IMS and technology

From Gunnar Heine

Translation of an article of German FUNKSCHAU magazine of April

What is IMS? Since quite some time, the IMS or rather IP-Multimedia Subsystem is on the radar screen of manufacturers and telecommunication operators.

To make a long story short, the IMS forms a "basis to provide IP-based communication services". The offer of this communication services ranks from ordinary calls to multimedia-based offers of all kinds.

Among other things, the IMS permits the provision of the frequently mentioned triple-play-services. In figure 1 a central characteristics of the IMS receives a definition: The IMS is situated between the ultimate user, who is behind the respective IP-CAN (IP Connectivity Access Network) and the real service (green network clouds) on a different point. The IMS takes over the function as an agent or intermediary between this service and the ultimate user in this process. We will have to point out the central definition of the IP-CAN later in detail. But here we reveal as much as the GPRS/UMTS- as well as WLAN/WIMAX-networks and also all usual DSL-connections are suitable and popular examples for IP-CAN's.

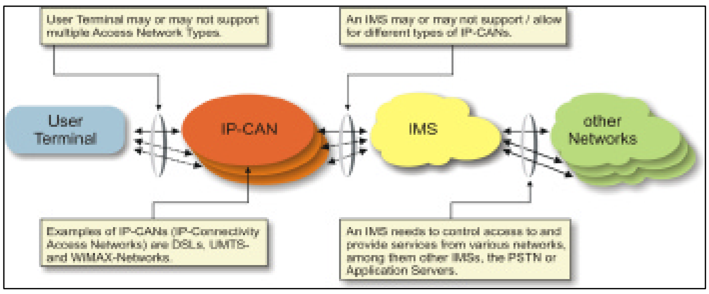

Figure 1: The IMS as agent between user

Why an IMS-based solution? Of course there are many good reasons to use the IMS as service-platform. Below we want to present to you the most important ones:

- First of all the all above mentioned agent-qualities are to be remembered. Here, the following aspect is important: A given service is only interesting for the provider if this service can be charged in the first place. And this is exactly what the IMS enables.

Charging works through the definition of appropriate interfaces like e. g. Ro and Rx and the definition of corresponding amendments to the DIAMETER-protocol.

- The IMS is an extreme open system. In the ideal case, the addition of new services takes place without big efforts on integration and test via the so-called ISC-gateway, which the IMS connects with different kinds of application servers. As a service provider you may want to consider the giant portals of pictures and mini-video appearing out of nowhere. Using the appropriate creativity, these portals can be diversified much better using the IMS, but also considerably safer and, above all, this way, you can make money with it.

FMC (Fixed Mobile Convergence) becomes a reality through the IMS. This aspect cannot be overrated for different reasons in its importance at all. The inherent FMC at the IMS is given a support by IP as transportation-network and protocol. Or in other words: For the connection between IP-CAN and the user, literally any transport network can be used that can provide IP-support. Whether such a transport network permits mobility or not is less important. This situation becomes especially understandable in figure 2. In a fascinating way the IMS allows also the realisation of another feature, the so-called Seamless Mobility which was praised for a long time through the Marketeers. With respect to seamless mobility, we talk about the (more or less) unnoticed change from one IP-CAN to another. This change is reported to the IMS via a SIP-based re-registration after the user has obtained a new IP-address in the new IP-CAN.

This type of mobility is likely called "IP-Mobility" or "Macro-Mobility" in contrast to the "Micro-Mobility", which refers to the mobility within one IP-CAN, e.g. inside a UMTS-network.

Why an IMS-based solution?

Of course there are many good reasons to use the IMS as service-platform. Below we want to present to you the most important ones:

First of all the all above mentioned agent-qualities are to be remembered. Here, the following aspect is important: A given service is only interesting for the provider if this service can be charged in the first place. And this is exactly what the IMS enables. Charging works through the definition of appropriate interfaces like e. g. Ro and Rx and the definition of corresponding amendments to the DIAMETER-protocol. The IMS is an extreme

open system. In the ideal case, the addition of new services takes place without big efforts on integration and test via the so-called ISC-gateway, which the IMS connects with different kinds of application servers. As a service provider you may want to consider the giant portals of pictures and mini-video appearing out of nowhere. Using the appropriate creativity, these portals can be diversified much better using the IMS, but also considerably safer and, above all, this way, you can make money with it.

FMC (Fixed Mobile Convergence) becomes a reality through the IMS. This aspect cannot be overrated for different reasons in its importance at all. The inherent FMC at the IMS is given a support by IP as transportation-network and protocol. Or in other words: For the connection between IP-CAN and the user, literally any transport network can be used that can provide IP-support. Whether such a transport network permits mobility or not is less important. This situation becomes especially understandable in figure 2. In a fascinating way the IMS allows also the realisation of another feature, the so-called Seamless Mobility which was praised for a long time through the Marketeers. With respect to seamless mobility, we talk about the (more or less) unnoticed change from one IP-CAN to another. This change is reported to the IMS via a SIP-based re-registration after the user has obtained a new IP-address in the new IP-CAN.

This type of mobility is likely called "IP-Mobility" or "Macro-Mobility" in contrast to the "Micro-Mobility", which refers to the mobility within one IP-CAN, e.g. inside a UMTS-network.

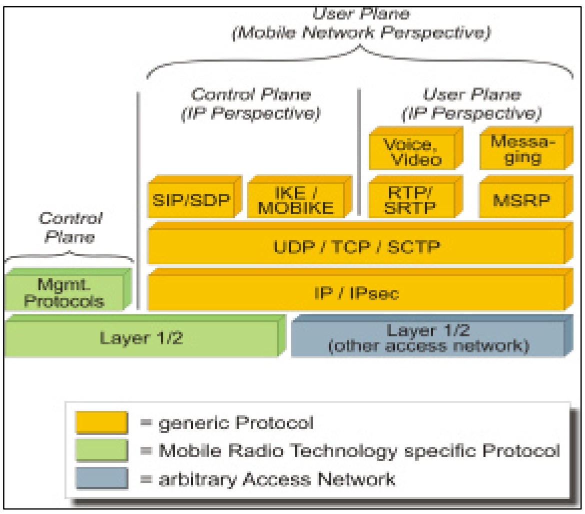

Figure 2: protocol stack of an IMS User Terminal

At the top we had pointed out "different reasons" for the importance of FMC. Now we will go into detail briefly: Certainly there is the perspective that the end costumer is available on one device at home as well as through another device while on the move. From this, most different services can be deduced and presented. However for the service-provider, the aspect of mergence of so far completely independent branches of business is much more interesting (PPT: landline and mobile phone), or the possibility of the development of completely new branches of business (e.g. WIMAX-network providers appear as mobile network operators). The appearance of the aspect depends on the current type of service provider in each case.

Another point is the question of expenses: For many operators the question about alternatives or the enlargement of the existing TK-infrastructure is raised today or in the near future. Without telling numbers now, you can say that the IP-based-equipment takes a lot less capital expenditure (CAPEX). The operation results in less costs already through the matter of fact that after the implementation of the IMS networks which have been separated before, are easily affiliated and so they can be used together. However, considering our experiences with IMS-commissioning and test, I like to warn the reader of too high expectations that die IMS "consists just of of a few computers" that do not cost that much.

"Last but not least" the standardisation has to be mentioned as an important reason for an IMS-based solution. The advantages and disadvantages of standardised solutions in the telecommunication have been discussed enough elsewhere and shall not be mentioned here again. However, it is remarkable that the IMS shows the synthesis between the hardly manageable and little standardised IP-world and the well-defined TK-world. Indeed, the IMS in its inside is nothing else than a composition of SIP-servers and media gateways. True: You can get that without IMS. But the special of the IMS is the standardised type of interior and external communication. We will treat the question of "how did that come to a solution?" in the following chapter.

History of IMS

The history of the IMS starts shortly after the turn of the millennium in the years 2002/2003 within the standardisation bodies of 3GPP. This umbrella organisation is responsible for UMTS and also works on the further evolution of GSM. In the 3GPP Release 5-standards the IMS is nothing else than an amendment of the existing packet-switched core network of GSM/UMTS access networks, To interested readers I recommend especially the specifications 3GTS 23.228, 24.228 and 24.229 for further reading. By this time the target inside 3GPP was the definition of a SIP/IP-based service-platform for the existing GSM- and UMTS-mobile-network. At that time again, they were not considering landline, DSL or other alternatives.

However the Release 5 reflects this focus clearly and you have to confess: Neither the IMS at this time or with this release was strong, nor any commercial IMS-service has been activated in these years until 2006 anywhere. However, the overall direction was specified and maintained until today. Technically, the system's core of Release 5-7 or even 8 did not change. However there have been serious changes and adoptions for to be able to provide for alternative and additional access networks like e.g. DSL or WLAN/WIMAX.

Since then, the more politically or economically important developments are the adoption of IMS as service-platform by various other groups. In this context, TISPAN and the North American dominated 3GPP2 have to be pointed out. Even organisations without any affiliation with 3GPP, like DSL- or cable-TV-forum, have discovered the IMS as interesting since then. Roughly you can tell the cooperation of the different standardisation committees in the following task sharing: While 3GPP is still taking care of standardising the actual IMS, the other organisations are working on the definition of adapters or IMS-amendments for tailoring the IMS to their particular access-networks.

At this time, the standardisation of the IMS still continues and is far away from coming to an end, even though the fact that already many commercial IMS-offers are already available to the customer. So we have to accept the matter that a standard does not work perfect before its implementation und during this, the standard must be improved constantly. At least a big part of the problems with the IMS, which are to be described, are due this history and the grown structure.

Architecture and protocols of IMS

In figure 3 you can find a strongly simplified illustration of IMS, which describes the main components. To allow the relation to figure 1, the IP-CAN, the user agent (UA), the application & server domain and the PSTN are also depicted. In the inside of the IMS-cloud is an IP-Backbone through which the various servers communicate.

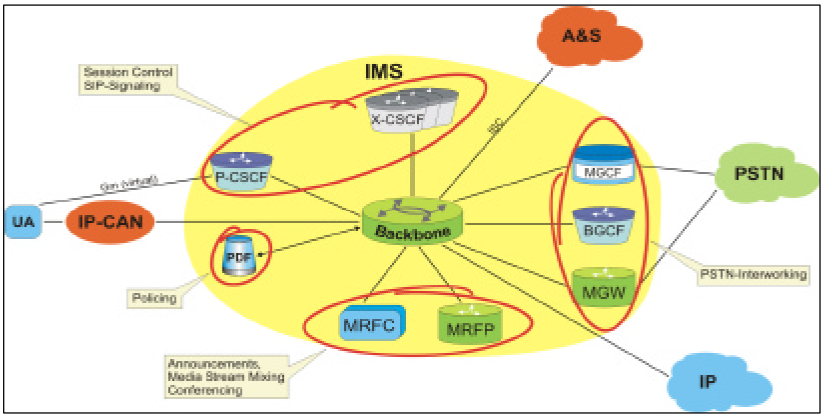

Figure 3: Internal Layout of the IMS

Here, various CSCF's or Call Session Control Functions are taking the most important function. There are three different types, named P-CSCF, I-CSCF and S-CSCF. Every single one is a functionally extended and specialised SIP-Server.

The P-CSCF stays in contact with the UA's and is able to establish and operate e.g. an IPsec-based tunnel.

The S-CSCF acts like a VLR in the mobile network and takes the function as SIP-registrar. Very important here: Every IMS-user registers always in a S-CSCF in his or her home IMS, even in the case of roaming.

The I-CSCF takes on the gateway-function in direction of foreign networks and acts as first point of contact for entering SIP-calls of any type. The only exception: Incoming PSTN-calls: Here we touch already one of the existing IMS-problems. At least today, the 3GPP-standards do not cover incoming IMS-calls from the PSTN. This lack obviously leaves room for proprietary solutions. Interesting is, that the other way around (IMS originating calls to the PSTN) is well described in the specifications already since Release 5. This leads us to the BGCF, which was defined for reducing Interconnection costs, and to the reasons for integrating MGC and MGW.

Probably but not mandatorily present is the MRF, which is divided in MRFC and MRFP. Primary tasks are the play-back of announcements as well as Media Stream Mixing or the preparation of virtual conference rooms or chat rooms. However, many IMS-providers realise the functions of the MRF via application servers without offering the MRF as such.

The PDF normally is a logical network-element only which ensures that in case of demand, real time-QoS is provided within the access network. "In case of demand" refers to the session which a client wants to setup. Typical examples for session-types with real time-QoS demand are "common" calls or video conferences. Here is another problem which relates to the IMS indirectly: The policing of the real time-QoS via PDF makes sense only if it is understood and provided by the providing access network. Of course the question is, how far this is true or if it is just "good luck"?

Watchful readers will have identified another problem, the determination of fees: IMS-operators cannot charge me for the provision of real time-services if the whole provider network cannot supply them in expected quality. A possible way out is the limit of the server-offer to handpicked access networks. But is a customer willing e.g. to be without a WLAN provider at home or to change to the IMS-operator WLAN offering??

With these open questions in mind we turn face to the protocol suite of the IMS a little more. The most important protocol in the area of IMS is the Session Initiation Protocol, called SIP. This SIP was standardised by the IETF in 1998/1999 and since that time many enhancements and changes have taken place. This phase is far away from being completed. SIP will be used for the communication between servers in the IMS as well as for the whole Session Management between IMS and the user device. Thinking about a session, you can visualise every possible task, going from Message-Transfers and Telephone-Calls to Video-Calls and interactive games. This flexibility is the biggest strength of SIP and its partner, the Session Description Protocol (SDP) which is used within SIP-messages for the definition and description of audio-, video und text-media. Furthermore the SIP serves the registration of terminal equipment in the IMS. In connection with IMS, registration does not mean anything else than relating the User-ID with an IP-address, so incoming session-demands can be routed to the correct IP-address. That the 3GPP has decided for the use of SIP and against the self-development at the definition of IMS, is somewhat irritating considering the structure of SIP and has rather political reasons. To understand this, you should know that SIP was defined by IETF as client-based-protocol with an extremely slim network. In other words: In the extreme case, SIP works even without network-server, although it uses these servers ideally only at the start of a session, particularly to locate the called partner. Afterwards, the servers disappear from the session. They are absolutely transparent for the user data (voice, video, audio) as well as during the session-release. For this type of server, a mid-class-computer which is installed anyplace, is sufficient. As well as many others, we carry on many servers like that for internal purposes like testing and training. In fact, this basic philosophy of SIP is controversial to the interests of any telecom provider. They also have to fulfil the demand for "Legal Interception" and other requirements of the legislature. But more obvious are commercial needs for possibilities of the operators, to be able to interrupt a session at anytime, e.g. if the credit is exhausted.

Therefore, SIP and SDP have been reengineered within 3GPP dramatically. Today the 3GPP-specified SIP and SDP for IMS vary extremely from the legacy-SIP/SDP and the IETF-SIP/SDP, especially through the fact that new and old ideal functions within the 3GPP-SIP/SDP are declared as "mandatory". Examples for such functions are the primary clarification of the QoS of a session or the presence of information elements and so called "Private Headers" in SIP-messages. Due to such discrepancies, the error-analysis in the first IMS-implementations are outmost complex, we are now preparing an additional SIP-signaling course because of this. Among other things, this course will deal with this problem in detail. More important in a long-term context is the question of interoperability of the expected implementations. A giant problem area may arise from today's method "Sell first and finalize the standard later". Without aggravating the truth we can assume that, e.g. SIP-terminal equipment from provider A or for TISPAN-based IMS-implementations are not or only restricted able to operate in IMS-implementations of 3GPP or 3GPP2. The superior panel of standardisation is lacking, common sense will need some time to win to agree on some joint set of mandatory functions. This problem may hinder a possible triumph of the IMS as for the GSM.

The next important protocol in the IMS-area is the real-time transport protocol, or RTP. Above all this protocol is placed between the transport-protocol UDP or TCP and the real codec, e.g. H.263 (video) or AMR (voice) to permit jitter-calculations. Alternatively to the RTP the SRTP can be used which allows additional functions authentication and data coding.

We want to point out especially that, despite their names, neither the RTP nor the SRTP provide real-time abilities. Actually, the necessary real-time resources have to be provided in some other way.

Although there are some other protocols in the IMS area, here the DIAMETER-protocol as a successor of RADIUS shall be mentioned. The name DIAMETER itself is program: DIAMETER= 2 x RADIUS.

The DIAMETER-protocol is a very efficient protocol for authentication, authorisation and fee (AAA), particularly because of its expandability.

One of most important accomplishments of 3GPP in the area of the IMS are these protocol expansions to enable the aforementioned functions.