Journal of Mobile Multimedia, Vol. 2, No.3 (2006) 189-224

PRACTICAL EXPERIENCES WITH AN IMS-AWARE LOCATION SERVICE ENABLER ON TOP OF AN EXPERIMENTAL OPEN SOURCE IMS CORE IMPLEMENTATION

PETER REICHL[1], SANDFORD BESSLER[1], JOACHIM FABINI[2], RUDOLF PAILER[4], ALEXANDER POROPATICH[5], NORBERT JORDAN[2], RAINER HUBER[4], HANNES WEISGRAB[1], CHRISTOPH BRANDNER[6], IVAN GOJMERAC[1], MICHAL RIES[3], FLORIAN WEGSCHEIDER[4]

[1]*Telecommunications Research Center Vienna (ftw.), Donaucitystr. 1, A-1220 Vienna, Austria*
[2]*IBK, Vienna University of Technology, Favoritenstr. 9/388, A-1040 Vienna, Austria*
[3]*INTHFT, Vienna University of Technology, Gusshausstr. 25-29/389, A-1040 Vienna, Austria*
[4]*Mobilkom Austria AG & Co KG, Obere Donaustr. 29, A-1020 Vienna, Austria*
[5]*Alcatel Austria AG, Scheydgasse 41, A-1210 Vienna, Austria*
[6]*Kapsch CarrierCom AG, Europlatz 5, A-1120 Vienna, Austria*

{reichl | bessler | weisgrab | gojmerac}@ftw.at; alexander.poropatich@alcatel.at;
{ joachim.fabini | norbert.jordan}@tuwien.ac.at;michal.ries@nt.tuwien.ac.at;
{r.pailer | r.huber | f.wegscheider }@mobilkom.at; christoph.brandner@kapsch.net

Received March 10, 2006
Revised June 26, 2006

The 3GPP IP Multimedia Subsystem (IMS) is currently expected to provide the basic architecture framework for the Next Generation Network which will bridge the traditional divide between circuit-switched and packet-switched networks and consolidate both sides into one single network for all services. Therefore, the imminent commercial roll-out of IMS will have immense impact both for the migration of the core network as well as the integration of future mobile services and applications. This paper presents an OpenSER-based experimental testbed which has been designed as a minimal standard-compliant IMS core network. We discuss major practical requirements and describe our implementation of this "IMS in a bottle" approach. Furthermore, we introduce a terminal-based native IMS location service enabler. We argue that physical location data can be regarded as a type of presence information and propose an architecture which reuses a large part of the IMS presence infra-structure by applying presence mechanisms, like notification handling, access control and privacy management, to location data. We demonstrate that the realization of this service can be integrated efficiently into the IMS core environment, and pre-sent initial evaluation results for the joint demonstrator. Finally, important current and future challenges including migration, interworking, charging, Quality-of-Service, identity management, security, and regulatory aspects, are discussed in detail, thus ending up with an up-to-date research agenda.

Key words: IP Multimedia Subsystem (IMS), Location Service Enabler, 3GPP

1 Introduction

On their way towards implementing the Next Generation Network (NGN), many mobile and fixed network operators have already started to migrate their telecommunication networks towards an All-IP infrastructure where voice loses its predominance and becomes just one amongst many services. The IP Multimedia Subsystem (IMS) [1][2], standardized by the 3rd Generation Partnership Project (3GPP) [3], is the most promising candidate for replacing legacy, voice-dedicated mobile networks

with an All-IP technology. The fundamental advantages of the IMS as opposed to traditional IP networks include guaranteed end-to-end Quality of Service (QoS) in the network and an infrastructure that enables the fast deployment and integration of new IP-based services and very flexible charging and billing while still maintaining compatibility with existing applications. Additionally, the new horizontal service architecture of IMS bridges the traditional divide between circuit-switched and packet-switched networks, consolidating both sides into one single network for all services, and thus will have an immense impact on the whole way future mobile business applications are designed, developed and deployed.

In the context of the imminent commercial roll-out of IMS, this paper presents an experimental test system which has been realized at the Telecommunications Research Center Vienna (ftw.) in the framework of the projects CAMPARI (Configuration, Architecture, Migration, Performance Analysis and Requirements of 3G IMS) and SIMS (Services in IMS). In both research projects, industrial participants include network providers as well as suppliers (i.e. Mobilkom Austria, Kapsch CarrierCom, and Alcatel Austria). Whereas CAMPARI investigates a "minimal-optimal" IMS core network configuration with respect to architecture and Quality-of-Service aspects, including measurement-based performance evaluations in a SER (SIP Express Router)-based IMS prototype, SIMS deals with interaction between applications and standardized or newly defined service enablers and develops dynamic service composition mechanisms that reduce significantly the time to market of future IMS applications. For further details on these two projects, please refer to [4] and [5] and references therein.

As our showcase we have chosen a Location Based Service (LBS), i.e. a service which requires information about the physical position of a user in order to provide 'added-value' to services in a 3rd Generation (3G) network. We focus on the Location Service (LS). a functional entity in the network enabling value-added services to query the current position of a user or to request a trigger when a specified area is entered or left Current standardization in the 3GPP [6] concentrates on LSs that use core network components to provide location data of a user's terminal. In contrast, we believe that the optimal source for location data is the user's terminal and show that a distributed, handset-based architecture scales better, is more accurate, efficient and cost effective. Our system design builds on the protocol, authorization, encryption and privacy mechanisms of the IMS presence enabler, but extends the existing specifications in order to support a distributed terminal-based LS.

The main contributions of this paper may thus be summarized as follows:

- We thoroughly investigate the requirements for a minimal-optimal fully IMS-compliant core testbed and discuss related design decisions.
- With the "IMS in a Bottle" approach we demonstrate the feasibility of an efficient 3GPP-compliant implementation of the IMS core network based on open source software.
- We define a novel location architecture based on application layer signalling which supports both terminal generated and network-calculated location information, where the control over the location information is either distributed (at the terminal) or centralized in the server entities, using the same mechanisms as for presence.

- We make the location service IMS-aware by using the SIP protocol and reduce the message overhead to a minimum through event triggering at the terminal. This allows for the first time to support applications that require advanced functionality like area location notification etc.
- Our Native IMS Location Service Enabler (NILS) is integrated into the IMS testbed, with IMS core infrastructure processing and routing SIP-based location signalling correctly.
- Finally, we discuss in great detail a number of important open issues for future research, thus giving a concise outline of the currently relevant research agenda.

The remainder of the paper is structured as follows: Section 2 provides a short survey on IMS core components and interfaces, and summarizes the basic requirements to the IMS core network prototype. In Section 3, we describe the implementation of our IMS testbed and present initial evaluation results. Section 4 introduces and evaluates the Native IMS Location Service Enabler (NILS) as our showcase whose realization on top of the IMS testbed prototype is a central joint result of the projects CAMPARI and SIMS. Then, based on the experiences collected in the course of both projects, in Section 5 we discuss in detail a couple of important and highly relevant open issues and thus determine a comprehensive agenda for current and future IMS research. Section 6 concludes the paper with a short summary and outlook; the appendix contains a list of acronyms.

2. IMS at a Glance

2.1. History and Standardization

The 3rd Generation Partnership Project (3GPP) [3] has been founded in 1998 by partners from Europe, America and Asia, with the objective to produce globally applicable standards for a 3rd Generation Mobile System. Additionally to its focus on UMTS, 3GPP inherited also the further development of 2G (GSM) and 2.5G (GPRS) technologies. Beginning with UMTS Release 5, the 3GPP started the standardization of a packet-switched Next Generation Network, i.e. the IP Multimedia Subsystem. The high-level requirements on IMS as summarized in TS 23.228 [7] are challenging: the IMS is supposed to provide end-to-end QoS, to fully replace current mobile circuit-switched voice networks without deteriorating voice quality, to be access-transparent, to enable flexible charging models, to enforce horizontal service architectures and to build a flexible infrastructure for cost-effective service and application deployment. Whereas today's mobile operators are required to install and maintain two distinct network infrastructures, a circuit-switched (CS) network for voice transmission and a packet-switched (PS) network primarily used for data, the horizontal service architecture of IMS provides the middleware for enabling rapid development of new services and allows to consolidate the infrastructure by deploying only one network for all services.

Note that, while pre-IMS standards focused on services for mobile and fixed networks, the 3GPP standardizes *service capabilities* instead and leaves the specific service realization to the implementer. There are several services that can contribute to the success of IMS, and some of them, like Location [6] or Presence [8], are explicitly supported by the IMS core architecture. In addition, TS 23.228 [7] mandates IMS support for the usage of non-IMS-specific Internet applications and thus increases drastically the number of applications that IMS users will be able to access from the very beginning (subject of course to the existence of QoS guarantees for the Internet applications).

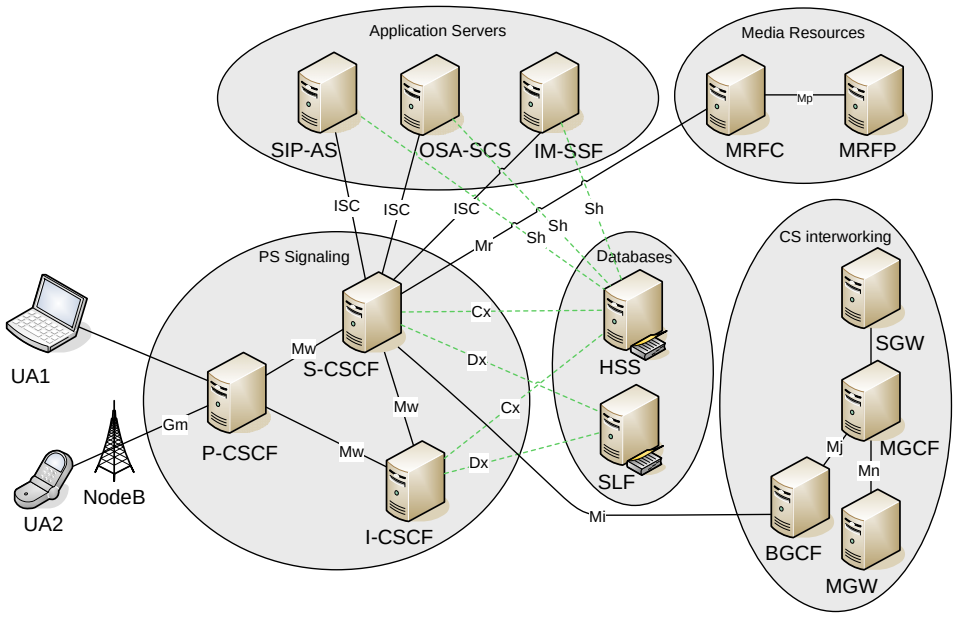

Figure 1: IMS Core Network Components and Interfaces
(SIP interfaces with solid lines, DIAMETER interfaces with dashed lines)

2.2 IMS Core Components

From a technical point of view, the 3GPP IMS is an IP-based overlay network that spans access and core network and provides for IP-based signalling re-using IETF protocols like SIP [9] and Diameter [10], both of which are enhanced by specific IMS extensions. The initial focus of the IMS specification was on UMTS and GERAN networks, but the concept as such is access-agnostic. Hence IMS can use also WiMAX, WLAN or fixed networks in the access providing guaranteed end-to-end QoS to customers. For network operators, IMS offers extremely flexible online and offline charging mechanisms as well as standardized interfaces towards application servers.

It is important to note that many of the features which make SIP an acceptable choice for commercial telecom systems are standardized by the IETF as (optional) extensions to SIP, whereas numerous such extensions are mandatory according to the 3GPP standardization and thus must be implemented by any 3GPP-compliant IMS system. Therefore, *IMS is more than SIP* in contrast to much of what is currently advertised by commercial IMS suppliers.

Figure 1 shows the main signalling components of a typical IMS network and their interfaces. The core IMS network uses three main component classes: the Call-Session Control Functions (CSCF) which are enhanced SIP proxies, Application Servers (AS) which implement services in the IMS network, and database components like the Home Subscriber Server (HSS). In addition, the IMS standardizes components for resource control and provisioning, e.g. Media Resource Function Controller (MRFC) and Media Resource Function Processor (MRFP), as well as components for interfacing with legacy CS voice networks like the PSTN, e.g. Breakout Gateway Control Function (BGCF), Media Gateway Control Function (MGCF), Media Gateway (MGW), and Signalling Gateway (SGW). In the remainder of the paper, we will however neglect interworking with circuit-switched networks and focus on packet-switched networking instead.

Call-Session Control Functions

IMS defines three types of CSCFs with exactly defined roles (a novel fourth type of CSCF, the Emergency CSCF (E-CSCF), is discussed later in Section 5.5):

- *Proxy CSCF:* The P-CSCF is the first point of contact to the User Equipment (UE) in terms of SIP signalling. The UE discovers its P-CSCF during IMS registration and stays with it during the lifetime of the registration. The P-CSCF acts as inbound and outbound proxy to the UE and as SIP security border gateway that protects the mobile operator's domain, e.g. by verifying the correctness and integrity of incoming SIP messages, establishing Security Associations (IPSec) with the UE, and compressing/decompressing the signalling over the Access Network (SigComp). Like many other IMS components, the P-CSCF generates charging records and forwards them to the Charging Collection Function (CCF).
- *Interrogating CSCF:* The I-CSCF is also a SIP proxy that is located at the edge of an administrative domain. It is registered with the DNS and acts as contact for IMS signalling messages coming from another IMS domain. The I-CSCF routes SIP requests to the destination within the administrative domain using routing information retrieved from the HSS and/or SLF through the Diameter interfaces. Typically, the I-CSCF assigns a S-CSCF to the user at registration time based on the user's HSS profile (e.g., support for subscribed services).
- *Serving CSCF:* The S-CSCF is a combined SIP Registrar and SIP Proxy. It is the „brain" of the IMS with a role similar to the MSC in mobile CS networks. After the I-CSCF has assigned a S-CSCF to a user, the S-CSCF authenticates and registers the IMS user and stores the binding between the user's address of record and his SIP contact. Then the S-CSCF is responsible for handling/routing all signalling traffic for the respective user based on the user's profile, specifically the Initial Filter Criteria (IFC), as downloaded from the HSS.

Home Subscriber Server (HSS)

The HSS is the main database component of IMS corresponding to the Home Location Register (HLR) in GSM/GPRS networks. It stores user profiles which contain all the user-related data required for IMS operation, e.g. public and private identities, subscribed services, shared secrets (corresponding to the data stored on the user's ISIM/USIM application), Initial Filter Criteria, etc. The HSS implements two Diameter-based interfaces: the Cx interface for CSCFs and the Sh interface for AS, enabling them to read and write HSS-stored user profiles.

Application Servers

Application Servers (see Fig. 1) are also SIP-enabled and host IMS service enablers, service components or even the applications themselves. In our showcase we envision that the location service runs on such a native SIP application server and exposes a web service (ParlayX) interface to internal applications or to third party service providers.

Example: IMS Registration Scenario

The interplay between the core components of IMS is demonstrated with the IMS registration procedure. To start with, the UE registers with the IMS network by sending its SIP registration request to the Proxy CSCF, which was assigned to the UE by the network. The P-CSCF forwards all UE mes-

sages to the Interrogating CSCF which serves the user's home domain. The I-CSCF contacts the HSS, selects a Serving CSCF based on the service subscriptions of the user, and forwards the registration message to the selected S-CSCF. The S-CSCF is the service access point and service dispatcher within the IMS network. It authenticates the user and registers him with the IMS network. Depending on the user profile, the S-CSCF may redirect calls from or to a specific user to one or more Application Servers in order to implement specific services requested by the user.

2.3 Requirements and Design Decisions for a "Minimal-Optimal" IMS Testbed

With the imminent commercial roll-out of IMS, network providers and suppliers are currently facing the challenge of migrating their current infrastructure towards this NGN architecture in a smooth and efficient way. In this context, the research project CAMPARI experiments with an open source-based "minimal-optimal" IMS configuration in order to determine the tradeoff between full standard compliance and practicability of the solutions. After a thorough investigation of the architecture options from a strictly pragmatic point of view, a test system emulating a minimal IMS has been set up and is currently used for measurement-based performance evaluations. Additionally, the project focuses on migration issues including IPv4/v6, the integration of IMS-specific billing options and security features and finally the identification, investigation and discussion of further IMS aspects which are not yet fully covered by the standards.

In order to deal with the complexity of an IMS core implementation, we have decided to avoid building our testbed from the scratch and have instead re-used existing open source software running on commercial off-the-shelf hardware. The following bullet points summarize our requirements for the testbed prototype and discuss related design decisions:

- *Basic Scenario:* The primary goal of the CAMPARI testbed was to signal an IMS-compatible basic call between two IMS/SIP UEs and to connect the media stream. The realization of this scenario requires the implementation of P-CSCF, I-CSCF, S-CSCF, DNS and a reduced HSS.
- *Access network emulation:* the complexity of a dedicated IMS access test network is out of the project scope, instead the existing Linux Network Emulator *Netem* [11] is extended for mimicking the behaviour of wireless networks (e.g. 2.5G (GPRS, EDGE) 3G (UMTS, HSDPA) or WLAN (802.11b, 802.11g) as closely as possible.
- *HSS subset emulation:* the planned performance tests require only two subsets of the Cx interface: Authentication (Diameter AVPs) and download of the Initial Filter Criteria (IFC) to the S-CSCF.
- *Charging interface:* we implement offline charging from the very beginning, with the option of later integration of other charging mechanisms like online charging and event-based charging.
- *Rudimentary IMS services*: the focus of CAMPARI is not on services but on the performance tuning of the core IMS. Nonetheless, CAMPARI integrates SIP ASs running a minimum service, e.g. Ping, SIP Presence etc., in order to perform IMS network optimization.
- *IPv4 and IPv6 support:* 3GPP R5/6 mandate IPv6 as the network-layer protocol and mention IPv4 only as an option. This conflicts with the current IPv4-dominated infrastructure. One focus of CAMPARI is to measure performance and propose migration paths for operators and vendors from an IPv4-only infrastructure towards mixed IPv4-IPv6, or IPv6-only networks.

- *Third-party IMS component integration:* While the implemented CAMPARI testbed is of limited functionality, a dedicated focus of the project is to enable integration of third-party IMS components. The currently planned interoperability includes a.o. third-party IMS domains, Charging Collection Function (CCF), Application Server (Presence), HSS, CS Gateways towards legacy circuit switched networks, dedicated user equipment etc.
- *Postpone IMS QoS and failover framework:* We replace the IMS QoS reservation framework, i.e. PDF, GGSN and SBLP, by a manually adjustable delay emulator. Registration for event notification is also considered not relevant for the main signalling path and has been postponed.
- *Postpone interworking with legacy networks:* the first testbed release focuses on PS topics only, gateways towards CS networks will be integrated by means of third-party vendor equipment.

Additionally, the CAMPARI testbed is designed to support the following performance tests:

- *Basic IMS performance:* comparison of Visited-GGSN vs. Home-GGSN scenario and of SIP/IMS signalling over TCP vs. UDP plus determination of bottlenecks and optimization proposals.
- *IPv4 vs. IPv6:* test configurations of IPv4-only, IPv6-only, IPv4-IPv6 mixed, interworking.
- *Application Servers*: dedicated AS vs. AS co-located with the S-CSCF.

Based on these requirements, we have implemented an experimental IMS core network as described in the next chapter.

3. Implementation and Initial Evaluation of the IMS Core Network

3.1. Testbed Version 1: Distributed IMS Core Network Prototype

The architecture of our IMS core network prototype is sketched in Figure 2 and contains a P-CSCF, a S-CSCF, an I-CSCF, an AS, the CCF, a scaled-down HSS, and two WAN emulators WE1 and WE2. The DNS implements SIP DNS resolution according to [12], the CCF collects charging data records (CDR) from all IMS components via the Diameter Rf interface. The HSS implements only the basic Cx interface functionality for authentication and Initial Filter Criteria download.

Two IMS-enabled terminals (UE1 and UE2) are located in the visited network, connected to distinct subnetworks in order to force the media traffic between these two IMS clients to be routed via WE1. The delay of packets exchanged between UE1 and UE2 equals the sum of the respective access network delays, whereas the IMS signalling traffic directed to the core network (UE1 to IMS, UE2 to IMS) suffers only from a single access network delay in one direction. The third IMS terminal shown in Figure 2 is a notebook connected via a Vodafone 3G UMTS/HSDPA card to the public Internet. An IPSec association is established between the laptop and the P-CSCF, enabling the mobile device to exchange signalling information with the IMS core and establish connections with the other IMS terminals in the network.

The testbed has been designed with respect to the two distinct IMS topologies defined by 3GPP Technical Specifications, i.e. the Home-GGSN scenario and the Visited-GGSN scenario. In the *Home-GGSN scenario*, the GGSN and the P-CSCF are both located in the user's home network, and packet data traffic is tunnelled from the visited network's SGSN via GTP to the home GGSN. In the *Visited-GGSN scenario*, the GGSN in the visited network terminates the user's packet data traffic.

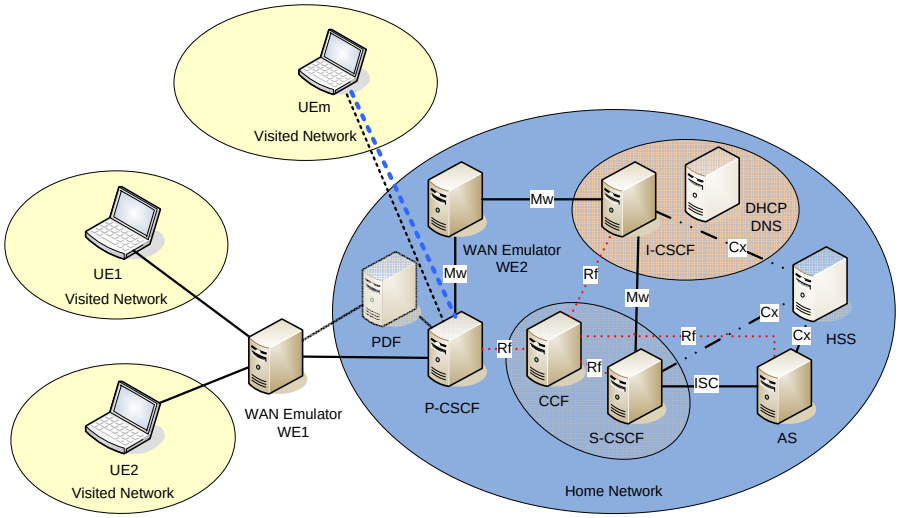

Figure 2: The Testbed Architecture

Our testbed can be configured for both cases simply by reconfiguring the two WAN emulators WE1 and WE2, i.e. without any changes in wiring. In the Visited-GGSN scenario, WE1 emulates only the access network delay, whereas in the Home-GGSN scenario, WE1 is configured to add the interoperator delay (e.g., Australia to Europe) to the access network delay. This differentiation is important for performance measurements, as in the Home-GGSN case not only the signalling, but also the media traffic passes the home GGSN and thus suffers from delay. WE2 can delay the signalling traffic between P-CSCF and I/S-CSCF. In the Visited-GGSN case, the P-CSCF is located in the visited network; it is thus WE2's responsibility to implement the interoperator delay. In the Home-GGSN scenario, the interoperator delay is implemented by WE1, so WE2 is transparent to IP traffic.

The rest of this subsection is devoted to important implementation details. As, in principle, an x-CSCF is a SIP-proxy (the S-CSCF acts additionally as a SIP-registrar), we have started with a thorough evaluation of the most prominent open-source SIP-proxies, before we came to the conclusion that the SIP Express Router (SER) [13] is the best choice for building an x-CSCF. SER is a rather fast, yet flexible, SIP-proxy which has been written in C and is easily configurable and already well-established. It can be configured by a C-like configuration file (but only with reduced functional range) and is available and supported for a variety of architectures, e.g. Linux/i386, Linux/armv4l, FreeBSD/i386, OpenBSD/i386, Solaris/sparc64 and NetBSD/sparc64. As operating system, SuSE Linux 9.2 has been chosen in accordance with the proxy software. On the S-CSCF, the database software "mysql" has been used for storing the user database and additional AVP data. The latter is an extension (module) for SER and is inalienable for storing data on a per-session basis, as e.g. necessary for adding the "Route"-header-field, etc. For easier access and maintenance, the software "phpMyAdmin" running on top of the php extension of the web-server Apache2 has been installed as a simple-to-use frontend with a nice GUI for the mysql database.

Originally, we started with SER v.0.8.14, but the lack of functionality for manipulating the AVPs forced us to move to openSER. The final testbed setup for the time being is based on the developer version of openSER as of Feb 16th, 2006. As SIP-(IMS-)clients, we use X-Lite Rel. 1105x on Windows and kphone v4.2 on Linux, which has been modified to support some of the IMS extensions like the "Service-Route" and some of the P-header extensions. Since X-Lite cannot be modified, the P-CSCF has been configured to take over this task if (and only if) the client does not support the necessary IMS extensions, thus allowing to use also non-IMS compliant clients (like X-Lite) in our distributed IMS testbed.

3.2. Testbed Version 2: Virtual Testbed ("IMS in a Bottle")

In parallel to the distributed testbed implementation which assigns a physical machine to each of the IMS components, a virtual testbed has been set up on only one physical x86 based Desktop PC. The advantages of this solution are lower costs, lower space requirements, lower energy consumption, and higher flexibility. The virtual environment is provided by the VMware Workstation 5.5.x software suite. In our scenario, the virtual machines ("Guests") are running SuSE Linux 9.2. The physical ("Host") operating system is Windows XP SP 2, but this irrelevant because any operating system VMware runs on could be used as host operating system as well. Figure 3 provides a sketch of the physical layering of our virtual testbed.

In terms of network connectivity, VMware offers 10 separated, virtual networks ("VMnets"), which can be imagined as an Ethernet connected via hubs. Each Guest and the Host is able to connect to one or more of those by utilizing one or more virtual network adapters. Additionally, any Guest can be connected directly to the physical network(s) to which the Host machine is connected to (this mode is called "bridging"). Thus any possible interconnection of virtual and physical networks is possible, which means that the virtual machines integrate smoothly into the existing physical environment. A VMnet behaves as if all hosts were connected via a hub. This means that all packets of the respective VMnet can be traced by any host, which is very useful because the message-flow between all machines connected to the same VMnet can be traced by a single machine without the need to correlate traces from distinct machines.

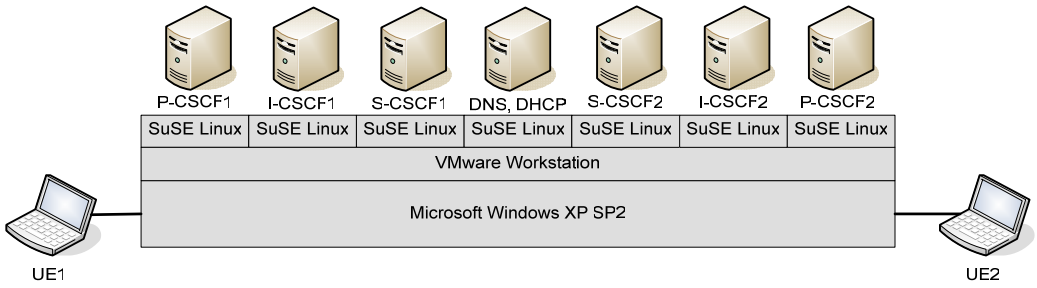

Figure 3: Physical Layering of the "IMS-in-a-Bottle" Testbed

Concerning performance, VMware offers good transmission rates of more than 200 Mbps with the cost of lots of processing power. Thus, our Host machine is equipped with an AMD 4400+ dual core CPU. Even though VMware could swap out some of the memory the virtual machine uses, this

would reduce performance dramatically. Thus we configured VMware to fit all Guest memory into the physical memory of the Host and equipped the machine with 4GB RAM, the maximum offered by affordable (=non-server) motherboards. Another critical performance factor is the HDD. It is true that space is not a major concern here because VMware uses technology which physically allocates only those parts of a virtual disc, which are actually used, and thus usually the (physical) image of the virtual disc is much smaller, but concurrent access has dramatically negative effect on the performance. Therefore, it is generally better to equip the machine with more small HDDs (which could also be rather slow) than with a few big and fast ones. Only if it is sure in advance that the Guest system will hardly use the HDD, two or three Guests could use one HDD. Nevertheless booting those systems could take a considerable amount of time if two or more Guests sharing the same HDD are being booted at the same time. Since our guests do hardly use the HDD, the machine is equipped with 5 HDDs such that a maximum of 3 Guests share one physical drive.

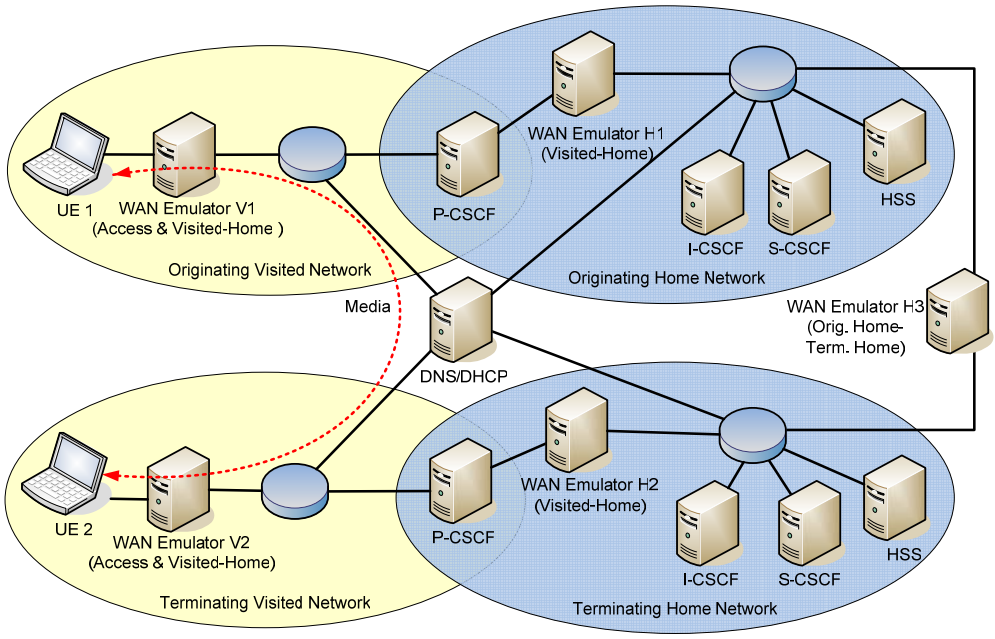

Figure 4: Components of the "IMS-in-a-Bottle" Testbed

Of course, there are some disadvantages of the virtual solution, most obviously the limited available performance and bandwidth. Furthermore, the clocks are rather inaccurate, thus timing measurements cannot be based on the clock of a virtual machine. Therefore, we connect the Host additionally to the VMnet that shall be traced and capture the traffic on the physical machine. Even though the clocks of the virtual machines disperse, the response time of the machines is good.

With this configuration, our “bottle” (Host-machine) is able to implement the whole 2(4)-operator scenario depicted in Figure 4. By tweaking the WAN Emulators V1/H1 and V2/H2, respectively, it is possible to change from the “Home-GGSN” to the “Visited-GGSN” scenario. Due to the

symmetric architecture, we can exploit the "clone" feature of VMware which enables us to clone each of the machines in a very elegant way. The clone automatically gets a new System Id and thus new MAC-addresses, but of course the system configuration (IP - adress(es), machine- and domain names, etc.) need to be changed manually.

Summarizing briefly, the virtual environment is a good solution for research, development, analyzing (call-flows, etc.), and testing functionality. It is moderately good for timing analysis in mid- and low load scenarios, but cannot be used for measurements in high-load scenarios and for performance measurements. Nevertheless, "IMS in a bottle" has been a successful proof-of-concept activity resulting in an alternative and cheap playground for implementation work.

4. The Showcase: NILS – a Native IMS Location Service

As already mentioned in the introduction, one key advantage of the IMS is related to simplified development and deployment of new services. In order to verify this, we have decided to realize as our showcase a native IMS-aware Location Service that reuses a part of the IMS presence infrastructure.

The concept of presence comes from instant messaging systems, where a "watcher" can subscribe to notifications about the online status of other users which are called "presentities". Classical presence information is defined as the willingness of the presentity to communicate. In general, presence and geographical location have semantic differences: presence attributes as defined by the Internet Engineering Task Force's (IETF) Rich Presence (RPID [23]), e.g. location type, activity, sphere etc., consist of a small number of discrete values, whereas geographical location has a continuous range of values, limited only by the location accuracy. This requires a different handling of subscriptions: whereas for presence all watchers subscribe to changes that occur to a subset of attributes, for location each watcher application may define different criteria for triggering events. Hence, publishing and storing location information on servers like for normal presence information is not useful, and our architecture takes these differences into account.

However, location and presence have sufficient similarities to approve our approach to reuse the IMS presence architecture [24]. Our analysis of the IMS presence system shows that the requirements regarding access authorization, subscription and privacy for presence are identical to those for location. Moreover, location, like presence information, can be collected both from the mobile terminal and from the network, and, most importantly, this can be done using the SIMPLE protocol stack. Our system design builds on the interaction mechanisms and data representation used for IMS presence, but extends the existing specifications in order to support a distributed terminal-based Location Service as we believe that the optimal source for geographical location data is a user terminal which is equipped with a GPS module.

Location information can provide considerable value to information and communication services. On the other hand, users are concerned about revealing their position data to others, especially to untrusted third party applications. Furthermore, most countries have legal restrictions that regulate processing of personal data and the protection of privacy in electronic communications. It is of utmost importance that the users can control who gets access to their location data and that the transport in the network of such sensitive data is protected by strong security mechanisms.

The Geographical Location and Privacy (GEOPRIV) working group of the IETF [17][18] has investigated a number of problems related to the distribution of geographical information on the

Internet. The result of this work is a generic framework for the creation and distribution of location information on the Internet that enables confidentiality and policy directives, which are abstracted from the format of the location information.

The Open Mobile Alliance, a large consortium that specifies standards for mobile services, has recognised the importance of terminal generated location and has defined a Secure User Plane Location (SUPL) enabler [19]. The SUPL architecture, in contrast to our approach, keeps network server components in visited and home networks, it does not use SIP, therefore it has to implement methods to route requests, to keep session state, to handle terminating asynchronous events using WAP push.

Further related work includes [20] where complex location update strategies in the mobile terminal are proposed in order to realize scalable Location Based Services. In [21], the SIP event mechanism is used to transport location data, and [22] presents a framework and requirements for usage of SIP to convey user location information and consider cases where message routing by intermediaries is influenced by the location of the session initiator.

4.1 An Overview on 3G Presence and Location Enablers

The *Presence Enabler* in the IMS is based on the SIP Instant Messaging and Presence Leveraging Extensions (SIMPLE) [25], which in turn uses the SIP event system [26] and its subscribe-notify mechanism. The presence system of the IMS is specified in [27], the presence event package in [28]. Figure 6 shows all elements of the presence system as described in the following list:

- The *Watcher Application* can run on an application server in the core network or on a user terminal in the access network. The application subscribes to the presence of presentities and receives notifications whenever the presence changes. It can supply a notification filter [29][30] at subscription time to limit traffic or request complex notification behavior.
- The (optional) *Resource List Server* (RLS, [31][32]) relieves the watcher's user agent from subscribing to and managing notifications from all addresses on his contact list. Instead, the contact list is stored on the RLS and the Presence User Agent (PUA) subscribes to the contact list and receives bundled notifications. This reduces traffic on the air interface.
- The *Presence User Agent* sets policies for subscribers (XCAP [33][34] over http) and sends notifications (SIP message: publish) if its presence state changes. Since the PUA does not know about the subscriptions to its presence data, it has to publish every single presence change of the user independently of the actual need for it, even if there are no watchers at all.
- The *Presence Network Agent* supplies presence information coming from the network (e.g. cell based location information) by sending PIDF [35] or RPID [36] documents to the presence server.
- The *Presence External Agent* supplies presence information coming from external sources such as a calendar.
- The *Presence Server* (PS), also known as *Presence Agent* (PA), is the entity in the IMS core network responsible for the presence information of all subscribers to this network. The PS maintains presence information, sends notifications to watchers and is responsible for the authorization of all subscriptions.

- The *Watcher Presence Proxy* provides watcher related functions such as authentication of watchers. It is a combination of SIP proxies, i.e. P-CSCF and S-CSCF, in the watcher's network.
- The *Presentity Presence Proxy* provides presentity related functionality such as determining the correct PS. It is a combination of I-CFCF and S-CSCF in the presentity's network.

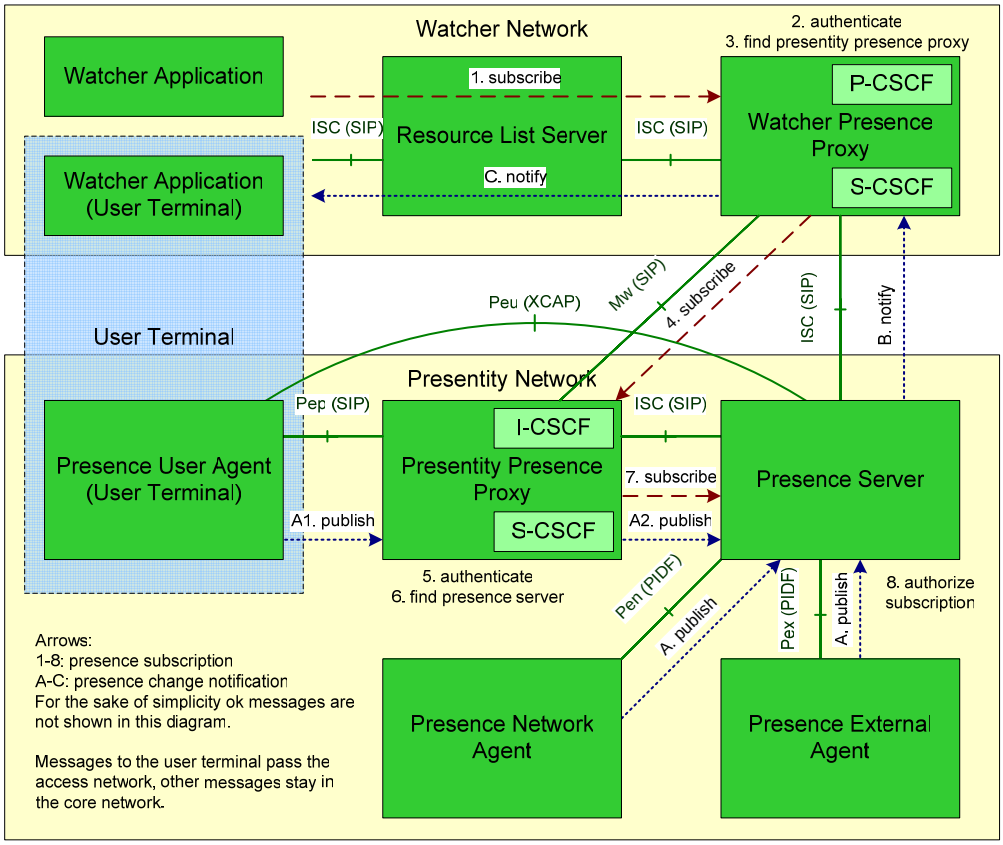

Figure 5: IMS Presence System Overview.

Figure 5 illustrates two message flows: a watcher subscribes for presence information (1 to 8) and a notification for new presence information (A to C). The authentication process (2, 5) involves the HSS (Home Subscriber Server) which is not shown in this diagram.

In order to provide location data to location applications, current 2G and 3G networks use a network-based LSE. 3GPP specification TS 22.071 [6] gives a general description of location services (LCS) and service requirements for 3G networks, and 3GPP TS 23.271 [37] specifies the mechanisms to support mobile location services for operators, subscribers and third party service providers. Interworking with the IMS has been introduced in version 6.7.0 for UMTS R6 in March 2004. The central component in the 3GPP LCS design is the Gateway Mobile Location Center (GMLC) that offers location services and communicates with the Location IMS Interworking Function (LIMS-IWF), see [41] for further details.

For determining the (geographic) location of the user equipment using radio signals of the wireless networks, the following network-based positioning methods are specified in the 3G documents:

- *Cell Coverage Based*. The cell ID is either known to the radio network or can be obtained by paging the terminal. The accuracy of this method depends on the cell size and is typically in the range of 300 m in urban areas.

Idle Period Downlink – Observed Time Difference Of Arrival (IPDL-OTDOA). This method measures the relative time of arrival of pilot signals from different base stations. At least three stations have to be visible to calculate a location. Network planning however optimizes a cellular network for available bandwidth which means reducing unnecessary overlapping of cells [38]. Thus the accuracy of the IPDL-OTDOA method is not satisfying in real networks (about 50m – 150m) and the technology requires a substantial up-front investment in the radio access network.

- *Network Assisted GPS (A-GPS).* The network sends assistance data to a GPS receiver in the terminal in order to speed up the position calculation and to provide service in areas where GPS signals are weak (e.g. inside buildings). Assistance data contains precise GPS satellite orbit and clock information, initial position and satellite selection. A-GPS offers good position accuracy (5 m) but requires a GPS assistance service and necessitates a medium up-front investment in the 3G network. Furthermore A-GPS enabled terminals are not available on the market yet.

4.2 Architecture of the Native IMS Location Service (NILS)

The NILS architecture is implemented within the presentity's home network and thus transparent to the watcher and his network. As already mentioned earlier, the proposed system architecture is based on the following two fundamental assumptions:

- *Geographical location is a special type of presence information that we call location presence data.* The access to both presence and location data has the same requirements regarding security and privacy. Therefore, this section will demonstrate how to reuse the well-specified presence system of the IETF (and thus also of the IMS), in particular the authorization, subscription and privacy mechanisms, in order to fit location presence data smoothly into the existing presence system which can be extended easily with the new location data type.
- *The optimal source for geographical location data is the user's terminal which is equipped with a GPS module*. Current network-based location methods do not offer satisfying accuracy, and enhancements would require expensive investments in the core network infrastructure. On the other hand, a terminal-based location enabler, such as a built-in GPS sensor or a Bluetooth GPS module, is a relatively inexpensive hardware feature and allows upgrading only those terminals whose users subscribe to or want to use location enabled services. It can be accessed in a peer-to-peer mode with minimal impact on network resources, thus granting scalability. Since a terminal-centric solution is totally independent of the underlying network technology or the network provider, it enables vertical handover (e.g. UMTS – WLAN) and works in roaming scenarios.

In order to create a terminal-based Location Service, we install a SIP Presence Edge Server (PES) at the GPS-enabled user terminal. A PES is a presence agent that is co-located with a presence user agent. Since the PUA manipulates the presence information, the PES can be aware of it ([28] sec-

tion 3). With the slight modifications described in the next sections, the PES, the PUA and the watcher UA are able to handle location presence data.

The standard source of location information in our architecture is the terminal supplied one, for example by a GPS receiver attached to the user mobile phone. For those users who do not have a GPS receiver or are not within GPS coverage (e.g. inside a building), our system connects to the network operator's GMLC to retrieve the location.

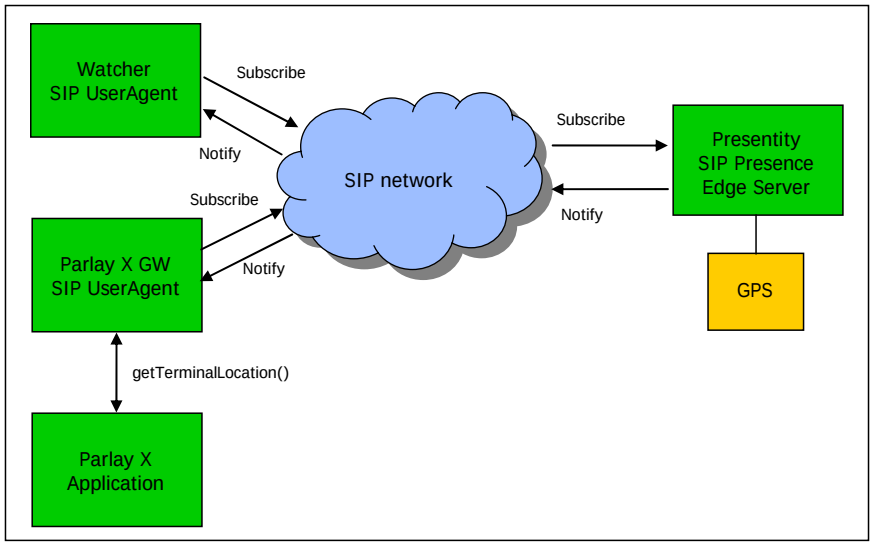

Figure 6: Peer-to-Peer System Architecture

The simplest form of the proposed system architecture is sketched in Figure 6, i.e. a *peer-to-peer location system.* The user's terminal hosts an edge presence server for location presence information. It allows 3rd party applications or other users to subscribe to location presence data notifications. A watcher application may use SIP messaging or interfaces like the ParlayX location API [39] to request presence location data. The client software in the user's terminal uses the J2ME Location API [40] to access the GPS module providing the physical location.

While it requires little infrastructure, the peer-to-peer architecture has several drawbacks: lots of notifications pass the radio network, the system does not support network-based location for terminals without a GPS receiver, the policy configuration done in current presence servers cannot be reused. As described in [41][42], these problems can be solved by the use of resource list servers, an aggregator component to switch between the peer-to-peer location described above and the network-based location and the reuse of presence servers.

Location Data and Notification Filters

To include location information in the PIDF document encapsulated in a SIP NOTIFY message, we propose the use of the Geography Markup Language (GML, [43]). Since both languages are XML based, GML elements can be integrated in PIDF documents easily.

Notification filters enable more complex notifications than pure location updates. The filter is sent within the *subscribe* message from the watcher to the presence edge server. For example the filter "(longitude TO centre BY distance) OR (latitude TO centre BY distance)" will send a notification

whenever the presentity enters a rectangle around the centre coordinates. This reduces traffic over the air interface (single notification instead of constant location updates) at the price of added complexity at the presentity's terminal.

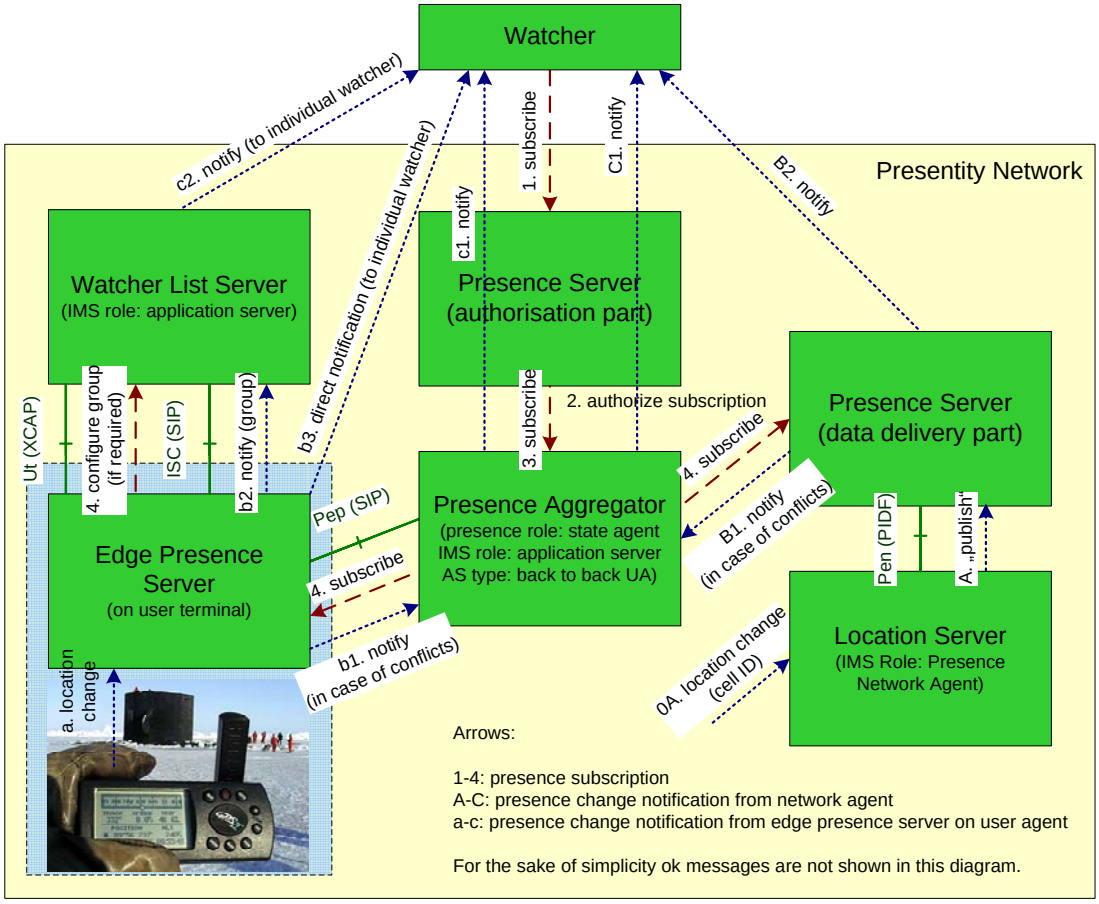

Figure 7: Full NILS System Architecture. (Photo of GPS Receiver taken from [44])

Watcher List Server

To reduce the bandwidth requirements of the edge presence server requires over the radio interface, we propose the Watcher List Server (WLS, see Figure 5). This element of the presence system manages lists of watchers subscribed to the same location events. The presence edge server at the user terminal uses the u_t interface to the WLS to configure these lists. Once the watcher list has been configured, any notification sent to it will reach all the watchers on the list (arrows b2 and c2 in Fig. 5).

The improvement achieved by a WLS is based on the assumption that the delivery conditions for location are the same for all the watchers. However, this is not true for all requests. For example one watcher might request periodic notifications while another wants to know, when the presentity reaches a specified location. In these cases direct notifications will be used (arrow b3 in Figure 5). While we distinguish between watcher list server function and resource list server (on the watcher side [31])[32]), we expect them to be implemented by the same software and run on the same ma-

chine. The watcher list server is transparent to the watcher and to the rest of the presence system. The only entity affected is the edge presence server on the terminal.

Presence Aggregator

The main remaining problems are the support of network based location for legacy terminals and the consolidation of information if both options (GPS and network) are available. Note that the presentity is the only entity able to resolve these conflicts ([28] section 6.9). This problem is resolved by the *Presence Aggregato*r (Figure 7). It intercepts all subscription messages on the presentity side of the IMS. Based on a user data base or on subscription policies, it decides how to locate the presentity. If the location is only available in the network, the subscription is forwarded to the presence server. If the user has an active GPS receiver, the subscription will be sent to the terminal. If both options are available the aggregator can check the data for reliability, i.e. whether the GPS coordinates are within the cell. Arrows B1 and b1 in Figure 7 show the (potentially conflicting) notifications to the aggregator, which forwards only one of them (C1 OR c1) to the watcher.

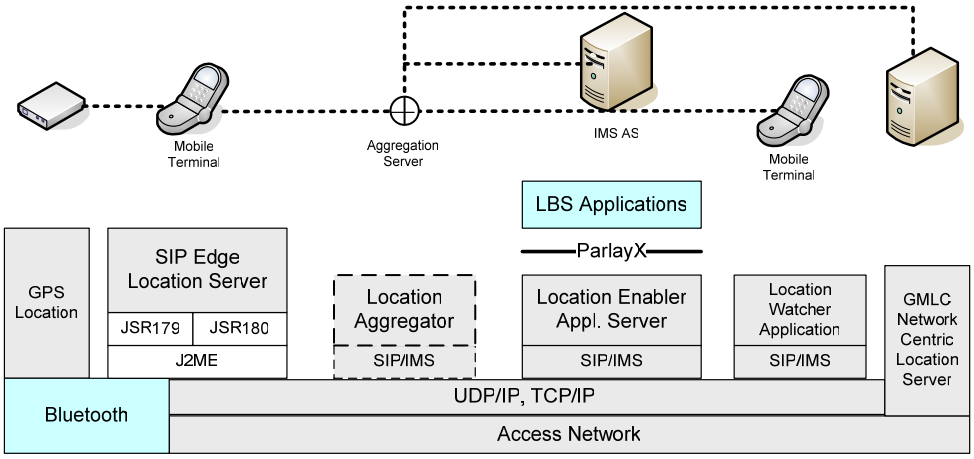

Figure 8: Location Enabler Software Architecture

Like the WLS, the aggregator is transparent to the watcher. It is also transparent to the presentity and all proxies. The presence server authorizes all subscriptions. We expect the aggregation function to be implemented within the presence server in the core network, especially since the aggregation task takes place after authorization, but before data can be fetched from any presence server. The resulting architecture relieves the terminal from publishing its location independently of actual watchers. Still, it is compatible with the IMS. It supports both terminal generated and server calculated location keeping the traffic in the access network low. Note that our architecture can be applied to any presence data originating at the user terminal

Routing and Addressing

Location presence data shows some essential differences from other presence data. First, location data is semantically different from classic presence data due to the almost continuous state space.

Second, location data originates from different sources than normal presence state. As location data is calculated in the presentity's terminal it requires a distributed approach. We therefore propose to separate subscriptions and notifications for location presence data from classic presence by introducing a new *locpres:* URI-scheme. The intended usage of the locpres: URI follows closely the usage of the pres: URI. However, it allows routing a subscription for presence location data to the presentity's terminal by addressing a subscription request to the locpres: URI of the target.

4.3. Implementation and Evaluation of the Demonstrator

The software layer architecture of the NILS enabler is depicted in Figure 8. With the "SIP proxy" we abstract the whole IMS transport overlay for the sake of simplicity. A simplified location watcher can be implemented in a user mobile phone as well, the area notification functionality however is best demonstrated by a LBS application running on an application server. We selected for this purpose the BEA WebLogic SIP application server. The application uses a ParlayX location interface and is notified when the user enters or leaves an area specified by its center coordinate and the radius. The presentity user is notified and eventually accepts the incoming subscription (assuming she is interested in the application).

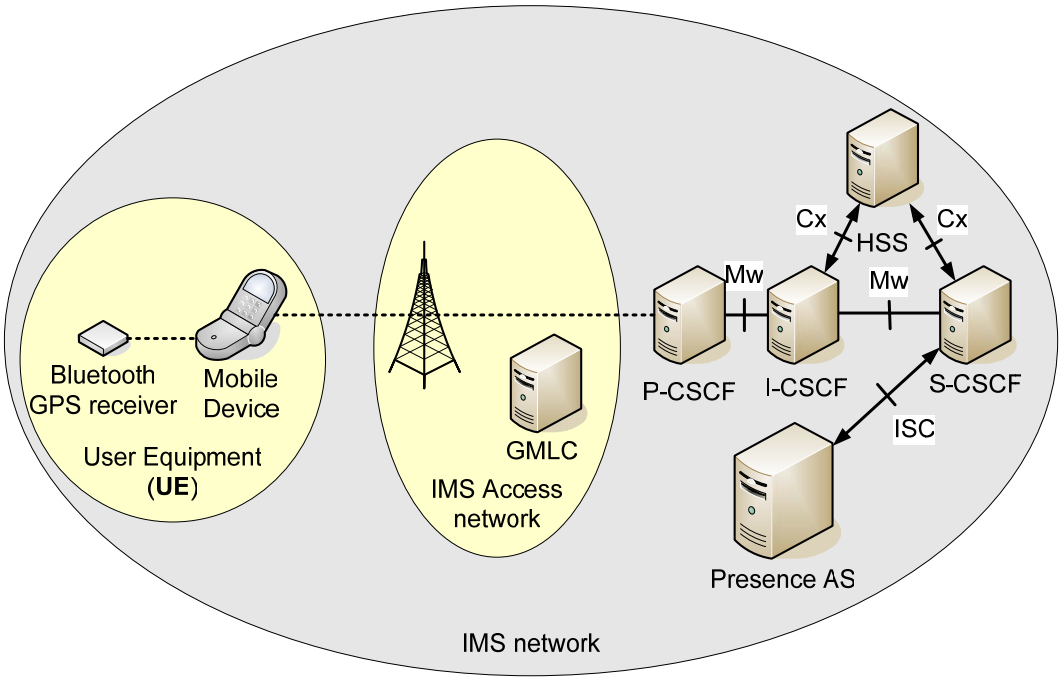

Figure 9: Final Joint Demonstrator Setup

Finally, Figure 9 depicts the joint demonstrator of NILS on top of the IMS prototype testbed.

Analysis of the 3GPP LSE Architecture

Our analysis of the 3GPP LSE architecture identifies several shortcomings:

- Complex request routing is required between GMLC and LIMS-IWF components in order to resolve the user's MSISDN and to accommodate roaming scenarios. In an IMS-enabled system this routing should be replaced by SIP message routing with its positive implications regarding performance, network management and security. Our alternative proposal is therefore to send a location request to the user's locpres: URI directly using plain SIP message routing instead of introducing complex interworking functions.
- Existing network-based positioning methods are either insufficient regarding accuracy or require up-front investments in the core network. Accessing a terminal-based GPS receiver as a source for location data instead is a network technology independent solution that offers satisfactory accuracy while requiring only investments in those customers who want to use a location service.
- The privacy requirements and mechanisms proposed by 3GPP standardization for Location Services are similar to those of presence systems [6][18]. Beyond the specified functions, a real world location service system requires components that allow provisioning, user access and user control of privacy parameters. It seems reasonable to re-use the IMS presence infrastructure for subscription, authorization and privacy management of LBS.
- A network-based positioning method is forced to implement triggered location updates through polling. In order to preserve battery power mobile terminals tend to be in an idle state most of the time. Hence position changes are not visible to the network. We propose to implement trigger logic in the terminal so that necessary location updates are kept to an absolute minimum and scarce resources in the radio access network are used in the most economic way.

Network Traffic over the Air Interface

In order to determine the performance gains of our architecture, we calculated the traffic over the radio interface for the following scenario: we assume that the watcher also supplies location information (Watcher plus Presence Edge Server in pure peer to peer mode, as in Figure 6) and that the system includes five presentities (few people are allowed to see a user's location). One notification is sent each ten seconds (the maximum allowed by the IMS presence system is 1/5 Hz), while notifications are about 1KB in length, OK messages about 300 bytes. Finally, SIP compression to 1/3 of the message size is assumed.

This scenario results in symmetric traffic (downlink: notifications from presentities, OKs from watchers; uplink vice versa) of about 3,2 kbit/s or two messages per second. Due to the high frequency of notifications resulting from the highly dynamic data, notification traffic outweighs subscription traffic by far. The traffic resulting from only five watchers is already more than the phone battery can sustain for extended periods of time. Since there is only one attribute in the presence document (the location), partial notifications do not save any traffic. Event throttling can reduce traffic substantially, but at the cost of reduced location precision (e.g. once every five minutes, instead of constantly).

The NILS architecture contains two main features for reducing this traffic:

- First, the Resource List Server specified by the IETF and the 3GPP [31][32] reduces the traffic on the watcher side: downlink messages are reduced in number by bundling and throttling noti-

fications from several presentities, this in turn results in fewer OK messages on the uplink channel [45].RLS reduces uplink traffic (OK messages) by a factor equal to the number of presentities.

- Second, the Watcher List Server introduced in this paper removes the need to send notifications to every single watcher. Instead, the WLS in the core network receives a single notification and forwards it based on the watcher list. Hence the notification traffic on the presentity side is reduced by a factor equal to the number of watchers.

In our case, the resulting traffic will be about 1 kbit/s at the full rate of notifications, or less than a third of our starting level. The rate of notifications (once every ten seconds) is still high and can be sustained for only short periods. The presence aggregator further reduces notification traffic from the Presence Edge Server, if it bundles subscriptions. In this case it works similar to the Watcher List Server.

Triggered Location Updates versus polling by a Network-Based Location enabler

A second simulation based performance analysis of NILS was done in order to investigate the quantitative advantage of triggered location updates of a terminal based location enabler compared to the polling of a network based location enabler. We have based our simulation of node movements on the Random Waypoint Model that was first introduced in [46] and is widely used as a mobility model to compare the performance of various mobile network protocols. [47] shows that a simulation based on the Random Waypoint Model will only reach a steady state if a minimum speed is defined.

In the Random Waypoint Model (RWM) a node moves from its current location to a new location by randomly choosing a new location in the simulation area and a speed that is uniformly distributed between [*minspeed*, *maxspeed*]. The RWM includes randomly chosen pause times between changes in direction and/or speed.

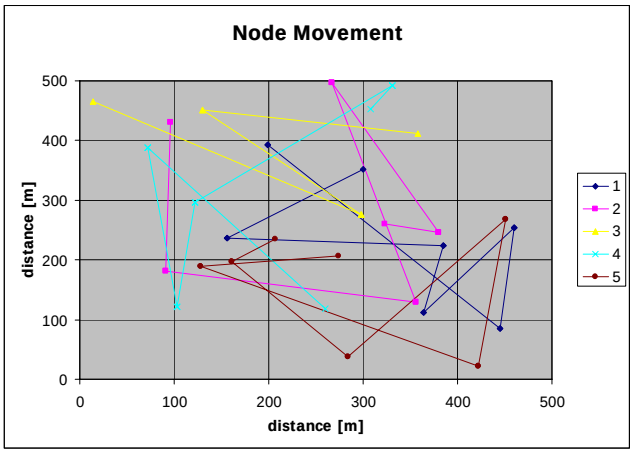

Figure 10: Node movement in the random Waypoint model

We have used the BonnMotion simulation software, a mobility scenario generation and analysis tool [48]. Our simulation targets a group of 100 pedestrians that move in a square with an edge length of 500 m. The node speed is chosen between [0.2, 1.5] m/sec. The maximum pause time is 60 sec. The

simulation was run for 7200 sec in order to reach a steady state and then measurements were taken for 1800 sec. Figure 11 exemplarily shows the movement traces of 5 nodes.We chose the simulation set-up in a way to estimate a lower bound on the reduction of necessary location update messages in a triggered scheme compared to a polling scheme. Our simulation observes nodes that are nearly constantly moving. Permanent movement is the worst case scenario for the number of triggered location updates, because location update messages have to be sent in regular intervals. In the real world, people do not move constantly but rather tend to stay in one place for long periods of time (e.g. commuters between home and office). As a triggered location update is not fired when a node does not move, it can be expected that savings in the number of necessary location update messages and thus in scarce wireless network capacity is much higher in real world movement patterns.

The node movement traces that come out of the RWM simulation were analyzed in two ways:

- *Count the number of triggered location updates that are necessary to report location changes for a given spatial resolution in meters.* The actual position was checked every second and if the distance of the actual position to the last reported position was greater than the given spatial resolution, a location updated was triggered. This implies that the maximum error in the location update is the distance that the node can move at maximum speed in one second. For a high spatial resolution of 10m and a maximum speed of 1.5 m/sec, the error results in 15% at most.
- *Compute the optimal polling time and the resulting number of location requests*. The optimal polling time is the interval between sending location requests to the network which results in a minimal absolute spatial error of all reported positions. Absolute spatial error indicates a deviation from the required spatial resolution in both directions – meaning that the polling happens either too often or too rarely. The optimal polling time is a function of the required spatial resolution and of the movement pattern and is therefore difficult to find for real applications. Even when choosing the optimal polling time, the average spatial error of all reported positions was about 40% compared to the required spatial resolution.

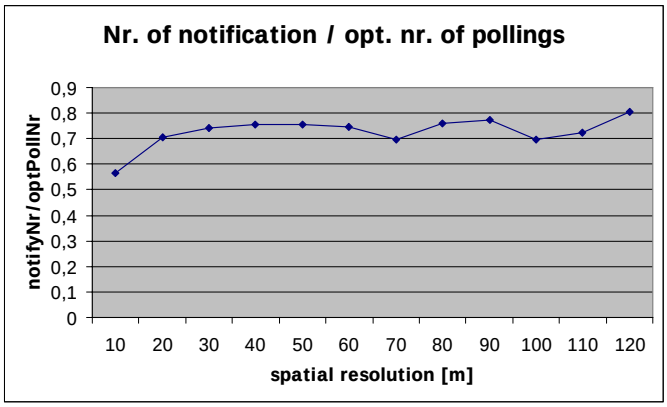

Figure 11: Triggered location updates versus location requests at optimal polling rate

We have finally compared the number of triggered location updates to the number of location requests at the optimal polling rate for different values of the spatial resolution (see Figure 11). This comparison shows that a triggered location update scheme, like that we have realized with the NILS, can save

at least 25% of messages, compared to a polling scheme working at the optimal polling time. Furthermore, the node location accuracy in a triggered scheme is at least double the accuracy of the polling scheme.

5. Experiences and Challenges

The massive body of project work described so far has not only resulted in realizing both the experimental IMS core network and NILS, but beyond these specific tasks has led of course also to a very detailed understanding of open issues and unsolved problems in the general area. Therefore, the goal of this chapter is to share some of these insights with the reader and thus to eventually outline a comprehensive agenda of important topics for current and future research on IMS in general.

5.1. Charging and the Evolution of IMS Billing Systems

One of the main benefits of IMS for both customers and operators is an extremely flexible charging architecture. While traditional charging systems are restricted to flat rate, per-time- or per-volume accounting, the IMS integrates Application Servers and Services within its charging architecture, thus increasing the complexity of the IMS charging system but at the same time providing satisfying transparency to the user.

The principles of IMS charging are standardized in [49]. IMS distinguishes between two fundamental charging approaches. *Offline charging*, often used for post-paid tariffs, requires a contract between the IMS operator and the customer prior to network usage. The operator later issues a bill to the customer, charging her for the services used during the last billing period. In technical terms, offline charging uses the dedicated Diameter-based Rf interface towards the CCF. Based on the IMS Charging Identifier (ICID), charging records are generated by IMS network nodes (e.g. BGCF, MGCF, MRFC, SIP AS, P-CSCF, I-CSCF and S-CSCF), sent to the CCF and correlated according to the specific service and billing plan. *Online charging* is required e.g. for pre-paid tariffs, where the user pays in advance a specific amount (“credit”) to the operator. Before consumption of a service, the operator verifies in real-time if the user’s account has sufficient credit and decides whether to grant access to the service or not. For services charged per time (e.g. voice calls), the IMS periodically checks the user’s credit and disconnects if there is insufficient credit for the next charging interval (e.g., 1 sec). Online charging uses its own charging framework where MRFC, S-CSCF and SIP-AS are connected via the Diameter-based Ro interface to the Online Charging System (OCS).

Note that orthogonal to the distinction between on- and offline charging, IMS offers different charging mechanisms: *Session-based charging* (per-time, per-volume), *Event-based charging* (e.g., per successful buddy localization or per presence-notification) and *Flow-based charging* (depending on the media using a specific bearer) [50].

The key to the flexibility of IMS charging is the ability to perform complex charging correlations, most notably the correlation of tickets generated by the GGSN and SGSN (i.e. non-IMS-components) with tickets coming from IMS network components. IMS separates signalling from data/media streams, so the GGSN (sometimes even the SGSN) is the only component that can record or charge data/media streams. The bearer-level charging functionality is basically identical to 3GPP Release 99 and must be changed to cope with the needs of multi-service networks.

The Diameter protocol has gained especial importance to an extent that, in short, without Diameter there will be no IMS. Although often referred to as real-time billing protocol, Diameter is in reality a way of retrieving and using customer data, i.e. it is not only able to check a customer's authenticity and credit status, but also, for example, to detect if a customer who is downloading a particular music track has previously downloaded one by the same artist. Thus, the user could, before making the purchase, perhaps be offered a desktop wallpaper if she buys two tracks instead of one of that artist's latest video. A part of the Diameter specification also takes care of the content-settlement function, ensuring that third-party providers get paid immediately.

This is just a relatively simple example of the potential of Diameter in conjunction with a SIP-based IP network, giving a hint of its ability to combine, correlate and charge for customer activity across a range of services and access clients. Functionally, Diameter relies on compliant applications running on the edge of the IMS network like the already fully developed Diameter Credit Control Application (DCCA). The DCCA defines how network equipment issues real-time charging requests to the network servers that control service delivery, it offers new credit-authorization models, and it provides operators with standard-based interface to deliver services to prepaid users. DCCA can be also of use for enabling postpaid services, e.g. if credit limits have been set or a customer needs to be told in advance of the cost of a particular transaction. In both these cases, the DCCA can provide the operator network and the customer with the relevant information and as such represents truly converged charging and billing.

Another interesting point about Diameter is that traditional billing solutions collect customer data for charging purposes in the form of CDRs from several points in the network. Since Diameter sits between the network databases and the application layer and works in such a different way from conventional methods, completely new solutions have to be developed. There are other protocols, such as Parlay and Web Services, but the consensus seems to be that they will at best be used by smaller vendors to expose objects for building applications.

Summarizing, there are still several challenges for IMS implementers and integrators left:

- *Correlation* of charging records sent by IMS and by non-IMS nodes in real time is non-trivial
- *Implementation* of dedicated charging mechanisms, like flow-based charging, requires significant architectural changes and infrastructure and may endanger compatibility to existing systems
- *IMS Roaming*, which requires correlation of third-party charging records with own records, raises highly complex problems
- *Service Convergence*: create more attractive packages for consumer and enterprise customers by combining voice, data, applications and media
- *Payment Convergence*, i.e. a single solution to effectively manage prepaid and postpaid services
- *Partner Settlement*: simultaneous calculation of partner payments / end-user prices in real-time
- *High Performance Rating*: rate millions of transactions per hour for all parties in the value chain
- *Profit Simulation*: give marketing teams the power to define and test business models for new services with integrated simulation tools

As far as the future of conventional billing platforms is concerned, the entire charging-related 3GPP specifications basically are a shift towards the network, and the billing platform will no longer perform mediation, rating and balance management. Of course, invoices will still be produced, but even that could be easily delegated to outside players like as SAP and Oracle as a complement to CRM services. Therefore, billing has to become part of understanding the real-time behaviour of the customer, especially when service providers move to a service delivery platform that is designed to be highly flexible. Thus, in general the IMS significantly raises the need for intelligent and accurate billing systems that communicate with the network in real-time. Nevertheless, the importance of charging and billing in IMS cannot be overestimated, as there is the common conviction that the success or failure of IMS will vitally depend on the availability of charging models that are transparent and attractive for the user.

5.2. Quality-of-Service

Since the early days of the Internet, Quality-of-Service (QoS) has been in the center of research and development activities both in industry and academia. Basically, the topic of QoS has emerged from the fundamental difference between the connection-oriented telephone network and connectionless packet-switched networks like the Internet. In the latter, resources are not explicitly allocated to individual pairs of hosts in the network, instead the available network capacity is shared at each link between several connections, leading to packet losses within connections at the links' output buffers and potentially causing quality impairments at the receiving side. Besides the traditional view on QoS in terms of connection parameters like packet loss or delay, also system availability and perceived quality become increasingly important, as bandwidth is often overprovisioned and the success of novel applications like Voice-over-IP (VoIP) mainly depends on user satisfaction.

As an IP-based system, IMS will inevitably inherit most of the QoS problems of the Internet. However, in contrast to the Internet where users accept occasional QoS degradations, IMS users will most likely expect the same high quality levels as in circuit-switched networks, based on their experience with traditional mobile terminals. Taking into consideration the architecture of 3G networks, we can summarize the most important aspects of QoS in the IMS as follows:

- *QoS in the strict sense,* i.e. packet loss, delay, delay variation, bandwidth: we may assume that 3G core networks will be sufficiently overprovisioned, and thus offer excellent QoS. Therefore, strict sense QoS problems will be limited to the radio access part, which is managed by means of PDP context activation, and therefore pretty much independent of functionalities defined in the IMS.
- *System availability:* IMS will be built on top of the existing IP infrastructure, and thus it will inherit all resilience problems of current IP networks. Therefore it is of utmost importance that the underlying core IP network is properly engineered in order to maximize IMS system reliability [51][52]. For the IMS components themselves, attention must be paid to failover scenarios in which the components' states need to be saved and transferred to backup working components. The goal is to advance the level of IMS availability to that of traditional telephony systems.
- *Perceptual QoS:* The crucial point for a successful deployment of multimedia services is the QoS as perceived by the user (also known as "Quality-of-Experience" QoE [53]). For instance, by making appropriate choices concerning the content type, compression and network settings of VoIP [54] or video sequences [55], the perceptual quality of the multimedia content can be opti-

mized. Even if we reduce the parameter space in our optimization to a minimum, the complexity of quality estimation and maximizing perceptual quality still remains very high [56].

- *Optimizing perceived QoS:* predict and maximize the perceptual quality for low resolution and low rate video and audio streaming services, with a focus on estimating the quality of IMS multimedia streams at the user-level and finding optimal codec settings for different IMS streaming scenarios. In order to select optimal codec parameters, it is important to consider corresponding quality requirements based on human perception [57].

5.3. IPv4/v6 Interworking

Initially IMS was strictly based on exclusively using IPv6 in the IMS core nodes and in the terminals. Only the CS domain and IP transport elements like the Radio Network Controller were allowed to either use IPv4 or IPv6 as transport protocol. But with version 6.3.0 of 3GPP TS 23.221 [58] the restrictive IPv6 policy has been softened and now it is also possible to use IPv4 for core elements and for the terminals. This tremendous IMS design modification allows implementing IMS networks which are completely relying on IPv4 alone or which are using a dual stack approach.

But if no homogeneous IPv6 architecture is used, there are many interworking issues emerging. In some cases this could even lead to an interworking incompatibility as the following example demonstrates:

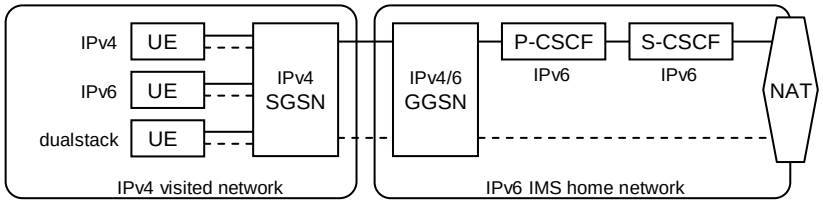

Figure 12: Simplified Home Roaming Scenario with IPv6-only IMS Nodes

Figure 12 illustrates a typical “home roaming” scenario with the P-CSCF located in the home network. This roaming type is important especially for early IMS deployments, since the visited network will be most likely a 2G network. As the IMS network in this example is IPv6 only, any IPv4 terminal would be unable to register to IMS (this situation could occur if a customer unlocks an external IPv4 phone which has not been provided by the operator).

But even IPv6 or dual stack terminals could fail to register if the visited IPv4 SGSN does not support the establishment of an IPv6 PDP context. Although SGSN and GGSN can communicate on a hop-by-hop basis via IPv4 GTP (GPRS Tunnelling Protocol), this is not sufficient for an end-to-end IPv6 connection. So maybe there is no IMS roaming possible in this scenario. According to TR 23.981 [59], this problem can be only solved by one of the following workarounds:

- If possible the home operator should try to negotiate an agreement with his roaming partner in order to support and allow the establishment of IPv6 PDP context over the visited SGSN.

- The terminal could establish an IPv6 tunnel on top of the IPv4 PDP context. But this solution needs a tunnelling gateway for unpacking and terminating the tunnel. Furthermore this procedure increases the network overhead and requires additional processing power at the terminal.

As a conclusion from this little exercise, an IPv4 or a dual stack approach may be only implemented if the network elements follow the recommendations and guidelines of TR 23.981 [59]. Note further that the use of IPv4 is only allowed for early IMS installations – in the long term the 3GPP will demand to migrate towards IPv6 in any case.

5.4. Identity Management and USIM / ISIM Migration

IMS introduces a novel user and service data model which defines the relationship between an IMS subscriber, its user identifier and the IMS service profiles. IMS subscribers are modelled as virtual entities with associated private user identifiers which identify a subscriber uniquely and are used for registration, authentication, authorization and accounting. Public identities following the SIP-URI [60][61] or the TEL-URI [62] specifications are linked to private identities and are utilized for requesting communication to other users. Compared to the telecommunications environment, the private identifier is similar to the International Mobile Subscriber Identity (IMSI), whereas the public identifier corresponds to the Mobile Subscriber ISDN Number (MSISDN). Note that service profiles contain service specific information such as properties for time depending call-forwarding or presence settings which can be linked to one or more public identities of the subscriber.

IMS standard-based authentication is done via the IMS Subscriber Identity Module (ISIM) [63]. Similar to the Subscriber Identity Module (SIM), the ISIM is an application which resides on the Universal Integrated Circuit Card (UICC). It stores one private User Identity (UID) and one or more public UIDs which are allocated to the subscriber. While the private UID is used for authentication, the numerous public UIDs are used for the SIP-compliant addressing of the subscriber. Figure 13 shows the correlation between the subscription and the several UIDs. Since IMS R6 it is even possible to use multiple private UIDs for the same subscriber which allows the subscriber to concurrently use multiple terminals. Furthermore the ISIM stores its home network domain name, as well as security parameters like the Cipher- and Integrity-Key.

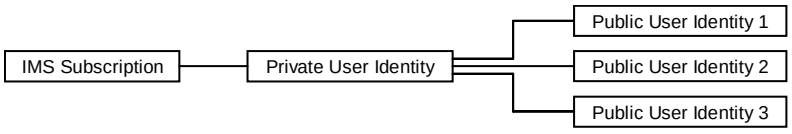

Figure 13: Relationship between Private and Public User Identities in IMS Release 5

Figure 14 illustrates the 3GPP Release 6 data model and shows a use-case where one subscriber owns multiple private identities (e.g. a private and a business SIM card) and concurrently uses several terminals. With the business SIM, he is reachable via his TEL-URI and also his business identity “user.business@a1.net”. If the subscriber is registered only with his private identity, he is reachable via his TEL-URI and his private user identifier, while call forwarding settings apply for the TEL-URI and his presence status is announced only for his private identity.

In early IMS installations the ISIM application might not be available for the mobile terminals, moreover also non-IMS subscribers should be able to authenticate and to use IMS services, allowing

for smooth migration during IMS rollout. This can be achieved by using parameters stored in the 2.5G Universal Subscriber Identity Module (USIM) application for generating temporary IMS authentication credentials [7]. The trick is to derive the required Private UID, Public UID and Network Domain URI from the International Mobile Subscriber Identity (IMSI) stored at the USIM. Thus, for the given IMSI "232011234567890" of mobilkom austria such a temporary version of an UID could for instance look like "sip:232011234567890@ims.mnc01.mcc232.3gppnetwork.org".

Another intermediate solution standardized by 3GPP re-uses information from the bearer level: if a user sets up a data call in a GPRS or UMTS network, he is already authenticated on the IP access level. [64] specifies that the authentication server of the access network sends the identity from the IP boundary to the IMS core network, where this information is reused, thus transforming all users of GPRS/UMTS data services to potential IMS customers.

Resulting from the open and extensible data model and the circumstance regarding the standardized user entities, the problems to be handled during the integration phase of IMS include:

- Most mobile operators do not have USIMs or ISIMs in the field up to now. Due to very high rollout expenses, intermediate solutions have to be utilized with limited functionality and security.
- Surrounding systems such as provisioning, billing and CRM engines have to be adopted to support the new data model and system capabilities which may result in high integration costs.

Simple and easy to use products have to be defined, where the powerful data model capabilities and the new service capabilities are bundled combined.

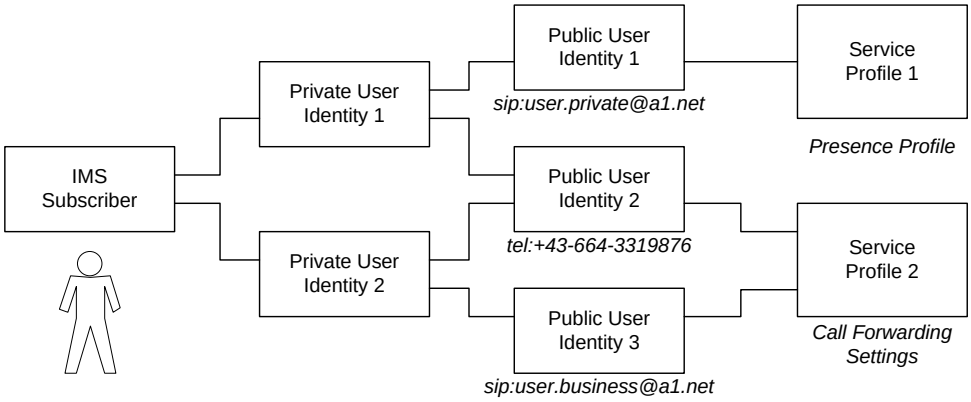

Figure 14: 3GPP R6 Data Model

5.5. Regulatory Aspects

Regulatory authorities impose similar obligations to the IMS as they do to an ordinary Public Land Mobile Network (PLMN). This includes number portability, reliability and voice quality, lawful interception as well as handling of emergency calls. Since many of the regulatory issues are still not fully standardized, this field is still providing a number of significant challenges for future releases.

Emergency Services

The integration of emergency services into the IMS has been originally announced for R6, but is now postponed to R7 [65]. The first rule for a mobile terminal setting up an emergency session is to

use the CS network. Only if this is explicitly not desired by the operator, an IMS based emergency session shall be established. To this end, the Emergency CSCF (E-CSCF) is defined in [66]. The E-CSCF is always located within the same network as the P-CSCF and routes the emergency session towards the emergency centre of its country. In the case of visited roaming this emergency centre will be the visited network, but for home roaming it would be in the home network, thus terminating the emergency call in the wrong country. In this case the (home) P-CSCF can reject the emergency session and force the terminal to place a CS emergency session in the roaming network.

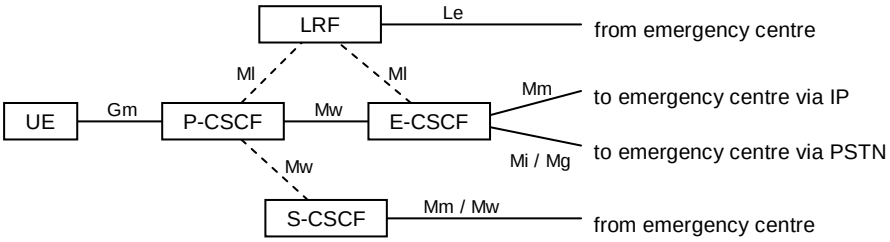

Figure 15: IMS Architecture with Emergency-CSCF (simplified)

Depending on local policies, the location of the terminal must be made visible to the emergency centre through the Location Retrieval Function (LRF). Furthermore the E-CSCF may query the address of the proper emergency centre from the LRF. Figure 15 illustrates the location of the emergency functions. The introduction of the LRF and E-CSCF brings up a bunch of new interfaces which are currently not yet defined by the 3GPP [65].

Lawful Interception

Almost every regulatory authority in any country claims to implement lawful interception. In order to fight or prevent crime it is necessary to reveal communication information of a monitored user, e.g. session events, participating addresses or even the content of a multimedia session.

To this end, the law enforcement agency is connected to several mediation functions via the HI1, HI2 and HI3 interfaces [67]. In order to avoid fraud, the access to these mediation functions must be only granted to authorized users. The mediation functions receive generic interception data and translate them into the country-specific format needed by the local law enforcement agency. The mediation functions decouple the network topology from the law enforcement agency, enabling several agencies to concurrently perform a lawful interception without interfering each other.

Figure 16 shows the concept of lawful interception. Requests from the law enforcement agency are transmitted via the Administration Function. Intercept Related Information (IRI), e.g. signalling events, is transported via the X2 interface, whereas Content of Communication (CC), e.g. media, is transported via the X3 interface. If transport security (IPsec) or compression (SigComp) mechanisms are applied, the interception should also work. Since the signalling messages are extracted at the P- or S-CSCF, they are in plain text format. If the media traffic is encrypted, the appropriate key should be also transmitted to the law enforcement agency. But if user-provided end-to-end encryption, encoding or compression is applied, it is not possible to trace the multimedia session [68], thus proprietary end-to-end mechanisms have to be barred to be able to intercept under any circumstances.

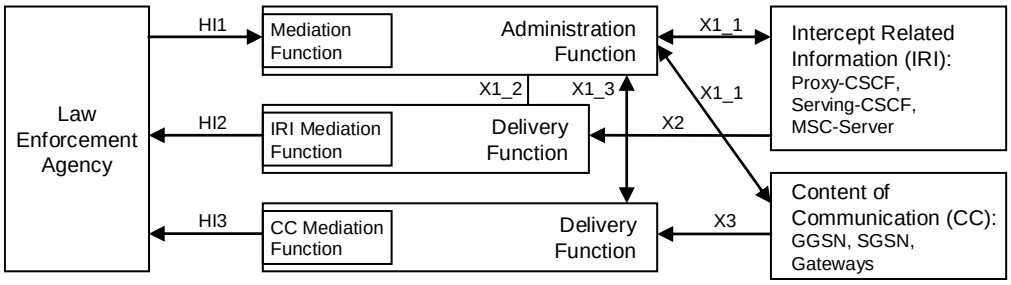

Figure 16: Lawful Interception Interface Overview

5.6. Security and Authentication

3GPP IMS provides a variety of measures to accomplish the high security requirements for the telecommunications domain. Most mechanisms use the "Authentication and Key Agreement" (AKA) mechanism executed between the customer's USIM/ISIM and the HSS on the SIP level. AKA implements secure mutual authentication and keying material derivation on the core and user equipment side, the provided keying material is utilized for integrity protection and encryption. SIP control data is protected by an IPSec tunnel between the user equipment and the P-CSCF, where the AKA-derived keys are utilized for IPSec tunnel keying.

At the time being, the first operators deploy IMS production environments but there are many reasons why the security measures are not applicable as they were specified by 3GPP initially:

- AKA-based mechanisms need USIMs/ISIMs which are used only in a minority of deployed cellular networks, due to the high migration costs for moving from GSM SIM cards to USIMs/ISIMs.
- Handsets do not support IPSec encryption.
- High bandwidth links like UMTS or HSDPA, over which encrypted SIP messages can be transmitted without harming the quality of service too much, are not widely available right now.

[64] has defined some intermediate security measures to overcome the mentioned problems. The main idea is to re-use identity information from the bearer level instead of AKA for authentication on the SIP level. Session setups on the bearer level are terminated on the home GGSN, which sends RADIUS messages containing user MSISDN, IMSI and IP-address to an external authentication server. RADIUS accounting messages are forwarded to the HSS, which re-uses the provided information for extracting the user identity to IP relation for the IMS application level. When a client issues a REGISTER procedure, the identity is registered implicitly, without deriving keying material, therefore AKA dependent mechanisms are not used. Thus, an intermediate security scenario is available, but of course the fundamental decision by the operators whether to go or not for USIM/ISIM depends on many more additional aspects.

5.7. User Equipment

Success of services offered in an IMS network will largely depend on the availability of terminals that can be used to consume these services. Apart from a few prototypes, there are currently no mobile terminals on the market that feature IMS-compatible user agent software. Mobile equipment

manufacturers face several challenges when developing terminals that support IMS based multimedia services, e.g. support of IMS authentication (USIM/ISIM authentication), IPv6, SIP/SDP/RTP and XCAP protocol stacks, XML parser, audio/video codec implementations (either in hardware or software), encryption, concurrent execution of several sessions/connections (e.g. voice, video, instant messaging, presence), integration of IMS call history, message history, contacts, buddy lists or presence data with standard mobile phone applications like address book, caller list, message store etc., and finally over-the-air provisioning of IMS configuration parameters.

This list of requirements shows that only terminal hardware platforms with excellent CPU performance, large memory size, support for many active I/O ports and above all sufficient battery capacity will be usable to implement IMS services. Furthermore operating systems of the mobile terminals have to support concurrency on a level similar to a standard PC operating system.

In this context, we have gained some practical experience by implementing a prototype IMS messaging and presence agent based on the Java 2 Platform Micro Edition (J2ME). By deploying this software on currently available phones we have learned the following lessons:

- Support for JSR-180 (SIP-stack) is still weak. We had to use a SIP stack that is available in source code (sip-for-me) in order to get presence and TCP support
- Available memory strongly restricts the number of java libraries (e.g. for XML processing).
- Processing delay of a large SIP message containing XML in the body can be larger that the retry timeout of the requestor.
- Do not expect that the Maximum Transfer Unit (MTU) of a UDP packet has the same size than in an Ethernet network. We were forced to switch our implementation to TCP because the UMTS network we used for our experiments supported only an MTU size of 512 bytes.
- TCP over an UMTS network adds considerable delay to the communication (1.5 RTTs at connection set-up time and 1 RTT for every packet sent). We measure 1 RTT of the UMTS network in our experiments to be about 150 ms.

User experience with respect to user interfaces, feature levels, quality of service or cost of packet-oriented multimedia services is currently built up in the Internet by closed services like Skype. The telecommunication industry is still facing a huge task to develop open standard-based IMS services to a comparable level, and IMS-enabled user equipment plays a crucial role in this task because the user will judge the IMS system by the user agent software.

5.8. Service Issues

Figure 1 has already depicted the technological alternatives for plugging application servers on top of the SIP ISC interface, namely OSA/Parlay, native SIP or CAMEL. Which service platform will be the appropriate one to satisfy the requirements for innovative services and short time to market? At the time of writing, IMS test applications are developed by network operators, like in the old times of IN, although IMS offers a unique chance of expose to third party applications important service functionality such as a negotiated quality of service, a flexible charging service, many valuable service enablers, roaming, service composition engines, etc.

Excluding the CAMEL interconnection, which is needed to support legacy services, our findings are, that OSA gateways cannot match well the ISC interface for the precise reason they have been

specified for: OSA gateways provide a given set of network independent capabilities like call control, messaging, etc, whereas SIP is extensible and can transport application specific data in additional headers and body of its messages. Thus, in a simple application the terminal may insert application commands or parameters in a SIP INVITE message, instead to develop an additional application protocol, however this information cannot be transported by the OSA capabilities to the application.

For the third technological alternative, the native SIP application server, we need to find more efficient ways to communicate with the IT world, using for example web service interfaces. Especially when these interfaces open up to third parties and become contract interfaces, we need a policy framework such as the one specified by OMA [69] .

6. Conclusion and Outlook

This paper has been devoted to main results of the two projects CAMPARI and SIMS, performed at the Telecommunications Research Center Vienna (ftw.). We focus on the "IMS in a Bottle" testbed and a novel architecture for a terminal-based IMS Location Service which scales better and is more accurate, efficient and cost effective than current network-based location architectures.

In the IMS core network, current and future work is focused on charging, migration, interworking, identity management, security, and regulatory aspects as well as on extensive measurement-based performance evaluations. As far as IMS services are concerned, we work on an extension of SIP event filtering specifications and mechanisms in order to express and implement complex triggering criteria in the mobile terminal. Furthermore, we are in the course of implementing a prototype system in order to prove our concept in a reference implementation, and finally we intend to contribute to the 3GPP/IMS standardisation process with the aim to add the IMS location enabler to the family of basic enablers such as presence and messaging.

Of course, work on IMS by far has not yet come to a conclusion. After 3GPP Release 6 (R6) has been completed in mid 2005, work on R7 has started and is expected to focus on leftovers from the R6 as well as defining fixed broadband access via IMS, policy issues, voice call handover between CS, Broadband Wireless Access/IMS and enhanced end-to-end QoS mechanisms. This work aims at a complete vision of IMS invoked by network operators and equipment suppliers where IMS is described as the major means to integrate voice and data services over a packet-based infrastructure, to deliver end-to-end QoS for high quality VoIP, and to converge wireline and wireless infrastructures under a common set of end-to-end signalling and billing mechanisms. The architectural complexity of IMS and its gradual growth from release to release is not just the by-product of its slow evolution through the standardization process, but also reflects the need to accommodate cellular providers as they gradually move away from their circuit-switched voice networks towards packet-oriented mobile data networks. Recently, cellular providers are also looking for new ways to integrate Wi-Fi and WiMAX (Broadband Wireless Access) technologies in their sweeping vision of IMS.

There is a unanimous worldwide support among existing telephony service providers, equipment suppliers and international standardization bodies for the IMS billing and control vision as the basis for ITU's umbrella concept of Next Generation Network (NGN). The process of fitting IMS into ITU's NGN is performed by the ETSI TISPAN group which is responsible for all aspects of standardization for present and future converged networks. This covers service, architecture and protocol aspects along with QoS and security studies as well as mobility aspects within fixed networks using

existing and emerging technologies and includes terminology adjustments as well as substantial changes aligning IMS to existing carrier architectures and planned migrations.

Finally, [70] provides possible IMS architectural enhancements necessary in the 3GPP system to support fixed broadband access to IMS, e.g. as stated in the initial ETSI TISPAN release 1. For the IMS core, 3GPP intends to develop specifications or changes to specifications necessary to enable reuse of IMS as a platform for session control in systems with fixed broadband access. R7 also specifies service requirements for "Combining CS and IMS services" using a CS speech or CS multimedia call in association with an IMS session, introducing mechanisms to manage and guarantee end-to-end QoS, packet-switched emergency calls over IMS, policy control and IP-flow based charging, and conference capabilities and group management. Of course, even if the list of new R7 features is already comprehensive, it will be likely extended further until standardization work of this release is frozen.

Acknowledgements

This work has been funded in the framework of the Austrian government's Kplus program. The authors would like to thank very much their colleagues Andreas Broch, Johannes Biedermann, Johannes Lehninger, Joachim Zeiss, Marco Happenhofer and Christoph Egger for their invaluable help with the implementation and all of their SIMS and CAMPARI colleagues for their continuous support and enthusiasm. Last not least our thanks go to the anonymous referees for the many detailed and helpful hints concerning the revised version of the paper.

References

[1] G. Camarillo, M. Garcia-Martin, *The 3G IP Multimedia Subsystem (IMS)*, Wiley 2006.
[2] Poikselka, M., G. Mayer, H. Khartabil and A. Niemi, *The IMS: IP Multimedia Concepts and Services in the Mobile Domain*, Wiley 2004.
[3] 3rd Generation Partnership Project (3GPP), see http://www.3gpp.org.
[4] http://ims.ftw.at
[5] http://www.ftw.at/ftw/research/projects/ProjekteFolder/P4
[6] 3GPP TS 22.071, "Technical Specification Group Services and System Aspects; Location Services (LCS); Service description; Stage 1", version 7.1.0, 2005-01
[7] 3GPP TS 23.228: "IP Multimedia Subsystem (IMS); Stage 2".
[8] 3GPP TS 22.141: "Presence Service, Stage 1".
[9] IETF RFC 3261: "SIP: Session Initiation Protocol", June 2002.
[10] IETF RFC 3588: "Diameter Base Protocol", September 2003.
[11] Kernel-based Network Emulator, see http://linux-net.osdl.org/index.php/Netem.
[12] RFC 3263, "Session Initiation Protocol (SIP): Locating SIP Servers", June 2002.
[13] SIP Express Router (SER), http://www.iptel.org/ser/
[14] http://www.ftw.at/ftw/research/projects/ProjekteFolder/A3
[15] P. Reichl, M. Umlauft, J. Fabini, et al.: "Project WISQY: A Measurement-Based End-to-End Application-Level Performance Comparison of 2.5G and 3G Networks", Proc. (IEEE) 5th Wireless Telecommunications Symposium WTS'05, Pomona, CA, April 2005.
[16] P. Reichl, N. Jordan, J. Fabini, et al.: "Wireless Inter-System Quality-of-Service: A Practical Performance Analysis of 3G and Beyond", Kommunikation in Verteilten Systemen (KiVS'05), 2005.

[17] Charter of the Geographical Location and Privacy (GEOPRIV) working group of the Internet Engineering Task Force (IETF), http://www.ietf.org/html.charters/geopriv-charter.html
[18] IETF RFC 4079 “A Presence Architecture for the Distribution of GEOPRIV Location Objects”.
[19] Open Mobile Alliance, Secure User Plane Location Architecture, OMA-AD-SUPL-V1_0-20060127-C, http://www.openmobilealliance.org
[20] A. Küpper, G. Treu: From Location to Position Management: User Tracking for Location-based Services. Kommunikation in Verteilten Systemen (KiVS) Kurzbeiträge 2005, pp. 81-88
[21] R. Shaham, H. Schulzrinne, W. Kellerer, S.Thakolsri, An Architecture for Location-based Service Mobility Using the SIP Event Model, International Conference on Mobile Systems Application and Services, (Mobys 2004), Boston, USA, 2004
[22] J. M. Polk, B. Rosen, “Requirements for Session Initiation Protocol Location Conveyance”, IETF Internet-Draft, draft-ietf-sipping-location-requirements-02.txt, October 25th, 2004 – expired April 25, 2005.
[23] H. Schulzrinne, V. Gurbani, P. Kyzivat, J. Rosenberg, "RPID: Rich Presence: Extensions to the Presence Information Data Format (PIDF)", IETF Internet-Draft, draft-ietf-simple-rpid-10 December 20, 2005.
[24] ETSI Standard ES 202 391-9, Open Service Access (OSA); Parlay X Web Services; Part 9: Terminal Location, V1.1.1 (2005-03)
[25] IETF RFC 2778, “A Model for Presence and Instant Messaging” , February 2000
[26] IETF RFC 3265, “Session Initiation Protocol (SIP)-Specific Event Notification”, June 2002
[27] 3GPP TS 23.141, “Presence Service; Architecture and Functional Description”, Version 6.7.0, 09/2004
[28] IETF RFC 3856: “A Presence Event Package for the Session Initiation Protocol (SIP)”, 2004
[29] H. Khartabil, E. Leppanen, M. Lonnfors, J. Costa-RequenaI, "Functional Description of Event Notification Filtering", IETF Internet-Draft, draft-ietf-simple-event-filter-funct-05, March 15, 2005 – expired September 16, 2005.
[30] H. Khartabil, E. Leppanen, M. Lonnfors, J. Costa-RequenaI, "An Extensible Markup Language (XML) Based Format for Event Notification Filtering", IETF Internet-Draft, draft-ietf-simple-filter-format-05, March 15, 2005 – expired September 16, 2005.
[31] A.B. Roach, J. Rosenberg, B. Campbell, A Session Initiation Protocol (SIP) Event Notification Extension for Resource Lists, IETF Draft, 15-Dec-05. http://www.ietf.org/internet-drafts/draft-ietf-simple-event-list-07.txt.
[32] J. Rosenberg, Extensible Markup Language (XML) Formats for Representing Resource Lists, IETF Draft, 9-Feb-05. http://www.ietf.org/internet-drafts/draft-ietf-simple-xcap-list-usage-05.txt.
[33] J. Rosenberg, “The Extensible Markup Language (XML) Configuration Access Protocol (XCAP)”, draft-ietf-simple-xcap-08, IETF Internet-Draft, October 24, 2005
[34] J. Rosenberg, "An Extensible Markup Language (XML) Document Format for Indicating Changes in XML Configuration Access Protocol (XCAP) Resources”, IETF Internet-Draft, draft-ietf-simple-xcap-diff-03, March 6, 2006.
[35] IETF RFC 3863, “Presence Information Data Format (PIDF)”, August 2004.
[36] H. Schulzrinne, V. Gurbani, P. Kyzivat, J. Rosenberg, "RPID: Rich Presence: Extensions to the Presence Information Data Format (PIDF)", IETF Internet-Draft, draft-ietf-simple-rpid-10 , 2005
[37] 3GPP TS 23.271, “Technical Specification Group Services and System Aspects; Functional stage 2 description of Location Services (LCS)”, version 7.0.0, 2005-03

[38] C. Johnson, H. Joshi, J. Khalab, “WCDMA radio network planning for location services and system capacity”, 3G Mobile Communication Technologies, 2002. Third Int. Conf. on (Conf. Publ. No. 489), 2002, pp. 340- 344

[39] ETSI Standard ES 202 391-9, Open Service Access (OSA); ParlayX Web Services; Part 9: Terminal Location, V1.1.1 (2005-03)

[40] Java Specification Request JSR-179, ‘Location API for the Java 2 Platform, Micro Edition (J2ME) version 1.0’, Nokia, 2003. http://jcp.org/en/jsr/detail?id=179

[41] R. Pailer, F. Wegscheider, S. Bessler, „ A Terminal-Based Location Service Enabler for the IP Multimedia Subsystem”, Proc. WCNC 2006, Las Vegas, April 2006.

[42] J. Fabini, R. Pailer, M. Happenhofer, "Terminal-Centric Location Services for the IP Multimedia Subsystem", Proc. IEEE VTC 2006-Spring, Melbourne, Australia, April 2006.

[43] OpenGIS, "Open Geography Markup Language (GML) Implementation Specification", OGC 02-023r4, January 2003, http://www.opengis.org/techno/implementation.htm.

[44] http://www.ldeo.columbia.edu/~dale/scicex99/photos/E.%20%20North%20Pole/5.%20%20GPS.jpg

[45] F. Wegscheider: “Minimizing Unnecessary Notification Traffic in the IMS Presence System”. Proc. 1st International Symposium on Wireless Pervasive Computing, January 2006, Thailand.

[46] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, J. Jetcheva, “A Performance Comparison of Multi-Hop Wireless Ad Hoc Network Routing Protocols”. Proc. 4th Annual ACM/IEEE International Conference on Mobile Computing and Networking, Dallas, TX, October 1998.

[47] J. Yoon, M. Liu, and B. Noble, “Random waypoint considered harmful.” Proc. Infocom '03, p. 1312-1321, San Francisco, California, USA, April 2003.

[48] BonnMotion, A mobility scenario generation and analysis tool, Communications Systems Group, Institute of Computer Science IV, University of Bonn, Germany. URL: http://web.informatik.uni-bonn.de/IV/Mitarbeiter/dewaal/BonnMotion/

[49] 3GPP TS 32.260: “Telecommunication management; Charging management; IP Multimedia Subsystem (IMS) charging”.

[50] 3GPP TS 23.125: “Overall high level functionality and architecture impacts of flow based charging; Stage 2”

[51] I. Gojmerac, F. Hammer, F. Ricciato, H. T. Tran, T. Ziegler: “Scalable QoS: state-of-the-art architectural solutions and developments”. Technical Report FTW-TR-2004-003, March 2004

[52] G. Iannaccone, C.. Chuah, R. Mortier, S. Bhattacharyya, C. Diot: “Analysis of link failures in an IP backbone”. Proc. ACM SIGCOMM, Marseilles, France, Nov 2002.

[53] P. Fröhlich, L. Baillie, P. Reichl, R. Schatz, F. Hammer, G. Niklfeld: „The HTI Lab @ ftw.: User Research for Telecom Systems“.CHI 2006, Conference on Human Factors in Computing Systems, Montreal, Canada, April 2006.

[54] F. Hammer, P. Reichl, Th. Ziegler: Where Packet Traces Meet Speech Samples: An Instrumental Approach to Perceptual QoS Evaluation of VoIP. Proc. of 12th IEEE International Workshop on QoS (IWQoS'04), Montreal, Canada, June 2004.

[55] M. Ries, R. Puglia, T. Tebaldi, O. Nemethova, M. Rupp: "Audiovisual Estimation for Mobile Streaming Services," Proc. International Symposium on Wireless Communication Systems IEEE Ed., Siena, Italy, Sept, 2005.

[56] M. Ries, O. Nemethova, M. Rupp: "Reference-Free Video Quality Metric for Mobile Streaming Applications," Proc. of DSPCS/WITSP 05, pp. 98-103, Sunshine Coast, Australia, Dec 2005.

[57] M. H. Pinson, S. Wolf: "A new standardized method for objectively measuring video quality," IEEE Transactions on broadcasting, Vol. 50, Issue: 3, pp 312-322, Sept, 2004.

[58] 3GPP TS 23.221: “Architectural requirements”.

[59] 3GPP TR 23.981: “Interworking aspects and migration scenarios for IPv4-based IMS implementations”.
[60] IETF RFC 2396: “Uniform Resource Identifiers (URI): Generic Syntax”, August 1998.
[61] IETF RFC 2778: “A Model for Presence and Instant Messaging”, February 2000.
[62] IETF RFC 3966: “The tel URI for Telephone Numbers”, December 2004
[63] 3GPP TS 31.103: “Characteristics of the IP Multimedia Services Identity Module (ISIM) application”.
[64] 3GPP TS 33.978: “Security aspects of early IP Multimedia Subsystem (IMS)”
[65] 3GPP TS 23.167: “IP based IP Multimedia Subsystem (IMS) emergency sessions; Stage 2”.
[66] 3GPP TS 23.867: “IP based IP Multimedia Subsystem (IMS) emergency sessions”.
[67] 3GPP TS 33.107: “3G Security; Lawful interception architecture and functions”.
[68] 3GPP TS 33.106: “Lawful interception requirements”.
[69] Open Mobile Alliance, Policy Evaluation, Enforcement and Management Requirements, OMA-RD-Policy_Evaluation_Enforcement_Management-V1_0-20050112-C, 2005.
[70] 3GPP TSGS#25(04)0686: “System enhancements for Fixed Broadband access to IMS”.

Appendix: List of Acronyms

3G	3rd Generation
3GPP	3rd Generation Partnership Project
A-BPS	Network-Assisted GPS
AKA	Authentication and Key Agreement
API	Application Programming Interface
AS	Application Server
AVP	Attribute Value Pairs
BGCF	Breakout Gateway Control Function
CC	Content of Communication
CCF	Charging Collection Function
CDR	Charging Data Record
COTS	Commercial Off The Shelf
CRM	Customer Relationship Management
CS	Circuit Switched
CSCF	Call Session Control Function
DHCP	Dynamic Host Configuration Procotol
DNS	Domain Name Server
DCCA	Diameter Credit Control Application
E-CSCF	Emergency CSCF
EDGE	Enhanced Data Rates for GSM Evolution
GEOPRIV	Geographical Location and Privacy
GERAN	GPRS EDGE Radio Access Network
GGSN	GPRS Support Node
GML	Geography Markup Language
GMLC	Gateway Mobile Location Center
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile communications
GTP	GPRS Tunneling Protocol
GUI	Graphical User Interface
HDD	Hard Disk Drive
HLR	Home Location Register
HSDPA	High Speed Downlink Packet Access
HSS	Home Subscriber Service
ICE	Intercepting Control Elements
ICID	IMS Charging Identifier
I-CSCF	Interrogating CSCF
IETF	Internet Engineering Task Force
IFC	Initial Filter Criteria
IMS	IP Multimedia Subsystem
IMSI	International Mobile Subscriber Identity
I/O	Input/Output
IP	Internet Protocol
IPDL	Idle Period Downlink
IRI	Intercept Related Information
ISIM	IMS Subscriber Identity Module
ITU	International Telecommunications Union
J2ME	Java 2 Platform, Micro Edition
LBS	Location-Based Service
LCS	Location Service
LIMS-IWF	Location IMS – Interworking Function
LRF	Location Retrieval Function
MGCF	Media Gateway Control Function
MGW	Media Gateway
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
MSC	Mobile Switching Center
MSISDN	Mobile Subscriber ISDN Number
MTU	Maximum Transfer Unit
NAT	Network Address Translation
NGN	Next Generation Network
OCS	Online Charging System
OMA	Open Mobile Alliance
OTDOA	Observed Time Difference of Arrival
P-CSCF	Proxy CSCF
PDF	Policy Decision Function
PDP	Packet Data Protocol
PES	Presence Edge Server
PIDF	Presence Information Data Format
PLMN	Public Land Mobile Network
PS	Packet Switched
PSTN	Public Switched Telephone Network
PUA	Presence User Agent
QoE	Quality of Experience
QoS	Quality of Service
R5/6/7	Release 5/6/7
RLS	Resource List Server
RPID	Rich Presence Information Data
RTT	Round Trip Time
SBLP	Service Based Local Policy
S-CSCF	Serving CSCF
SDP	Session Description Protocol
SER	SIP Express Router
SGSN	Serving GPRS Support Node
SGW	Signalling Gateway
SIM	Subscriber Identity Module
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions
SIP	Session Initiation Protocol
SLF	Service Location Function
TCP	Transmission Control Protocol
UA	User Agent
UDP	User Datagram Protocol
UE	User Equipment
UICC	Universal Integrated Circuit Card
UID	User Identity
UMTS	Universal Mobile Telecommunication System
URI	Uniform Resource Identifier
USIM	Universal Subscriber Identity Module
WAN	Wide Area Network
WE	WAN Emulator
WLAN	Wireless Local Area Network
WLS	Watcher List Server
XCAP	XML Configuration Access Protocol
XML	Extensible Markup Language