SIP: The Killer Application in the Future Converged Networks

M.K Sidiropoulos and M.D.Logothetis[1]
WCL, Dept of Electrical & Computer Engineering
University of Patras, 265 00 Patras, Greece.
Tel +30 2610 996 433 Fax: +30 2610 991 855
E-mail:msidiropou@upnet.gr, m-logo@wcl.ee.upatras.gr

***Abstract*- While technology continues to evolve rapidly around us and competition is more and more intensifying, telecom companies should search for innovative technology solutions that will enable them to offer to the market smart services. On the one hand these services should provide efficient and easily to handle solutions to the users, and on the other hand increase the company's revenues. This can be done by minimizing unnecessary costs and by deploying service-generating new technologies. Towards this trend SIP protocol can and should possess a major role. The time for the combination of the mobility provided by mobile networks and the Internet has come and therefore an evolution of a wide range of new services is inevitable. In this sense we call the SIP protocol killer application. This paper is an overview of the SIP protocol and its promising features. Moreover we give an emphasis on SIP's numerous applications and we conclude by considering the advantages of SIP's protocol from an economical point of view.**

Keywords: SIP, convergence, carriers.

I. INTRODUCTION

A. IP convergence

Over the last ten years the telecommucations industry has been the witness of an unprecedented change in the way people communicate. The evolution of the internet and the rapid growth of mobile communications have without doubt changed the way we conceive communications. The traditional circuit-switched telephone network is threatened by the emeging IP network and as a result the idea of converged networks, capable to deliver concurrently media of all types (voice, data, and video) any time, anywhere, appears to be a true challenge. But what is an IP converged network and how is it related with SIP?

IP convergence has come to blend together the two popular domains for communications the mobile and Internet domain. It uses packet-based networks as a common infrastructure, and it offers a rich end-user experience. This is achieved by the combination of voice, video, text, and content seamlessly in a person-to-person communication. Its ultimate goal is to integrate together the wide range of access technologies that have dominated the telecom market and offer to the end users a seamless connection. That is, wireless access technologies such as Wideband CDMA (WCDMA), CDMA, GSM, Enhanced Data Rates for Global Evolution (EDGE), and WLAN have to be combined with wire line technologies, such as DSL, in order to offer to users connectivity without boundaries.

However, this IP convergence platform, which makes possible numerous services, cannot stand alone. SIP (Session Initiation Protocol) is the enabling protocol. This specification (RFC 3261) [1], was released by the IETF in June 2001and because of its simplicity, power and extensibility, SIP was rapidly adopted as a Voice Over Internet Protocol (VoIP) standard.

SIP has been embraced by the IP telephony industry. Device manufacturers, software vendors, and service providers have started integrating SIP into their own products. Since November 2000, the 3GPP wireless initiative has adopted SIP as its protocol. Consequently, SIP has been the foundation for the IP Multimedia System (IMS) industry. IMS is an emerging standard architecture with tremendous potentials. It is an "open", multi-media service architecture for mobile and fixed networks and it uses SIP as a unifying signaling protocol for voice and multimedia sessions. IMS allows different services, e.g. Presence, Push To Talk, to share common components and enables carriers to develop a dynamic service environment full of new customized services quickly and economically.

Therefore, the challenge is unique. All what carriers have to do, is to adopt this so powerful toolkit and employ strategies that can generate new revenue streams.

B. Why convergence withSIP?

The answer to this question describes the whole concept of the goals that converged networks have come to fulfill. To be more specific a SIP-enabled network offers:

- opportunities for new revenue streams as it enables a plethora of enhanced services that were not previously possible or were too complex and expensive to implement.
- Peer-to-Peer communication. It is the key factor for achieving richer communication between two end-users. After the establishment of the connection there are a whole range of IP protocols that can deliver their services in the session. For example, two people can play an interactive game directly between terminals, making comments during the game, using voice and sharing files or video with each other.
- Numerous services, easily implemented. SIP co-operates with other IP protocols to offer services to the user. Therefore, the creation and integration of these

[1]member of FITCE

services is fast and easy. Moreover developers have at their disposal “open” SIP APIs. This allows applications to be developed and deployed much faster than in the past.
- Integration and interaction of services. Such services are voice, video, instant messaging, and presence. All can interact with each other to create services that are both beneficial and easy to use.
- Co-operation with the IMS. This way users will be able to access, create, consume and share digital content using interoperable devices. For example, a videoconference could be held with different participants using a fixed line connected to a PC, a mobile phone connected via a WCDMA network, and a laptop connected via WLAN.

C. *SIP and telecom operators: a first approach.*

The SIP protocol enables telecom operators to adopt strategies that:

1. Reduce costs.
2. Increase revenues.
3. Develop a flexible network, easily extensible.
4. Boost customers’ loyalty.

The above goals have been originated by the traditional circuit-switched telephony era. However, nowadays the evolution of technology has paved the path to networks that combine data, voice and video and therefore the above goals are feasible. This is easily justified if we think that converged networks, which main feature is voice over IP (VoIP) technology, are more economical to build and maintain than legacy circuit-switched networks, which have greater CAPEX and OPEX. That is because a VoIP network cuts down on:

- hardware costs by using servers.
- maintenance, training and administration costs.
- phone charges due to total or partial VoIP transmission.

According to a recent study conducted by research firm Integrated Research [2], 78% of large companies are already deploying IP telephony. The study has revealed that this trend towards an all-IP network is not necessarily due to the replacement of the old PBX systems because they didn’t work. The main drivers of this trend are IP applications and enhanced communication capabilities, such as IP-based video conferencing. This survey shows that companies are willing to exploit for their own benefit the enhanced capabilities that IP telephony provides over traditional telephony.

II. OVERVIEW OF SIP

A. *What is SIP?*

SIP is an application-layer signalling protocol that can provide all the necessary functionality to establish, modify and terminate sessions with one or more participants [1]. Its main role is to help session originators deliver invitations to potential session participants wherever they may be. These sessions can be voice (internet telephone calls), multimedia distribution and video sessions (conferences) across packet networks. Members in a session can communicate via multicast or via a mesh of unicast relations, or via a combination of these.

SIP supports session descriptions that allow participants to agree on a set of compatible media types. It also supports user mobility by proxying and redirecting requests to the user's current location. SIP is not tied to any particular conference control protocol. SIP provides the following functions:

- Ensures that the call reaches the called party wherever they are located and that details of the nature of the call (Session) are supported.
- Permits any feature negotiation. This will allow the group involved in a call (this may be a multi-party call) to agree on the features supported - recognizing that not all the parties can support the same level of features. For example, video may or may not be supported.
- Manages the number of participants, as during a call a participant can bring other users onto the call or cancel connections to other users. In addition, users could be transferred or placed on hold.
- Support possible feature changes. A user should be able to change the call characteristics during the course of the call.

SIP fulfils these functions and re-uses other web elements to make it flexible and scalable. That is, SIP is a request-response protocol which closely resembles the famous internet protocols HTTP and SMTP. For example, SIP uses URLs for addressing. Each user is identified through a hierarchical URL that is built around elements such as a user's phone number or host name (for example, sip:user@company.com). This means that it is just as easy to redirect someone to another phone as it is to redirect someone to a webpage. Furthermore, unlike HTTP and SMTP, SIP can run on top of either TCP or UDP, providing its own reliability mechanism. SIP can work with both IPv4 and IPv6.

In order to provide telephone services carriers should use together a set of different standards and protocols However SIP alone, can’t do everything. That is, SIP works with other IETF protocols to build a complete multimedia architecture and accomplish its job. Fig.1 illustrates the overall IETF SIP protocol suite. This suite apart from the well known protocols (IP, TCP, UDP, TLS and SCTP) includes protocols such as the Real-time Transport Protocol (RTP) (RFC 1889) for transpoting real-time data and providing QoS feedback, and the Real-Time Streaming Protocol (RTSP) (RFC 2326) for controlling delivery of streaming media. Note also that whereas the SIP protocol contains the parties’ addresses and protocol’s processing features, the description of the media streams, exchanged between the parties of a multimedia session are defined by the Session Description Protocol (SDP) (RFC 2327). SDP is in fact, not a protocol, but a structured, text-based media description format that can be carried in the SIP message. Since the message body is transparent to SIP any session description can be transferred. Therefore, SIP can be used in any kind of session and in conjunction with other protocols in order to provide complete services to the users. However, the basic functionality and operation of SIP does not depend on any of these protocols.

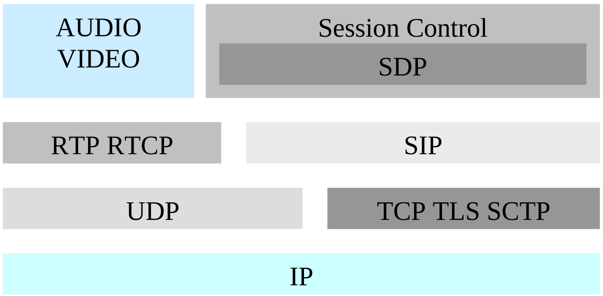

Fig1: SIP and other protocols

B. *Basic Features of SIP*

SIP's features are numerous. We highlight the following:

- It is simple meaning that it specifies only what it needs to specify. As it has been built as a pure mechanism that establishes sessions, SIP does not need to know about the details of a session. Its job is to just initiate, terminate and modify sessions. That's why it is so popular among different architectures and deployment scenarios.
- It is scalable, meaning that many people can participate in spontaneous, media-rich and interactive activities at the same time. Examples are: audio and video conferencing, application sharing and any other collaborative activity.
- It is extensible in the sense that its design allows for the addition of new features, while continuing to supporting existing ones. Extensibility stems from the fact that the intelligence involved in the SIP applications is located at the endpoints of the network and not in the core network. Therefore new features can be added without having to upgrade infrastructure components.
- It is flexible. Note that with the term flexibility we refer to a system that might support many features, or adapt itself to different conditions, but its total capabilities remain fixed. SIP is flexible, in the sense that it can be used to establish many different types of sessions and integrates well with the Web, e-mail, streaming media applications and existing protocols used on the Internet.

C. *The SIP Architecture*

The basic elements that form the buildings blocks of a SIP network are:

1. SIP User Agent (UA): any network end-user device that can act as both a user agent client and user agent server. That is, it can originate and terminate a SIP session. Such devices are: SIP-enabled telephones, SIP PCs (called "softphones"), cell phones, PDAs, etc.
2. SIP Proxy Server: An intermediary entity that acts as both a server and a client for the purpose of making requests on behalf of other clients. A proxy server primarily plays the role of routing, which means its job is to ensure that a request is sent to another entity "closer" to the targeted user.
3. SIP Registrar Server: is a database that contains the location of all User Agents within a domain and dynamically updates its data via REGISTER messages.
4. SIP Redirect Server: this element provides routing information to user agents, when requested and allowes SIP Proxy Servers to direct SIP session invitations to external domains.

The messages that SIP uses are called methods and are shown in the following table 1:

SIP Method	Description
INVITE	Invites a user to a call
ACK	Acknowledgement of an INVITE request
BYE	Indicates that one endpoint is terminating the session
REGISTER	Registers a user's current location
CANCEL	Terminates a request, or search for a user
OPTIONS	Solicits information about a server's capabilities

Table 1: SIP requests

D. *SIP applications and services*

SIP is a signaling protocol that differs from the traditional signaling protocol (ISUP) used in PSTN. This difference is mandated by the distributed nature of Internet telephony and the fact that Internet telephony is able to deliver combined services [3]. SIP can easily support converged services firstly because it uses MIME, the de facto standard for describing content on the internet, and secondly because of its ubiquitous support for URLs and URIs as an addressing format.

SIP provides an open platform, which leaves plenty of room for the development of a wide range of new applications. Some examples are: unified messaging, simultaneous ringing, and intelligent call routing to any device. In addition non-voice applications such as instant messaging, web collaboration and conferencing are also capable with SIP. The most notorious SIP applications are:

- Unified Communications: A SIP session can contain any combination of media (voice, data, video, etc). These sessions can be modified at any time by adding new parties or by changing the nature of the session.
- Unified Messaging: email, voice-mail, faxes and phone messages are accessible from the same box. Note that people use many different SIP devices for communication. Unified messaging promotes communication via different devices at any time.
- Directory Services: They are to a network what white pages are to the telephone system. They are used for storing information about various objects. People can use the service to search for objects by name, like the yellow pages. For a network manager directories are a mean to manage user accounts and network resources.
- IP-PBX functionality: Any IP-PBX that is compliant with the SIP standard can be utilized offering flexibility and options for future expansions.

- Third party Call control: It refers to the ability for a device that is not one of the SIP endpoints to affect a SIP dialog. Many services are possible through 3PCC. An example is Click-to-dial.
- Click-to-dial: It describes the ability for a hyperlink on a web page to commence a telephone conversation to the referenced destination. Consider for example a direct link from a company's website to a sales representative. The web server then creates a call between the two entities. The call can be between two phones, a phone and an IP-host, or two IP-hosts. .
- Presence gives a new dimension to communication. With presence callers can learn about the availability of the person they are trying to reach and how he wishes to be contacted - before they call is initiated to that person. For example, when a user is on a phone within his company network, his status ("on the phone") can be automatically updated and communicated to a central presence server. This presence information is then distributed to other users through the integrated telephony/IM contact list.
- Auto-conference application: A user views a web page containing a form which has the email addresses and phone numbers of the conference participants. Once submitted, the application waits until all participants are available, as determined through the presence service. When all are available, the application sends each participant an instant message. The message asks each participant if they actually wish to join the conference. It contains two hyperlinks, both of which resolve to the controller. One is clicked if the user wishes to join the conference, and the other, if they do not. If all click the hyperlink to join the conference within a specified time interval, the application rings the phone of each participant, and then connects each of them into a conference.
- Instant Messaging (IM): It provides the ability to send messages to other individuals. The underlying requirement is very similar to email. Instant messages are analogous to the sentences in a conversation, as they are short and expect a quick response. As an example of IM consider an IP phone which sends a note through the IrDA port as an Instant Message and can receive a quick response.
- Mobility: SIP client software can be embedded in mobile phones and PDAs so that these services can cross all platforms. With SIP sessions can be established between different devices that then negotiate the appropriate media capability.
- Wireless LAN VoIP Telephone Handsets: Portable telephone handsets support Voice over IP through an 802.11 wireless LAN connection.
- Push-to-talk: It is a Voice over IP streaming service that turns a mobile phone into a Walkie-Talkie, which operates smoothly in GPRS and UMTS networks. It provides end-users with the ability ro communicate over the wireless medium, one-to-one, or one to many, via a real -time always-on connection over the network. With a simple push of a key, users can initiate an instant voice conversation with any available users on their contact list. They are able to talk to one or to groups of people simultaneously or directly deposit a message into recipients' mailboxes, without having to call or to speak to them. PTT service uses cellular access and radio resources more efficiently than circuit-switched cellular cervices, as it reserves network resources only for the duration of the talk spurts instead of for an entire call session.

III. SIP: BUSINESS/FINANCIAL PERSPECTIVE ?

- *The financial impact of SIP's innovative services on telecom operators*

Today noone can deny the impact of SIP on both wireless and wireline carriers. However, it is still unknown how big this effect of SIP is going to be.

The first years of the launch of SIP-based products in the market were not so lucrative for the telecom industry. The impression that SIP was a simple protocol to develop resulted in the development of many SIP implementations, which were providing different levels of quality and completeness. As a result, were raised interoperability problems which could have "killed" the protocol. Moreover, the fact that traditional telephony services became a commodity, offered by many companies increased competition. Consequently, mobile services markets showed a severe decline in their voice ARPU (average revenue per user). That happened because initially carriers, saw SIP and VoIP in general, as a threat. However, this notion is changing. The strong features of SIP and the enhanced services that can offer give to all players in the telecom industry an immediate revenue opportunity.

There are numerous examples of devices and applications that employ SIP and are expected to shake the market with their sales. Consider for example the smart SIP-enabled phones which have already launched in the market and their popularity is steadily increasing. These smart phones are being sold at a greater price than the current dumb devices, but those costs will come down as volume production goes up. On both circumstances carriers will be benefited. These devices have the potential to communicate directly with the internet using IP-packets. Therefore, cellular operators can use as an alternative solution IP-routers in their networks instead of using the expensive, proprietary circuit-switched technology.

Moreover, a new family of phones, named dual-mode phones is entering the market. These handsets are SIP compliant and can use either the cellular network or VoIP over WiFi/3G/WiMAX, etc. As Wi-Fi capability is becoming a common-feature in cell phones we strongly believe that carriers have a lot to earn if they pursue this share of the market. A recent survey has shown that the number of cellular/VoWLAN subscribers will reach 256 million worldwide by 2009. That's 12 percent of the total cellular subscription base. In that year, the number of people using VoWLAN will surpass the number using Wi-Fi for data only.

However, it is the application that makes any SIP device popular. Consider for example, push-to-talk, which is expected to be the voice form of SMS (which until now, without doubt, is the killer application as it has generated streams of revenues). Carriers believe that the teen and young adult populations will be a big market for this capability. Moreover, PPT is believed to be the replacement of the traditional walkie-talkies that families used to stay in touch, and therefore PPT is seen to have a big potential market. In the USA mobile operators have already seen that PTT users generally spend more money every month and are more loyal than people just using voice services.

Apart from the plethora of SIP products and applications, it is the type of the network that SIP fits in, that makes it so popular. To manage a converged infrastructure, makes life easier for carriers, as they are now able to drive costs down. This is justified if we think that by this way any administrative and management complexity is reduced, as converged networks provide immediate cost reduction in both staffing and maintenance of hardware and software. Let's think for example of VoiceXML application. When it came out, as it was based on web technologies, it gave the user the ability to separate the application for the telephony infrastructure.

Additionally with the replacement of the traditional circuit switched trunks costs declined. The major driver for that was that deploying new services requires significantly lower investments in time and money in a converged network than in legacy TDM systems. By this way carriers could significantly reduce their CAPEX and OPEX. That is because, in a SIP-enabled network, intelligence is no longer restricted into the PBX, but instead it has moved to the end-points. As a result employees are now able to move and use any plug-and-play SIP-based end-user device. Therefore new applications are replacing existing expensive services and infrastructure. For example, IP conferencing is able to reduce costs compared to the high costs of existing conference-call services. It is the flexibility of SIP that allows for the addition of new features on the end-user devices without having to upgrade infrastructure components such as proxy servers. Hence, developers do not need in-depth knowledge of the SIP infrastructure in order to write SIP-enabled applications.

This opens up the application development process to more players in the market and productivity is increased. Although, by this way competition also increases, SIP enables the development of innovative services in less time and at less cost. Therefore there is a kind of price/mass-production trade-off. On the one hand increased competition drops the prices of SIP devices, but on the other hand an innovative service that has conquered the market can grow revenues.

Carriers can also aim at enterprises to offer their services and boost their profits. As an example, consider that SIP trunking services can allow two SIP-empowered enterprises to communicate from endpoint to endpoint without the need for a gateway between them. The explosion of services of this kind helps on reducing costs, as the load on packet switched gateways is now reduced.

- *Possible methods to increase revenues*

Today's business landscape is rapidly changing. Traditional carriers offering voice and mobile services face threats from new players who are entering the telecom market and by utilizing new technologies bypass legacy telephony and reduce carriers' revenues. This poses the question of how carriers will manage to gain back the large portion of the revenue-rich traffic which seems to be sliding out of their control. There are many methods to confront this trend. Some of them are:

1. Search for new cutting-edge services that can deliver significant business gains in the areas of productivity, efficiency and customer satisfaction. Carriers should look for new SIP-based services that generate revenues. These new services might take quite some time to be absorbed by the market.
2. Cutting costs by improving operational efficiency: This is feasible by replacing existing services, such as, PSTN/PBX telephony with IP voice communications. Rich end-to-end controlled IP communications can be introduced and supported at a cost that is truly negligible compared to the huge investments and operational cost required by traditional PSTN networks. Note also that improved operations means that companies will profit more because of the success of their products.
3. Another operational win is to eliminate any possible inaccurate billing processes: this will increase customer satisfaction and most importantly will put an end to carriers' losses. Note that SIP was not designed to provide a direct source of billable usage data. It is an event-based protocol that produces many reporting events for a single usage session. Therefore the flood of SIP events for a single session could end up costing the service provider more for the billing system to process than it receives back in revenue.
4. Improve customer loyalty: This can be achieved by delivering services that enable mobility, and feature consistency in a continuously competitive marketplace.
5. By a slow migration of the services offered by carriers to an IP backbone. A company can migrate its communications to IP on a site-by-site or application-by-application basis. This initiative will lower overall cost through the elimination of multiple service technologies and will set the basis for a future where all services will be delivered in IP form as part of the fixed mobile convergence (FMC) trend. It seems that carriers have no alternative other than embrace FMC, and find a way to profit from it.
6. By using technology solutions that allow carriers to charge subscribers when they connect to services via third-party network access points such as WLAN hotspots. On the other hand users enjoy a high quality connection through seamless handovers between the WLAN and the mobile network.

IV. CONCLUSIONS

In this paper we presented a discussion about the role of the SIP protocol in the future converged networks. We highlighted the drives that led carriers to its adoption and the needs that SIP came to meet. In addition, we gave emphasis on the most compelling services and applications that SIP enables. Finally, we described how telecom market is going to be financially benefited from SIP and proposed some possible methods that can return profits back to carriers. Our belief is that SIP can pave the way towards an all-IP network, where seamless communication will be a reality.

REFERENCES

1. J. Rosenberg, et al, "SIP: Session Initiation Protocol", RFC 3261, IETF, June 2002.
2. Integrated Research : www.asx.com.au
3. J. Rosenberg , "Distributed Algorithms and Protocols for Scalable Internet Telephony",Columbia University, 2001
4. White Paper: IP Convergence Based On SIP - Enhanced Person-To-Person Communications, Nokia, www.nokia.com
5. White Paper :IP multimedia –a new era in communications, Nokia www.nokia.com
6. White Paper: IXIA: Session Initiation Protocol (SIP) Technology.
7. White Paper: SIP Market Overview
8. White Paper: Evolving to Converged Communication with Session Initiation Protocol (SIP). www.avaya.com
9. www.sipcenter.com
10. www.tmcnet.com
11. www.sipforum.org

Michael K. Sidiropoulos was born in Athens, Greece, in 1979. He is a graduate student of the Electrical & Computer Engineering Department, University of Patras, Greece, and is going to receive the Dipl.–Eng. Degree in Electrical & Computer Engineering, in June, 2006. Currently, he is carrying out his diploma thesis at Wire Communications Laboratory of the aforementioned Department. His main scientific interests are on Wireless LAN and related Protocols. He is a member of the IEEE Student Branch, University of Patras.

Michael D. Logothetis was born in Stenies, Andros, Greece, in 1959. He received his Dipl.-Eng. degree and Ph.D in Electrical Engineering, both from the University of Patras, Patras, Greece, in 1981 and 1990, respectively. From 1982 to 1990, he was a Teaching and Research Assistant at the Laboratory of Wire Communications, University of Patras, and participated in many national and EU research programmes, dealing with telecommunication networks, as well as with natural language processing. From 1991 to 1992 he was Research Associate in NTT's Telecommunication Networks Laboratories (Tokyo). Afterwards, he was a Lecturer in the Department of Electrical & Computer Engineering of the University of Patras, and since 2003 he is an Associate Professor in the same University. His research interests include traffic/network control, simulation and performance optimization of telecommunications networks. He is a member of the IEEE, IEICE, FITCE and the Technical Chamber of Greece (TEE).