The FOKUS Open SIP AS - A Service Platform for NGN

Elmar Fasel, Karsten Knuettel, Thomas Magedanz

{fasel|knuettel|magedanz}@fokus.fraunhofer.de
TU Berlin, Lehrstuhl AV – http://www.av.tu-berlin.de/
Fraunhofer FOKUS NGNI – http://www.fokus.fraunhofer.de/

Abstract: Evolutionary, a Next Generation Network (NGN) is a significant improvement of a successfully existing network. The third Generation Partnership Project (3GPP) aimed to merge two of the most successful paradigms in communications: cellular networks and the Internet. Within the IP Multimedia Subsystem (IMS) the 3GPP specified a comprehensive and service oriented architecture that includes session based Quality of Service in IP, charging mechanisms and offers standardized interfaces for application service integration and seamless interactions with legacy services. Under this technical premise a variety of promising value added services suppose to attend our entire communication life. Fraunhofer Institute FOKUS designed and implemented a 3GPP compliant Application Server for converged services called the "Open SIP AS" which is deploying services within the "National Host for 3Gb Applications Testbed" on top of an IMS architecture. This paper describes details of the implementation and the usage of the "Open SIP AS" and gives an example of a NGN service.

1 Introduction

Third Generation (3G) mobile networks aim to merge two of the most successful paradigms in communications: cellular networks and the Internet. The IP Multimedia Subsystem (IMS) is the key element in the 3G architecture that makes it possible to provide ubiquitous cellular access to all the services that the Internet provides. Picture yourself accessing your favourite web pages, reading your email, watching a movie or taking part in a videoconference wherever you are by simply pulling a 3G hand-held device out of your pocket. The concept of IMS was developed by 3GPP as an extension of the Global Packet Radio System (GPRS) Core Network towards the "all IP" architecture to allow for smooth integration of new IP based services with cellular architectures. As mentioned before the idea of IMS is to offer Internet services everywhere and at any time using cellular technologies, but you may ask now the legitimate question: Cellular networks already provide IP over Circuit- and Packet Switched, so what do we need the IMS for? First of all there is a variety of advantages using Packet Switched (PS) for all services including voice. To name only a few of them, there is much more efficient usage of resources, maintenance in a single PS network, faster data transmissions, higher bandwidth and so on. But this is no sufficient answer to the question "why IMS?", because all the advantages could be gained with a cellular VoIP system. There are four additional key functionalities that mark the IMS as the future technology in a comprehensive service and application oriented network.

- The IMS provides event oriented Quality of Service feasibilities.

- The IMS provides event oriented charging mechanisms.

- The IMS provides easy and efficient ways to integrate different services, even from third parties. Interactions between different value added services like voice, web and precence are anticipated.

- The IMS enables the seamless integration of legacy services and is designed for consistent interactions with circuit switched domains.

All those techniques and methodologies are not new, but IMS provides the integration and the interaction of all key functionalities and point out a migration path for existing cellular networks. Still, there are many open issues and the 3GPP IMS is not closed in specification. None of the commercial operators is running an IMS yet, but academics are doing research and the industry provides partial solutions. Against the background of parallel engineering of academia and industry within the same area of interest, the Fraunhofer FOKUS got the mandate from the German ministry of Education and Research to establish a "National Host for 3Gb Applications Testbed" based on traditional experiences and successes in the research area of VoIP and Next Generation Networks as well as the strong linkage to industrial partners. With the aim of shortening the time to market of IMS and the exchange of knowledge and requirements between academia and industry FOKUS designed an IMS Testbed named IMS Playground. FOKUS implemented all core components of the IMS and enriched this base infrastructure with components from commercial vendors. Different programming paradigms and enabling technologies like OSA/Parlay [parlay], Web Services, Servlets, CPL [IETF CPL], CGI [IETF CGI] or JAIN SLEE [JAIN] are equivalent considered, implemented and validated.

The focal point of this paper is the "Open SIP AS". It is a service platform for NGNs and was designed for service provisioning within the IMS. The "Open SIP AS" is based on SIP Servlet technology specified in the JSR 116 SIP Servlet API [JSR 116]. Through the relationship between SIP and HTTP Servlets the JSR 116 is a very convenient and efficient enabling technology for converged services. A converged service in this context defines a combined service that makes similarly and seamlessly use of SIP, as the protocol for the session initiation and of HTTP in an advanced way.

2 IMS specifications review and SIP AS options

The 3rd Generation Partnership Project was founded in 1998 as a collaboration agreement between regional telecommunication standard bodies. Current partners are:

- Association of Radio Industries and Business in Japan (ARIB)
- China Communications Standards Associations (CCSA)
- European Telecommunications Standards Institute (ETSI)
- Committee T1 in the United States of America
- Telecommunications Technology Association of Korea (TTA)
- Telecommunication Technology Committee in Japan (TCC)

The consequent realisation of the aim to establish a globally applicable third-generation mobile system based on GSM resulted, beside some other attainments, in the IMS. 3GPP does not produce standards but Technical Specifications (TS) and Technical Reports (TR). TS and TR are grouped in Releases. 3GPP Release 5 is frozen and contains a first version of IMS. Release 6 is under specification and contains enhancements to the base infrastructure. The FOKUS IMS Playground and the Open SIP AS is an implementation of TS Release 5. 3GPP [3GPP.x] specifies the IMS Core Network (IM CN) functionalities and interfaces to all entities as well as protocols (adopted from IETF) and protocol enhancements. In the context of application service provisioning interfaces to the signalling plane and to the management/control plane are sufficiently defined. Interactions to existing and approved systems like OSA/Parlay and IN or CAMEL are also defined, but the implementation path has been left unclear on purpose. Thus many options and programming paradigms for service provisioning compete for establishment in the IMS.

2.1 IMS

This chapter provides you with a brief overview to the IP multimedia core network subsystem. See figure 1 “IMS overview”. For clarification: Detailed specification of IMS functionalities and interfaces is omitted.

P-CSCF The Proxy Call State Control Function is the first contact point within the IM CN subsystem. Its address is discovered by User Entities (UE) following the Packet Data Protocol context activation. The P-CSCF behaves like a Proxy accepting requests and services them internally or forwards them on. Performed functions are: Authorize the bearer resources for the appropriate QoS level. / Emergency calls / monitoring / header (de)compression / identification of I-CSCF.

I-CSCF The Interrogating Call State Control Function is the contact point within an operator's network for all connections destined to a subscriber of that network operator, or a roaming subscriber currently located within that network operator's service area. There may be multiple I-CSCFs within an operator's network. Performed functions are: assigning an S-CSCF to a user performing SIP registration / charging and resource utilisation: generation of Charging Data Records (CDRs) / acting as a Topology Hiding Inter-working Gateway (THIG)

S-CSCF The Serving Call State Control Function performs the session control services for the endpoint. It maintains session state as needed by the network operator for support of the services. Within an operator's network, different S-CSCFs may have different functionality. Performed functions are: User Registration / Interaction with Services Platforms for the support of services. The S-CSCF decides whether an Application Server is required to receive information related to an incoming SIP session request to ensure appropriate service handling. The decision at the S-CSCF is based on filter information received from the HSS. This filter information is stored and conveyed on an application server basis for each user.

The ***Home Subscriber Server*** (HSS) is the equivalent of the Home Location Register. It stores subscriber profiles which may be requested from other nodes (AS or CSCF).

The ***SIP Application Server (AS)*** is the service relevant part in the IMS. Applications providing value added services are deployed on the application servers. The well defined APIs enable developers to use almost any programming paradigm. The SIP AS is triggered by the S-CSCF which redirects certain sessions to the SIP AS based on filter information obtained from the HSS. The SIP AS itself comprises filter rules to decide which of the applications deployed on the server should be selected for handling the session. During execution of service logic it is also possible for the SIP AS to communicate with the HSS to get additional information about a subscriber or to be notified about changes in the profile of the subscriber.

The ***Media Server*** can be split up into the Media Resource Function Controller (MRFC) and the Media resource Function Processor (MRFP). It provides media stream processing resources for e.g. media mixing, media announcements, media transcoding.

The triple of ***Border Gateway Control Function*** (BGCF), ***Media Gate Control Function*** (MGCF) and ***Media Gate*** (MG) is abstracted in the figure to B/MGF and performs the bearer interworking between RTP/IP and the bearers used in the PSTN/ISDN/PLMN networks.

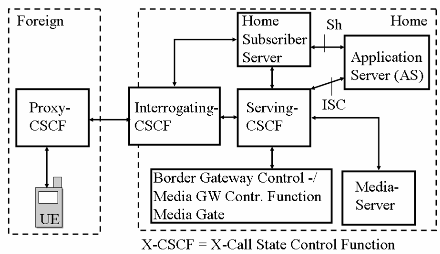


figure 1: IMS overview

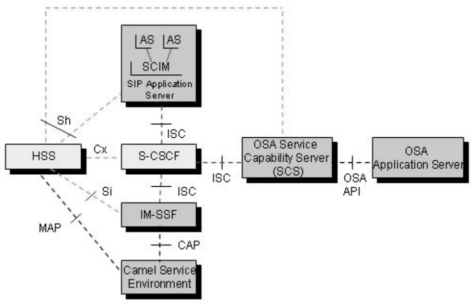


figure 2: SIP Application Server Options

2.2 SIP AS options

Applications providing value added services are deployed on the application server. Well definied APIs enable developers to develop new applications which may be deployed on any SIP application server implementing the same API. 3GPP suggests the following technics and APIs for implementing a SIP application server:

CPL Call Processing Language

Applications are triggered by user-created program written in a simple, static, non-expressively-complete language. It is meant to be simple, extensible, easily edited by graphical clients, and independent of operating system or signalling protocol. It is suitable for running on a server where users may not be allowed to execute arbitrary programs, as it has no variables, loops, or ability to run external programs.

SIP CGI Common Gateway Interface

CGI is a generic interface for implementing services on Application Servers. Unlike the CPL, it is a very low-level interface, and would not be appropriate for services written by non-trusted users. Also, SIP CGI [IETF CGI] can be implemented in languages that allow recursion and complex loops. Hence, there is no guaranty that the CGI will terminate in short time. This aspect is very important for a commercial service provider as it can affect call setup time and the reliability / robustness of its service.

SIP Servlet API

The Servlet API defines ways of interacting programmatically with requests and responses on the server side. A known derived technology is the HTTP Servlets API for HTTP requests and responses. In imitation of the HTTP Servlets the SIP Servlet specification was developed for the creation of server based SIP applications.

OSA/Parlay application server

Parlay is an open multi-vendor consortium developing - in cooperation with 3GPP - open technology-independent application programming interfaces (APIs). The OSA/Parlay APIs can be used to develop mobile data services, as well as fixed network services and services for a converged Next Generation Networks. OSA/Parlay includes specific capabilities designed to enhance mobile data services, including location, terminal capability, connection management and charging. It provides a secure framework for third party application providers.

IMS-SSF

IP Multimedia Subsystem - Services Switching Function provides the interworking of the SIP message to the corresponding Customized Applications for Mobile Networks Enhanced Logic (CAMEL), ANSI-41, Intelligent Network Application Protocol (INAP) or Transaction Capabilities Application Part (TCAP) messages [3GPP.x].

2.3 Convergence of SIP & HTTP

The WWW is perceived by most people as „the Internet“and is a well known and accepted communication medium for over 800 million people worldwide. The mainly use HTTP for their sessions. By comparison SIP is a relatively new technology and fairly accepted. To gain all the advantages of VoIP there is the need to merge these two powerful technologies (SIP & HTTP). Enhanced and converged applications riches our entire communication process. By the combination of SIP and HTTP VoIP can gain a strong impulse to spread out. The concepts behind this bridge are applications which are build on both protocols similarly. A convergent application in combination with an advanced user devices may switch seamlessly or uses in parallel the two paradigms to resolve the features not possible by a sole technology.

Subsequently an example will be described and explained whereas a caller gets a pre-filled web page instead of a busy tone if the B-party is not available.

3 FOKUS Open SIP AS

The Open SIP AS is implemented in the Java Programming Language, since servlets are a Java related specification. The SIP application server has two interfaces to other elements in the IMS. It mainly communicates over the IMS Service Control Interface (ISC), which is basically SIP with some extensions. Therefore it makes sense to introduce a separation between the low level communication and the servlet execution environment which uses the SIP communication facilites. Like this, our environment is build upon a JAIN SIP stack which fully implements the SIP protocol and handles all low level affairs. The NIST SIP stack [NIST] is a well known JAIN SIP compliant SIP stack. As a bonus feature the used stack may be exchanged with any other JAIN SIP stack with ease. As displayed in figure 3 “FOKUS SIPSEE” implementation the second communication possibility is over the Sh interface. The Sh protocol is a Diameter based protocol. Once again we depend for this part on a well known and approved solution called Open Diameter. As Open Diameter is a C implementation of the Diameter base protocol we developed a JNI wrapper to use OpenDiameter within the Java Programming Language. Based on this wrapper we created a Diameter infrastructure which allows us not only to implement commands and AVPs of the Sh protocol in a rapid way, but every other Diameter application as well. On these communication facilities we created a servlet container which is a suitable environment to deploy SIP servlet applications. To allow the deployment and execution of converged applications we use Jetty as HTTP servlet container and integrated our SIP servlet execution environment into it. Both parts have access to a common application session repository for means of exchanging data between the HTTP and the SIP servlets. Currently the SIP servlet specification proposes a way to get a spanning application session from a SIP servlet. But as the concept of application sessions is unknown to the current version of the HTTP servlet specification, there is no standard way to get hold of an application session from a HTTP servlet. The Open SIP AS features a proprietary method to get an application session by means of a "ConvergedApplicationHelper" instance. This helper object is stored in the servlet context and accessible by every servlet. Additionally, a method is provided to SIP servlets for encoding the application session ID into an URL in such a way, that the application server can execute a HTTP servlet in the context of the requested application session.

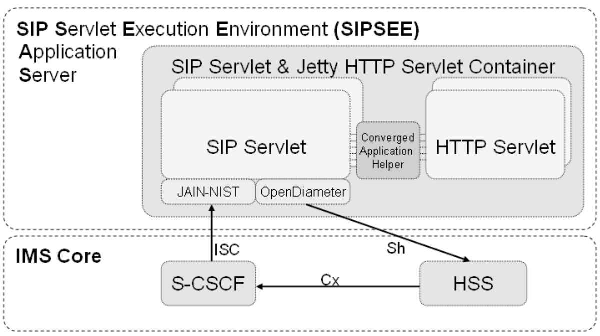


figure 3: FOKUS SIPSEE Implementation

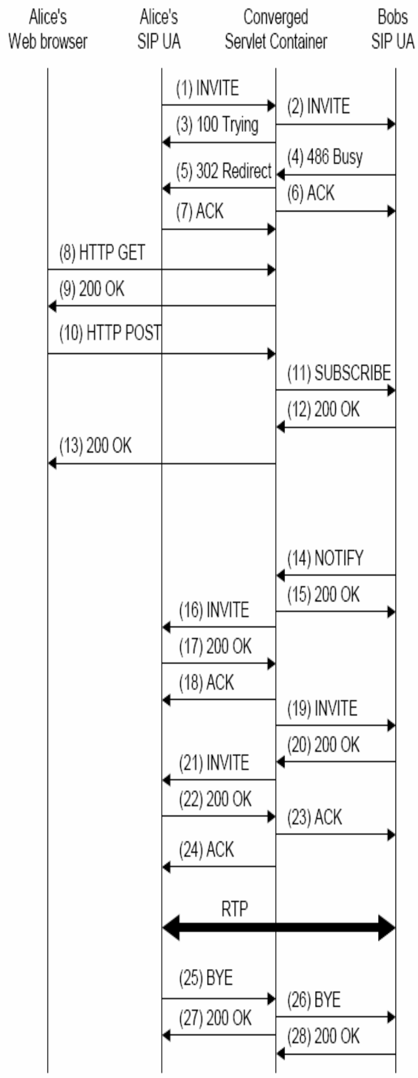
4 SIP AS application example, as implemented in the SIPSEE

This section provides an example for a converged application: Call Schedule on Busy or No Answer (CSBNA). In this example user A calls user B, but user B is busy or does not answer. So the intermediate SIP application server sends a message to user A redirecting him to a web page where a pre-filled form is displayed. So user A just has to send the form if he agrees with the input. The application may now register to user B to be notified when user B becomes available again so the application server may instantiate a call between user A and user B to establish a communication as requested by user A at the beginning. The steps in detail are as follows (according to example in SIP Servlet API Version 1.0):

1. Alice makes a call to Bob. The INVITE is routed to the converged container.
2. The CSBNA SIP servlet acts a B2BUA and sends an INVITE to Bob.
3. It also sends a 100 (Trying) informational response upstream.
4. Bob's phone returns a 486 (Busy Here) as Bob is currently in a phone call.
5. The CSBNA servlet returns a 302 (Temporarily Unavailable) response containing a single Contact with an HTTP URL pointing back to an HTTP servlet that is part of the CSBNA application.
6. The SIP application server sends an ACK to Bob.
7. Alice's SIP UA sends an ACK to the appl. server.
8. Alice's SIP UA launches a Web browser to retrieve the Contact URL.
9. The HTTP CSBNA servlet returns an HTML page containing a form allowing Alice to have a call automatically setup based on Bob's availability. The form is pre-populated by the servlet with all required information, in particular Alice and Bob's SIP addresses.
10. Alice hits the Submit button on the HTML page thus causing an HTTP request to be sent to the application server.
11. The service sends a SIP SUBSCRIBE request to Bob's SIP UA.
12. Bob's UA accepts the subscription.
13. The HTTP servlet receives the posted form and returns a Web page to Alice which she can subsequently use to modify or cancel the scheduled call.
14. Bob hangs up and his UA sends a notification of his new availability status to the CSBNA subscriber servlet in a SIP NOTIFY request.
15. The CSBNA application responds to the NOTIFY. (unsubscribe is not shown in the diagram)

16-24. The application establishes a call between Alice and Bob.

25-28. The call is terminated.



5 Summary & Outlook

The SIP Servlet API is an emerging specification, for which enhancements and extensions are in progress. There are issues to enrich the generic part of the servlet specification with parts only defined in the SIP Servlet API in order to standardise interactions between SIP and HTTP servlets. Thus an alignment of the generic servlet specification, the HTTP servlet specification and the SIP servlet specification is needed. This will lead to converged application servers which will not depend on proprietary ways of connecting HTTP servlets and SIP servlets. The "Open SIP AS" is the living proof of the efficiency of converged services, but the designers had to develop a proprietary interface to enjoy all the advantages of co-operational usage of SIP and HTTP. Consequentially applications have to be slightly modified according to the manufacturer's API if they want to be transferred from one SIP application server to another SIP application server. Ongoing work at Fraunhofer Institute FOKUS is to design for the "Open SIP AS" a graphical user/administrator interface for Rapid Application Development. The aim is to deploy new personalized services within a few clicks. All "Open SIP AS" efforts and achievements are embedded into the Open IMS project as the driver for Next Generation Networks. Within the 3Gb test- and development-centre at FOKUS software and applications are designed, implemented and validated. Due to all the technical advantages of IMS, there are no doubts that IMS will become a success story within operator networks. But telecommunications is a very disruptive industrial sector. Unfortunately, operators cannot buy customer loyalty by offering low prices in this highly competitive market. Attractive, customer-friendly and committing services based on innovative, efficient and powerful application servers attract new customers and help to retain existing ones.

References

[parlay] The Parlay Group, http://www.parlay.org/

[IETF CPL] J. Lennox, H. Schulzrinne, Call Processing Language Framework and Requirements, RFC 2824, Internet Engineering Task Force, May 2000

[JAIN] JSLEE and the JAIN Initiative, http://java.sun.com/products/jain/

[JSR 116] A. Kristensen and Expert Group, JSR 116: SIP Servlet API, Mar 2003

[IETF SIP] M. Handley, H. Schulzrinne, E. Schooler and J. Rosenberg, SIP: Session Initiation Protocol, RFC 3261, Mar 1999

[IETF CGI] J. Lennox, H. Schulzrinne, J.Rosenberg, Common Gateway Interface for SIP, RFC 3050, Jan 2001

[3GPP.x] The 3rd Generation Partnership Project, http://www.3gpp.org, TS 22.228, TS 23.218, TS 23.228, TS 23.278, TS 24.228, TS 24.229, TS 29.328, TS 29.329.

[NIST] National Institute of Standards and Technology, http://snad.ncsl.nist.gov/proj/iptel/

[FOKUS IMS] Open IMS @ FOKUS, http://www.open-ims.org/

[satnac04] T.Magedanz, D.Witaszek., K.Knuettel: "Service Delivery Platform Options for Next Generation Networks within the national German 3G Beyond Testbed"

[CCN2004] T.Magedanz, K.Knuettel "IP Multimedia Subsystem, a System Description for a comprehensive service and application oriented network architecture based on IETF protocols and paradigms." http://www.iasted.org/pastconferences.htm

[tridentcom05] T.Magedanz, D.Witaszek., K.Knuettel: "THE IMS PLAYGROUND @ FOKUS – AN OPEN TESTBED FOR NEXT GENERATION NETWORK MULTIMEDIA SERVICES", http://tridentcom.org