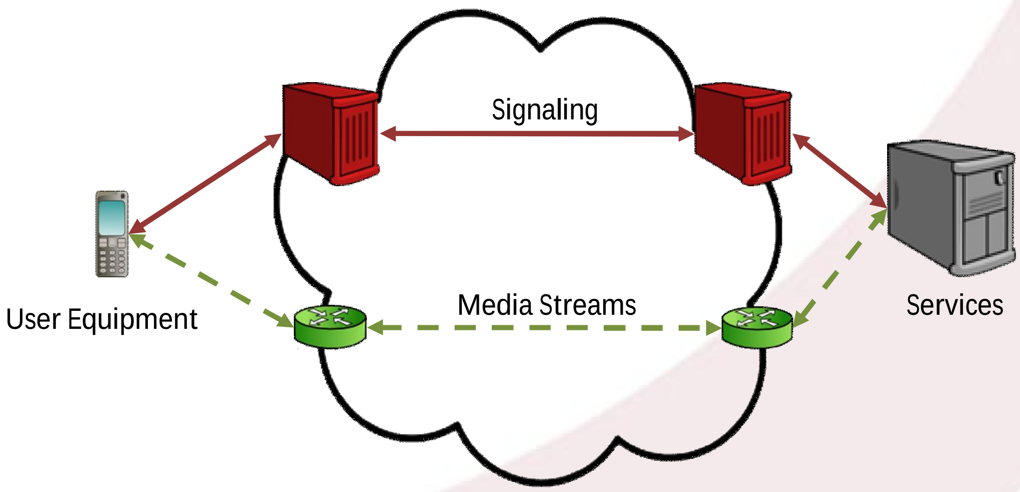


Figure 1 The different parts of the IMS architecture

2.4.1 Signaling

Signaling is the core of the IMS architecture. This is where the switching and routing of messages happen. This is also the part that handles security and verifies the user’s identity and rights. All sessions are set up through signaling, which in turn controls the media streams.

2.4.2 Media Streams

The content of the sessions is not sent through the signaling nodes, but instead in the part specialized for transportation of media streams. It consists of an IP backbone that is capable of transmitting large amounts of data in a lot of concurrent media streams. The backbone must also have good traffic characteristics with respect to delay and jitter. Furthermore the backbone should have the possibility to differentiate between the packet streams and give priority to those belonging to premium services to control the QoS. Which streams to prioritize is determined by signaling nodes at session setup and communicated to the transport nodes.

2.4.3 Services

This is where all the services, provided by the operator and potential third party providers, reside. These services can contribute with pure signaling (e.g. redirecting a session), content only (e.g. a weather service) or a combination (e.g. video conference system)

3 Protocols in IMS

This section describes the most central protocols used in IMS. Such protocols include SIP for session control and RTP for media transport. Further, other protocols for data transport and security used in IMS are briefly described.

3.1 Session Control

In IMS, SIP is used as signaling protocol for session handling. IMS utilizes both the basic functionality of SIP and also a number of extensions to SIP. In the following subsections the basic SIP protocol and its accompanied protocol Session Description Protocol (SDP) is described.

3.1.1 Overview

SIP is an application layer control protocol for initiating, managing and terminating multimedia sessions across IP networks. The participants in a session range from people to various types of devices, like signaling switches in the core network or application servers. SIP also provides methods for user registration and location.

With SIP an email-like address is introduced. An example is sip:alice@attentec.se where Alice is the registered user in the Attentec SIP domain. Based on this address, SIP message routing is performed between intermediate SIP servers.

3.1.2 Components

There are a number of SIP specific components to realize a multimedia session. The two main logical components are the User Agent (UA) and the SIP server. In IMS a UA may, for example, be a mobile phone. The SIP servers are located within the IMS core network. In SIP a UA is in turn divided into a User Agent Client (UAC) and a User Agent Server (UAS). Session establishment is initiated from the UAC towards the UAS.

There are different kinds of SIP servers and the two main servers are described here. The first one is the intermediate Proxy server that enables message routing between UAs. The Proxy server can act either as an outbound or inbound Proxy server depending on if it forwards SIP messages from or to the UA respectively. The second server is the Registrar server. A Registrar server functions as a location server where users register their SIP address together with their IP address.

Figure 2 shows a typical SIP network setup and communication paths between the different nodes. Along with the two UAs the network consists of an outbound SIP proxy, a Domain Name System (DNS) server, an inbound SIP proxy that also functions as a SIP Registrar and a location server database.

The UAC to the left sends a request towards its outbound SIP Proxy. For a UAS in a different domain, a name-address resolution is performed by a DNS service. The resolution mechanism enables routing of SIP messages from the outbound SIP Proxy to an inbound SIP Proxy in the external domain wherein the endpoint resides. The outbound SIP Proxy uses the DNS server to resolve which domain the request should be forwarded to. An inbound SIP Proxy receives the request and connects towards the location server database to find the IP address that corresponds to the SIP address in the request. The request is then forwarded to its final destination UAS based on this IP address.

A response is routed back in the exact reverse route path. At this point the UAC and the UAS have exchanged each other's IP-addresses and any consecutive request(s) between the UAC and the UAS can be sent directly between the two.

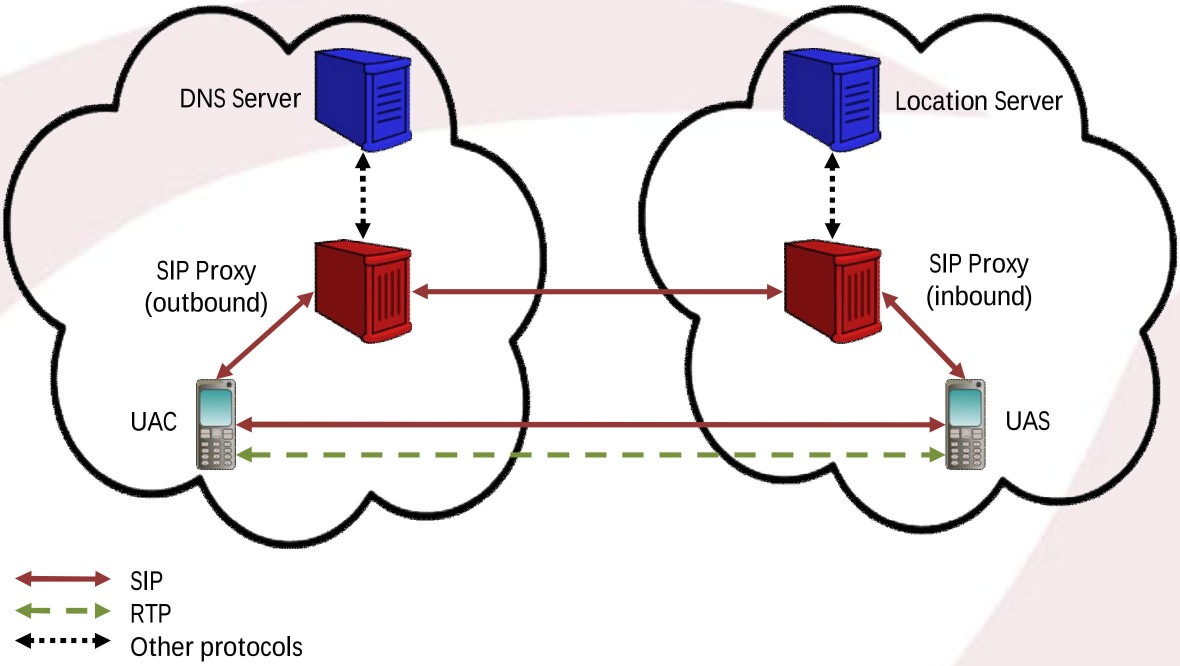


Figure 2 Protocols and components commonly involved in a multimedia session that uses SIP for signaling

3.1.3 Messages

SIP is a text encoded protocol and is based on request messages initiated by the UAC and response messages from the UAS.

Each request and response message consists of a header and a body. The messages are highly influenced by Simple Mail Transfer Protocol (SMTP) and Hyper Text Transport Protocol (HTTP) and the header contains similar clear text fields as in those protocols. For example, the header fields state information about source and destination addresses as well as information about the body's content and length. The body can carry any type of data. However, the most common one is SDP, which describes the individual media streams used in the session.

There are a number of SIP request messages that are used to manage a session. Below is a short description of the most common ones:

- REGISTER is sent to the registrar server, which in turn stores the IP address currently used by the UA. The IP address is later used when a session is initiated towards the registered UA.
- INVITE is sent from the UAC to the UAS to initiate a session.
- ACK concludes the INVITE. After an ACK the session is established and the actual media session can start.
- BYE will terminate the session and all included multimedia streams.

- CANCEL terminates any previous request.
- OPTIONS asks for information about the capabilities, for example SIP extensions, of the other party.

Apart from the request messages there exist a number of response messages. The response messages are highly influenced by HTTP and the most common SIP response is the 200 OK response.

3.1.4 SIP Registration

Figure 3 shows a user registration request. The REGISTER request is sent from the UA to the Registrar server where in turn the UA's SIP address together with the UA's IP address is stored. A 200 OK response is sent back upon successful registration.

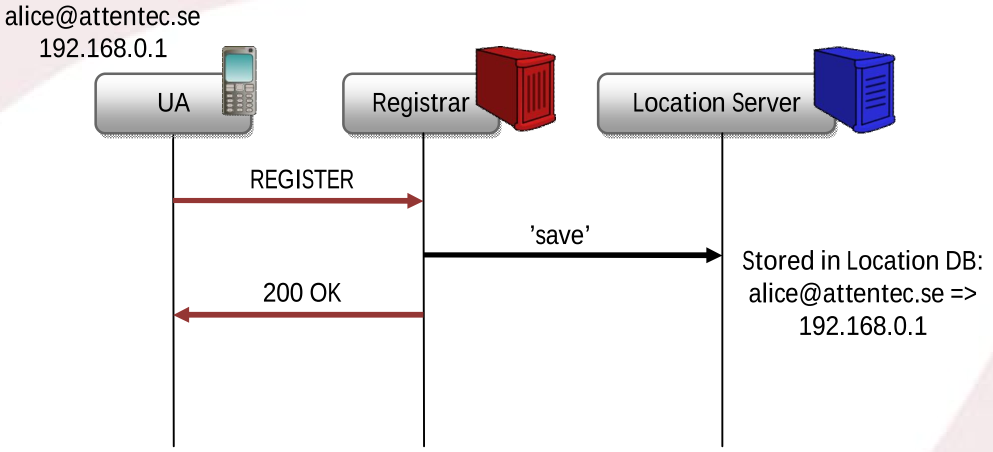


Figure 3 SIP Registration

3.1.5 SIP Media Session

Figure 4 depicts a basic session setup between a UAC and a UAS. An INVITE is sent and responses that follow indicate the progress. Upon an ACK request the data traffic can start. When one of the session participants issues a BYE request the session is terminated.

In IMS the basic SIP protocol is extended to cover all the aspects of a complete session protocol suitable for the demands required by a telecommunication network. Therefore a number of SIP standards exist that defines extensions that add functionality to SIP either by adding new messages or by adding fields to the SIP header.

SIP is often accompanied by SDP. SDP describes the different multimedia streams, their capabilities and stream addresses of the session. The protocol implements an offer-answer mechanism which negotiates multimedia capabilities of the participants. With SDP it is possible to define the media types, format and transports attributes which in turn makes it possible to combine, for example, voice and video in a single session.

The SDP data is sent as SIP payload data during the session setup phase. As depicted in Figure 4, the SIP INVITE request includes the SDP information which defines individual media streams and the UAC's media compression support. The SIP response messages 200 OK provides the answer about which media streams are supported and accepted by the UAS. After the setup the parties can start their media streams.

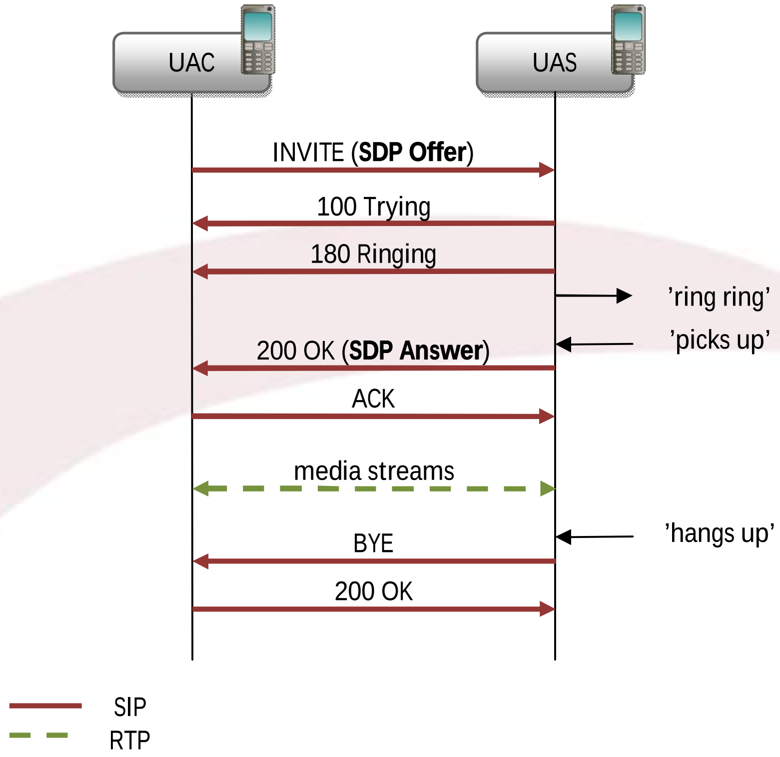


Figure 4 Session setup with SDP offer-answer mechanism

3.2 Real-Time Multimedia Transportation

The Real-Time Transport Protocol (RTP) is an application level protocol specifically designed to transport real-time multimedia. Each RTP packet contains media samples generated by the participants involved in the multimedia session. For each media type present in a multimedia session a separate RTP stream is set up. Based on the SIP signaling and the SDP information the participants can negotiate which media compression techniques to use. When the participants have agreed which compression technique(s) to use, the communication can start over RTP. Due to time constraints, packet retransmissions are not feasible in RTP media streams. Because of this, UDP, which is a best-effort protocol, is preferred over reliable protocols with packet retransmissions such as TCP.

Each packet in RTP is implemented with media transport information, like timestamps and sequence numbers. This information enables the client to adjust the media quality based on network conditions such as packet loss, reordering and jitter. To carry this control information and statistics RTP Control Protocol (RTCP) is used. In multimedia sessions where multiple media sources exist, RTCP can also be used to carry information to allow synchronization between the different sources.

The use of the protocols RTP and SIP makes the signaling and media transportation disparate. Thus different routing paths are possible. In IMS, where these protocols are used, the IMS network only handles SIP traffic where as the RTP traffic traverses other network nodes. Thanks to this, unnecessary routing delay for the media traffic can be kept to a minimum.

3.3 IP Version 6

The successor of IP version 4 is the IP protocol version 6. Foremost, IP version 6 meets the new needs for a larger address space, better support for emerging multimedia applications and simplified routing.

IP version 6 has a number of new features but the most important ones are the following. First, the address field is extended from 32 bits to 128 bits. Second, the header format is simplified compared to IP version 4 which makes IP version 6 more bandwidth efficient. Fields in IP version 4, such as the checksum, are removed to enable faster processing and routing of the packets through the network routers. With the checksum field removed the packet verification is performed by the lower layer link protocols or higher layer transport protocols.

Phasing out IP version 4 requires a transition period where both versions can coexist. Thus interoperability techniques between the two protocols are needed during this period where hardware and software are updated. Some common techniques already deployed, and some also present in IMS, are for example dual communication stacks, protocol tunneling and Network Address Translators (NAT).

3.4 IP Security

IMS utilizes the IPsec facility to enable access and network security. IPsec has support for authentication, data integrity and confidentiality. In IMS, IPsec is used both when a user authenticates towards the network and for general traffic protection. Also, within the IMS network and between IMS networks IPsec is used for traffic protection.

The IPsec facility consists of two protocols, Authentication Header (AH) and Encapsulating Security Payload (ESP). AH provides authentication, where the authentication covers both the IP header and the payload. ESP provides authentication, data integrity and confidentiality but covers only the payload data.

Two communication modes are supported by IPsec, transport mode and tunnel mode. In tunnel mode the payload consists of an entire IP packet, whereas in transport mode the payload only consists of the IP payload.

Only ESP is used in IMS, hence protection for the payload data. For access security and user authentication, ESP in transport mode is used. To enable secure communication between different IMS domains ESP in tunnel mode is used.

3.5 Manage Network Access

In a telecommunication network there is a need for management of Authentication, Authorization and Accounting (AAA) services. In short, with Authentication the user's identity is verified. The Authorization determines the user's access privileges to the network resources and services. Finally, the Accounting keeps track of a user's resource usage, for example to enable correct invoicing.

Diameter is a signaling protocol that provides the AAA services in an IMS network. Some of the main features found in Diameter are for example that it uses reliable transport protocols for sending its binary information, it has support for information security aided by IPsec and it is easily extended with new signaling capabilities.

The binary information, usually sent over TCP, consists of a container that holds attribute-value pairs for AAA data. To communicate the AAA data between the network entities, Diameter defines a number of commands to signal the requests and responses that contains the AAA data.

Diameter is a general base protocol defined by IETF. However, it is extended with Diameter applications that provide more specific AAA data containers and commands suitable for the purpose. In IMS, Diameter is used for authentication and authorization of the clients and accounting to correctly handle charging, policy enforcements and QoS. A number of Diameter extensions, also defined by IETF, are used in IMS, for example extensions for SIP and policy handling.

4 The Design of IMS

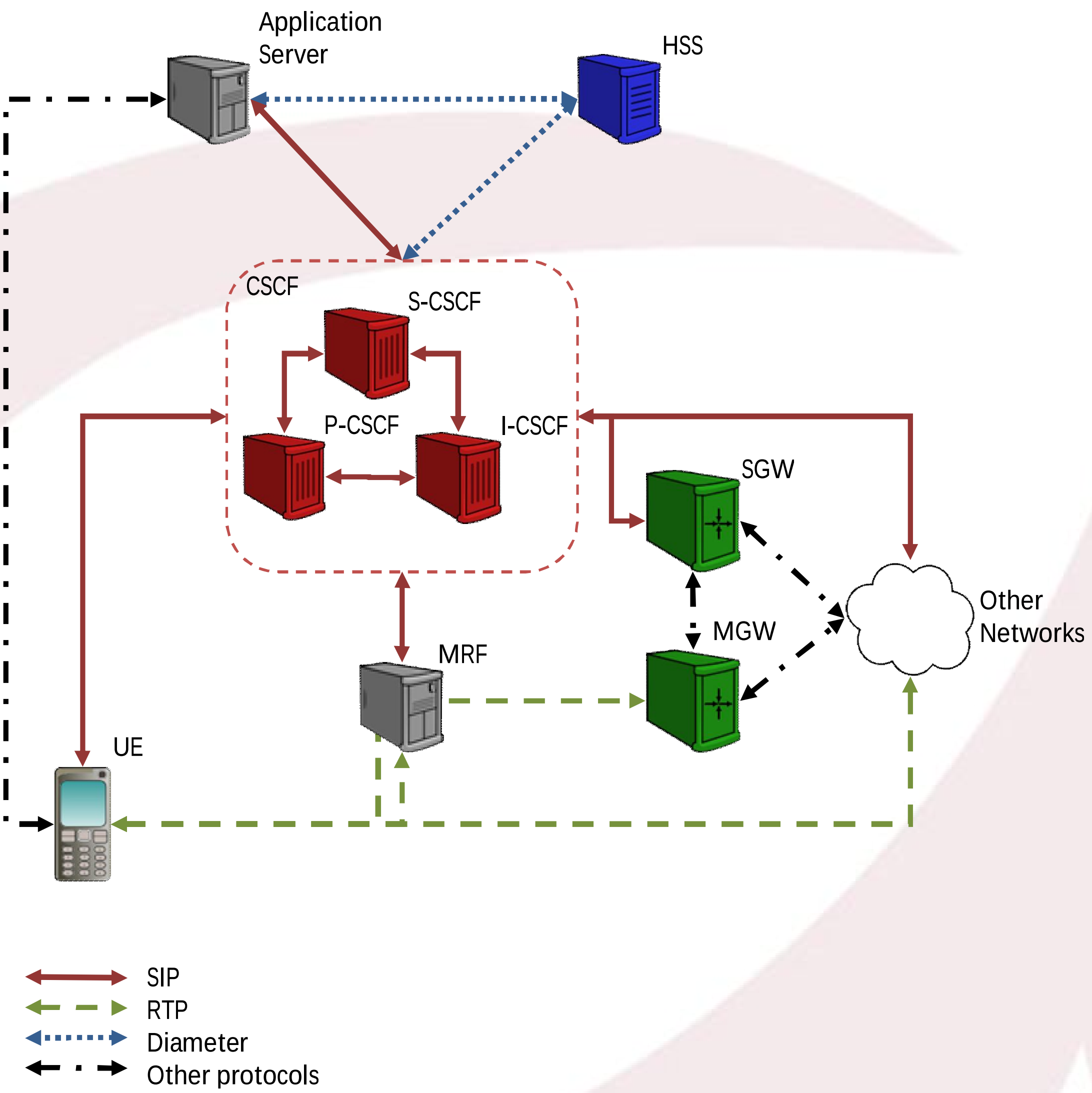
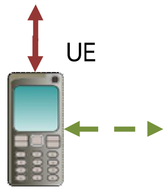


Figure 5 High level design of IMS, end user equipment (UE), switching infrastructure (CSCF), databases (HSS), gateways (xGW) and application servers

The IMS design is, as most telecommunication architectures, based on a number of logical functions and standardized interfaces between these functions. Compared to traditional data communication architectures, that usually are no more complex than a server and a client, the IMS architecture may seem very complex. However, one must remember that IMS is designed to meet the telecommunication requirements derived from the use cases described in 2.2.

The different functions in IMS can be grouped into five main categories. These are end user equipment, switching infrastructure, databases, gateways and application servers. In this section the most central functions in every category, and their interaction with each other, will be presented. Figure 5 shows a high level picture of the IMS design where the most important functions are present.

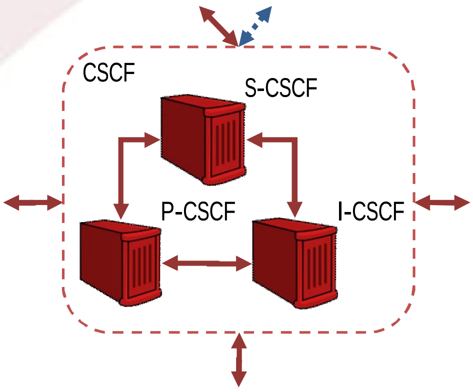
4.1 User Equipment



A User Equipment (UE) in IMS is any device that can be used for electronic communication. This spans the range from plain analog telephones to full featured computers and anything in between, cell phones, IP-TV boxes, etc. The range of services possible to deliver from the network will of course differ a great deal between the different types of UEs though.

Basically any native IMS UE should be able to communicate directly with the network over SIP by acting as a SIP UA. The native IMS UEs can also be expected to handle most services such as voice and video. Legacy systems that use gateways to connect to IMS may only take part of a subset of services such as basic voice calls.

4.2 Switching



The switching core is the brain of IMS. It consists of a number of Call Session Control Functions (CSCFs) which together make it possible to set up multimedia sessions through signaling with SIP. There are three different CSCFs: the Proxying CSCF (P-CSCF), the Interrogating CSCF (I-CSCF) and the Serving CSCF (S-CSCF), which all perform different and well defined tasks in the network. The CSCFs are all different types of SIP proxies. It should be noted that these logically separated functions may be implemented in the same node.

4.2.1 P-CSCF

The P-CSCF is the first function in the IMS network which the UE interacts with over SIP. In SIP terminology the P-CSCF is an outbound SIP proxy.

The P-CSCF routes SIP messages from the UE further into the IMS network. It also authenticates the UE and establishes IPsec connection to it for signaling upon connection. It may manage the setup of QoS parameters in the network to which the UE is connected to secure that enough resources are present in the network to deliver the requested service with the quality intended. If narrow-band radio access is used by the UE to connect to the IMS network, compression and decompression of SIP messages can be performed by the P-CSCF. To protect the rest of the IMS network against malformed signaling the P-CSCF often performs syntax validation to verify the correctness of SIP messages sent from the UE. It may also collect and report charging information, based on the session that is set up.

The P-CSCF determines where to route the SIP messages based on DNS lookups of the user's domain name in the SIP address.

The location of the P-CSCF can be in the home network, the network to which the user is a subscriber, or in the visited network, for example if the user is abroad[1].

4.2.2 I-CSCF

The I-CSCF is the gateway into the user's home network. It is the address of the I-CSCF that is registered in DNS for the operator's domain. The I-CSCF is an inbound SIP proxy.

The I-CSCF forwards SIP messages to a suitable S-CSCF depending on a number of factors. Such factors may include the type of subscription that the user has. Some S-CSCFs may, for example, be more powerful and suitable for delivery of advanced services to premium subscribers while others only can provide basic call services. This type of information is fetched from the user databases by the I-CSCF over the Diameter protocol.

After the initial signaling of a session setup, the I-CSCF can be configured to stay in the signaling path. If it remains in the path it can perform the work of a Topology Hiding Inter-network Gateway (THIG). This means that it can encrypt or even remove some parts of the SIP messages which are not vital to the network elements outside of the operator's network, but which could expose information about the topology of the network.

The I-CSCF is usually located in the user's home network.

4.2.3 S-CSCF

As the name implies the S-CSCF is the node in the network that serves the user, hence it has a very central role. In SIP terminology the S-CSCF is a combined SIP Proxy and Registrar.

When the user registers with the IMS network the S-CSCF downloads authentication information from the user database over Diameter. This information is sent as a challenge to the UE. The S-CSCF also downloads the user's service profile from the same database.

The S-CSCF downloads the user information. Thereafter the S-CSCF stores its own address together with the user information in the user database. This is done to allow any I-CSCF to contact the user database for the duration of the user's registration to find out which S-CSCF that has been allocated for that user.

When the user performs session setup, or receives a session setup request, the S-CSCF examines the user's profile to determine if any additional services should be executed. It also performs session control, that is, verifies that the user is allowed to perform the requested actions.

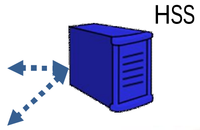
If the S-CSCF finds that additional services should be executed for a user the request is routed to an application server which executes the service. For example, the S-CSCF could detect that the user to which a call is destined is not online. This would then make it re-direct the session setup request to an application server that handles voice mail.

---

[1] If GPRS is used as access method the P-CSCF will be located in the same network as the GPRS Gateway Support Node (GGSN).

The S-CSCF is always located in the user's home network.

4.3 Databases



There are two standardized databases in IMS which are used for different purposes, the Home Subscriber Server (HSS) and the Subscription Locator Function (SLF).

4.3.1 HSS

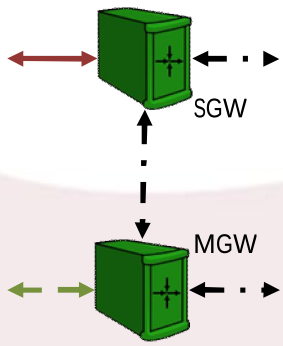
The HSS is the database for user related information in IMS. Such information includes the following:

- Authentication information, for which the HSS holds a copy of a shared secret key. The other copy of this key is stored in the UE. Upon registration the UE and IMS network prove to each other that they know the secret key.
- Authorization information used to verify what the user is allowed and not allowed to do.
- The user profile, which includes information about the type of subscription the user has and which services the user has activated.
- Application specific data that trusted application servers are allowed to store data in the HSS.
- Information about which S-CSCF the user is currently registered with.

In case the operator has a big network, more than one HSS will probably be needed to store all user information. In this case a special database is needed to keep the mapping from user to HSS. This is the sole purpose of the SLF.

The HSS uses Diameter to communicate with other functions in the network. It has interfaces towards the S-CSCF, the I-CSCF and trusted application servers.

4.4 Gateways



The gateways are central parts of the IMS architecture for interaction with other types of networks. The gateways can be divided into two main categories, Signaling Gateways (SGW) and Media Gateways (MGW). Apart from these gateways there are also additional functions defined. For example, to convert between IPv4 and IPv6, and security gateways to set up encrypted connections over IPsec to other IMS and IP networks.

4.4.1 SGW

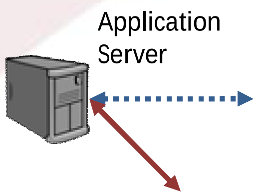
The signaling gateways are used to translate signaling to/from SIP. In this way the interfaced network can use IMS and its CSCFs for session establishment. From an economical point of view this is very beneficial since it allows the operators to get rid of costly switching equipment in other networks. Instead the switch can be replaced by a (in comparison) rather simple gateway to IMS.

For interaction with the Public Switched Telephone Network (PSTN) the signaling gateway has been divided into two logical functions. The SGW, which takes care of conversion between IP and non-IP networks and the Media Gateway Control Function (MGCF) which converts between conventional telephony protocols and SIP.

4.4.2 MGW

The MGW converts user media data between non IP-based networks and RTP, which is used for media transport in IMS. This conversion allows users not connected to an IMS network to establish multimedia sessions with IMS users. The MGWs are controlled by a signaling function, for example an MGCF, which instructs it to activate or deactivate sessions.

4.5 Application Servers



The application servers are used to provide the user with services in addition to basic multimedia session setup. Such services may include voice-mail, instant messaging, presence, etc.

The typical IMS application server communicates over SIP. These may function as any type of SIP entity, for example UA or Proxy Server.

Besides SIP the application server could implement several other protocols. For example, the UE may be allowed to access the application server over HTTP to allow the user to configure the service.

Together with the application server, which in IMS is defined as a signaling function, there is often some type of media function which provides the media content of a service. These may be collocated with the application server and controlled directly by the application server. The IMS specification specifies a more general solution for this called a Media Resource Function (MRF). The MRF communicates over SIP and can be interfaced via the S-CSCF to provide functionality such as playback of media streams.

The application servers can be located in the operator's network or in the network of a third party. If the application server is within the operators network it is usually referred to as trusted and can access the HSS over Diameter to store and retrieve application data.

5 Conclusions

In this paper IMS has been discussed from a number of viewpoints. To begin with, the perspective of the operator, service provider and the end user was discussed. Further the organizations and committees involved in the standardization work were briefly described. Finally a technical description of IMS was presented where the different protocols and nodes involved were thoroughly described.

For the operator IMS provides a possibility to make large savings in both operational and capital expenses. By the use of one converged IP based network for switching and service delivery, instead of many different technologies in separate networks, the amount of equipment needed can be reduced significantly which in turn should lead to reduced costs. The design of IMS and the protocols used allows the operator to fulfill the tough requirements on availability and reliability normally present in telecommunication. Furthermore, simpler and standardized interfaces facilitate fast service development and deployment. For the end user this is likely to result in a higher number of attractive services available.

The IMS design and the protocols create a solid platform on which arbitrary multimedia sessions can be realized. The legacy telephony use cases such as basic session handling, roaming and technology interoperability, as described in this paper, are all well supported. But the supported use cases are not limited to telephony only, with IMS any type of service and content can be delivered with the same flexibility.

All this put together makes IMS an attractive option for any telecommunication operator that wishes to cut the costs and deliver more attractive services.