Experiences with Blending HTTP, RTSP, and IMS

Sohel Q. Khan, Sprint-Nextel
Robert Gaglianello, Alcatel-Lucent Technologies
Michael Luna, Openwave Systems, Inc.

The concepts presented are authors' opinion and do not necessarily reflect Alcatel-Lucent, Openwave, Sprint-Nextel, and Comcast technological directions.

Dr. Khan was at Sprint-Nextel when the manuscript was submitted. Now he is at Comcast.

ABSTRACT

Although the Session Initiation Protocol (SIP) based IMS network has attained attention for IP voice and video telephony networks, the popular and prevalent Web and streaming video applications are supported by real time streaming protocol (RTSP) and HTTP proxy networks. The current trend of maintaining disparate SIP-based IMS, RTSP, and HTTP proxy networks introduce added cost due to the duplication of network components, separate maintenance functions, and scaling inefficiency. Service providers' requirements to reduce such costs, call for the development of a solution that integrates the three proxy networks into a single hybrid platform. This article depicts the vision of such a hybrid platform, exemplified by two blended prototypes — shared streaming video application and the multimedia proxy. It is through the development of such applications that a path to an enhanced or extended IMS architecture encompassing the three proxy networks is possible. This article describes approaches promising the synthesis of IMS, HTTP, and RTSP based networks into integrated platforms that pioneer convergence to the vision.

This article investigates a future, possible expansion of IMS to harmonize SIP, HTTP, and RTSP service delivery. The objective of this article is to initiate discussion and research.

INTRODUCTION

IMS [1, 2] promises to enable rich blended services such as voice, video, data, and multimedia applications within a converged IP platform. Such services would simplify the communication and entertainment experience of users and create a value-based market for service providers. IMS intends to enable the creation of innovative personalized applications and the faster deployment of these applications.

Although the current IMS standard defines mechanisms for initiating and negotiating sessions that contain multimedia content, it has not specified a well defined generic coordination point [3] to blend voice, video, data, and multimedia applications and services. This lack of coordination has lead to management of separate service specific networks that traditionally are deployed in parallel.

Application decomposition for integration with the IMS environment is important for the development of commercially viable services. However, the current application development environment consists of many different, but capable standards that are not included in the IMS. Thus, it is important that techniques be developed for including these in IMS applications. Many vendors are developing various techniques to integrate RTSP and HTTP based services within the IMS platform. In this article, we illustrate only two examples — the shared streaming video (SSV) [4] application prototype and multimedia proxy (MMP) prototype [5].

The organization of this article is as follows. We start by presenting an overview of the current network architecture and its merits. We then depict our envisioned methods for hybrid networks: an extended IMS architecture and an enhanced IMS architecture. Next, we discuss the shared streaming video (SSV) application prototype. Then we describe an alternate approach using the multimedia proxy (MMP) prototype. Both solutions focus on blending IMS and non-IMS based services to provide a seamless user experience. We conclude with observations of both prototypes and possible directions for future work.

Note that this article does not include a discussion of security architecture.

CURRENT APPROACH

The current deployment strategy for a generic wireless provider is to maintain separate signaling networks for push to talk (PTT), voice over IP (VoIP), video streaming, and Web applications. Due to the high delay sensitive requirements, wireless providers either deploy a proprietary signaling protocol supported distinct

 0163-6804/07/$20.00 © 2007 IEEE

PTT network or a SigComp/SIP supported PTT over cellular (PoC) network. For the VoIP and Video over IP telephony network, the trend is to deploy a SIP-based IMS network. An RTSP proxy-based network supports video streaming, video playback, and video on demand. An HTTP proxy-based network services wireless Web-based applications.

These networks operate in parallel. Each network has its own subscriber profile databases, signaling functions (i.e., signaling proxies), application policy filter, media transfer function, billing function, and security function. One notable advantage of current deployment is that each network can be fine-tuned to meet the application type. Another advantage is that separate departments can maintain each network. The notable disadvantage of maintaining separate networks is the cost of redundant equipment and operational maintenance. In addition, a handset client has to register in the databases of each network and must perform authentication for each service type in each network. Maintaining concurrent services, coordinating across services, performing harmonious scheduling, and providing seamless application mobility are difficult.

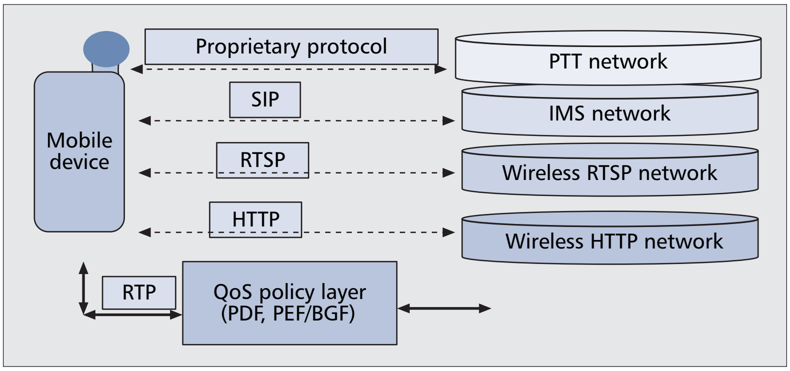


Figure 1. *Current parallel networks.*

EXTENDED IMS ARCHITECTURE

Functions of an IMS network can be subdivided into subscriber profile function, signaling functions, QoS policy function, media function, and charging function. Subscriber profile function is a user subscription profile database known has Home Subscriber Server (HSS). The signaling function of an IMS network is based on the Session Initiation Protocol (SIP) and Session Description Protocol (SDP). The signaling functions consist of SIP servers mainly performing authentication, registration, location discovery, call routing, and call re-direction. The signaling functions are referred to as CSCFs (proxy, serving, and interrogating call session control functions). The QoS policy function creates and enforces QoS specific policies. In some IMS infrastructure, they are referred to as policy decision function (PDF) and policy enforcement function (PEF). The charging function either performs on-line or off-line accounting for billing purposes.

In some IMS implementations, the media functions are combined with QoS policy functions in IP routers. In this article, we will refer to it as QoS policy and media function. Some providers deploy a media transfer function that performs tasks such as transcoding, translating, and framing media to fit in the wireless handset. We will refer to this function as media transfer function.

Current IMS applications are mainly voice and video telephony centric. There are a few vendor and standard initiatives to extend IMS by adding an application policy function that interfaces with RTSP and HTTP based applications and enforces RTSP, HTTP, and WAP (Wireless Application Protocol) application specific policies. For example, it would filter application requests and deny access to applications that are not subscribed to or prohibited. The application policy function enables providers and subscribers to create, manage, and implement policy rules such as parental controls, privacy controls, and other application sharing polices. In some marketing circles, this is referred to as service delivery platform and would sit above the S-CSCF and span other network or network access technologies. Note that these application policy rules are in addition to the SIP-based filter criteria (a.k.a. *initial filter criteria*) specified in 3GPP TS 23.218 [6].

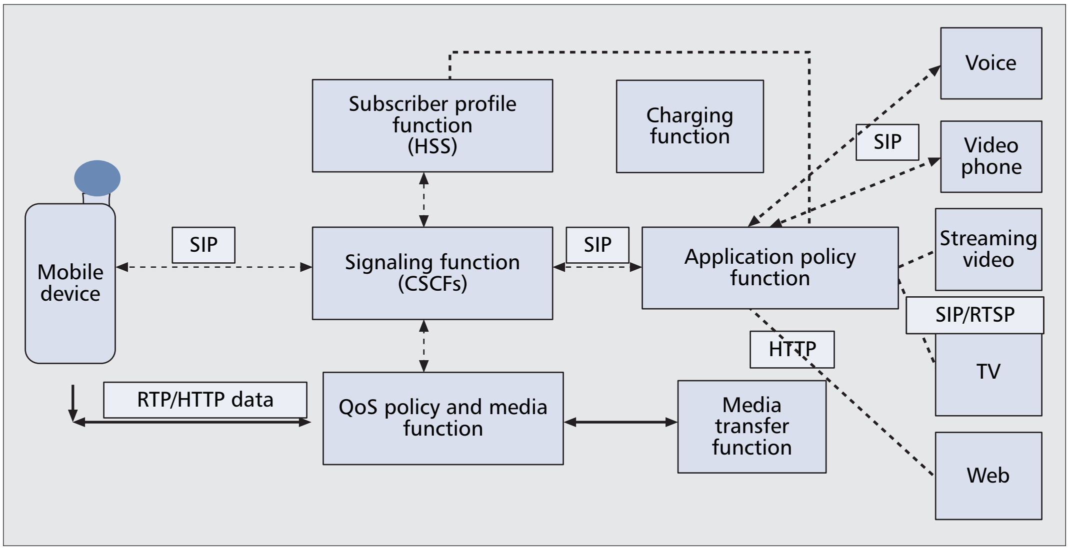
The extended IMS architecture appears in Fig. 2 and promises to provide streaming video, IP TV, and Web applications from a converged platform.

Some advantages of this architecture are maintenance of single platform, ease of service creation, reduction of redundancy, improvement of architectural scalability, and support of seamless application mobility. Some disadvantages are degraded signaling load scalability, multiple points of provisioning and support, and a degraded customer experience due to inconsistent content rendering.

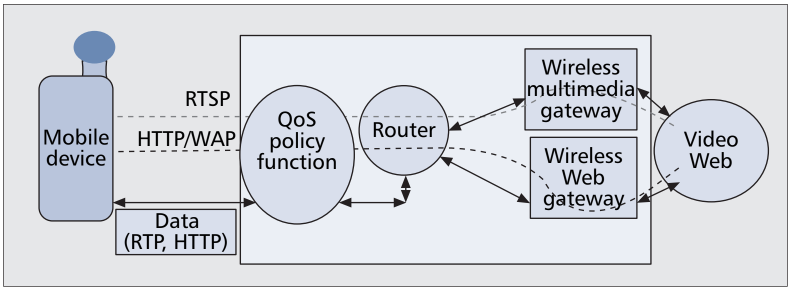
Note that an element of the application policy function in Fig. 2 can be the service capability interaction manager (SCIM) [6]. It is an entity for managing interactions among SIP application servers and between SIP features and legacy signaling system components. It invokes service logic as per SIP request. The 3GPP/3GPP2 designated the SCIM as a functional element within the IMS that sits between the S-CSCF and application servers. The SCIM also can be implemented as a stand-alone function. It has the opportunity to modify the messages of a session/flow as required and can understand different protocols such as HTTP and RTSP. Thus, it can be used to enhance the underlying functionality by enabling the interaction between different standards such as HTTP and RTSP with the SIP based IMS core. Capabilities of a SCIM include support for multiple dialog sessions, dynamic blending of applications services, and personalized feature interaction management.

ENHANCED IMS ARCHITECTURE

The IMS architecture is well standardized to provide VoIP and video over IP telephony services. However, RTSP and HTTP proxy-based wireless/wireline architectures are not well standardized yet. At a high level, an RTSP and HTTP/WAP centric architecture is shown in Fig. 3.



■ **Figure 2.** *Extended IMS architecture.*



■ **Figure 3.** *Current wireless/wireline video streaming and wireless/wireline web architecture.*

A close study of the wireless/wireline multimedia gateway (WMG) and wireless/wireline Web gateway (WWG) reveal their functional similarity with that of IMS architecture. The WMG and WWG were developed prior to IMS standardization and deployments. The WMG and WWG respectively, contain RTSP and HTTP proxies. Both WMG and WWG contain subscriber profile databases that are similar to the HSS in an IMS network. Both contain authentication, authorization, and accounting functions. Analogous to the application policy function depicted in the extended IMS network (Fig. 2), both WMG and WWG contain application filter functions that control user access to application servers and provide appropriate warning messages. Thus, there is an opportunity to map these functions with the extended IMS functions of Fig. 2.

We depict an enhanced IMS architecture that harmonizes SIP, RTSP, and HTTP network functions as depicted in Fig. 4. A subscriber profile function contains SIP, HTTP, and RTSP specific subscriber profiles. This blended approach requires augmentation of the current IMS HSS to support a generic, flexible (e.g., changeable schema) directory service to support generalized subscriber subscription information. The signaling function combines SIP, RTSP, and HTTP proxies onto either the same platform or separate co-located platforms invoked by a unique multiplexer. Notice that HTTP is a stateless protocol. On the other hand, SIP is a stateful protocol. This difference in state management must be accounted for when designing an enhanced IMS network. Application policy related functions of IMS, RTSP, and HTTP networks are combined into a unique application policy function. An element of the application policy function can be the SCIM [6].

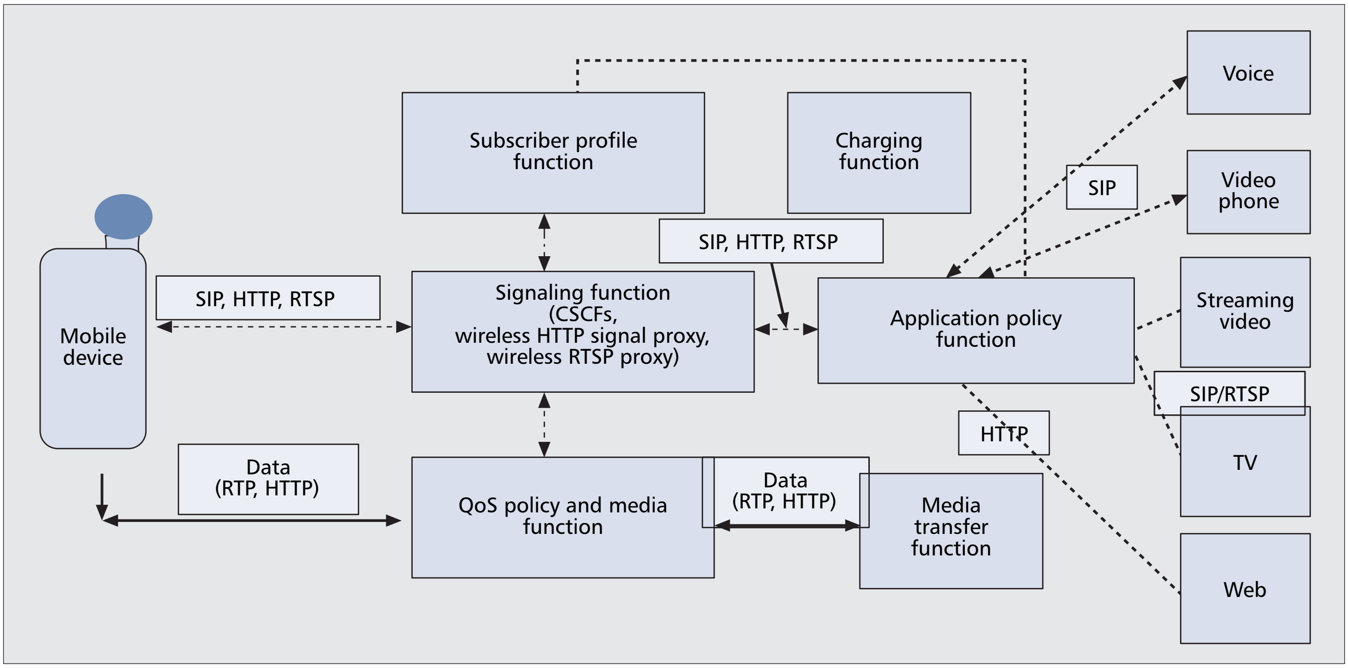
This synthesis enhances service creation, seamless service continuity, and concurrent service availability. It enables wireless and wireline providers to deliver coordinated "IMS ready content," Web, and video streaming content from a single platform; thus, eliminating duplicate components in the network. It also reduces duplications of border security and privacy devices. It enables single point coordination for charging, monitoring, and operation-maintenance.

In an evolution towards an enhanced IMS, many vendors are developing solutions to synthesize IMS-based applications with HTTP and RTSP based applications. We only describe two current blended prototypes — shared streaming video application [4] and Openwave's multimedia proxy [5].

SHARED STREAMING VIDEO APPLICATION PROTOTYPE

The shared streaming video (SSV) [4] application prototype enhances IMS VoIP calls by enabling users to share concurrent video applications. The SSV glues SIP-IMS with non-IMS HTTP and RTSP based applications.

The prototype SSV application uses SIP/IMS for call setup and session management, RTSP for streaming video, and HTTP to create a rich user interface for conference control sharing the streaming video. There are several extensions that would enhance the prototype. Shared Web browsing and click-to-dial functionality integration into the SSV prototype would add to a user's experience by enabling the user to make a



■ **Figure 4.** *Enhanced IMS architecture.*

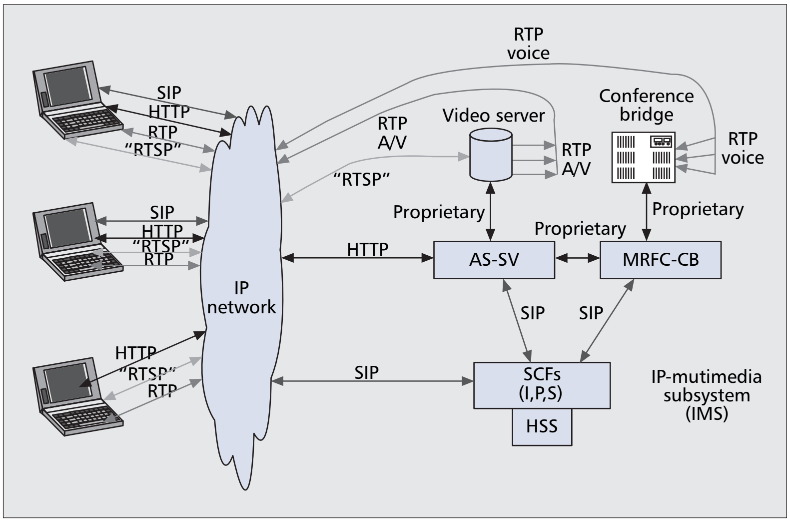
phone call by clicking a phone number located within a browsed page. Other extensions include creating a SIP based technique for general purpose sharing, enabling applications such as the sharing of Web pages, slide presentations, and documents.

The IMS architecture is defined by the 3rd Generation Partnership Project (3GPP), the European Telecommunications Standards Institute (ETSI), and the Parlay Forum. The SSV prototype contains decomposed 3GPP IMS services architecture [7] units, non-3GPP IMS functional units, and hybrid elements of shared-video application servers and conference servers. The decomposition and integration of non-3GPP-IMS elements highlights the need for enhancements to IMS and illustrates the success of blended applications.

Figure 5 shows this functional decomposition, with the data and control paths for the SSV application. Each arrow is labeled with the protocol "spoken" on that channel. On the right side of the diagram are the 3GPP IMS elements, the non-3GPP-IMS elements, and the hybrid elements. The hybrid elements consist of a SIP interface and one or more non-IMS standard interfaces. On the left side of the diagram are the client end-points. Connecting both sides is a standard IP network infrastructure. It should be noted that the functional elements do not necessarily reside on separate physical entities.

The 3GPP-IMS elements in this prototype are the call session control function (I, P, and S CSCFs) and the Home Subscriber Server (HSS) database.

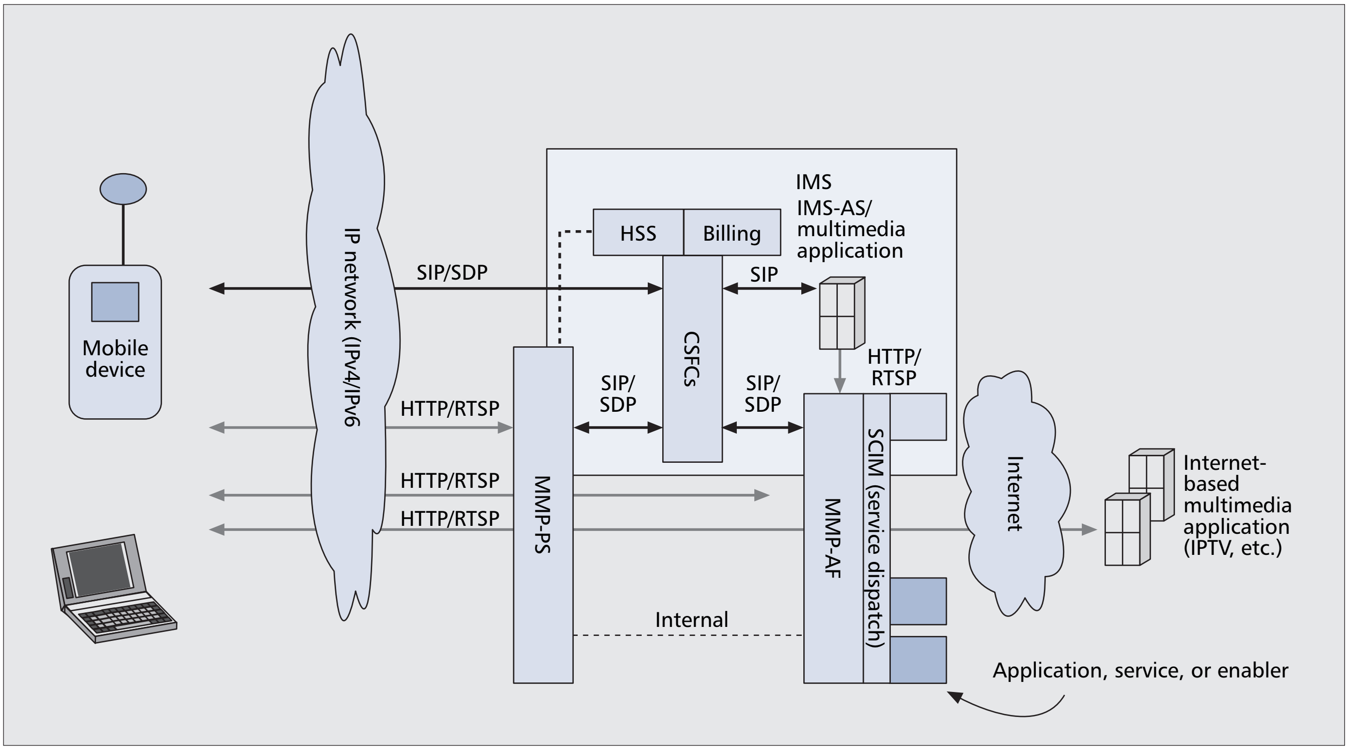
- The CSCFs and the HSS database comprise the call/session control layer. The CSCFs provide the registration of the endpoints and routing of the SIP signaling messages to the application server (AS-SV) and MRFC-CB (media resource function controller-conference bridge). The HSS database maintains profiles for each user. The user's service profile stores all of the user service information and preferences in a central location.



■ **Figure 5.** *Shared streaming video architecture overview.*

The hybrid elements are the shared-video application server (AS-SV), and the MRFC-CB.

- The AS-SV is the heart of the application in that it implements the logic/code of the application and handles the interactions between the application and the clients. It interacts with the various functional elements using SIP, HTTP, and several proprietary protocols. It includes a SIP stack, an HTTP Web server, and application specific code to tie all these pieces together. It has the responsibility to control the video server and communicate with the MRFC-CB.
- The MRFC-CB provides a SIP communications interface to the conference bridge



■ **Figure 6.** *MMP prototype architecture.*

using a proprietary protocol. Due to implementation time constraints, the MRFC-CB is implemented as part of the AS-SV, and as such, it is easier to implement the interactions between these two elements through a proprietary interface. The AS-SV and MRFC-CB function similarly to gateways, in that they have a SIP interface for interactions with the IMS elements and a media specific interface for controlling various media resources.

The non-3GPP-IMS elements in the prototype consist of video servers and the conference bridge servers:

- The video server provides synchronized streamed video content to the clients.
- The conference bridge is the media-handling device that sums the incoming Real-Time Transport Protocol (RTP) voice streams and distributes them to the users. As mentioned previously, an internal proprietary interface is used to interact with the MRFC-CB.

Readily available, off-the-shelf software packages for softphone, video server, and video player were used in lab trials that exhibited enriched user experiences in various scenarios.

In the first scenario, two clients are in an existing point-to-point VOIP call. During the conversation, client-1 wishes to show client-2 a video. The SIP phone interface of client-1 is augmented with a Share-Video button. Clicking this button causes a Web browser to start on the client-1 end-point. A page appears that contains information on various video clips available for sharing. Client-1 chooses the appropriate video clip and clicks it. A video player is started on both clients, and the video, including the audio track, is simultaneously streamed to both clients. The voice call remains active during the playing of the video enabling both clients to interact via voice. Upon completion of the video, client-1 can request another video by choosing another selection from the Web page.

The second scenario is an extension of the previous one, where two people connected in a point-to-point call and sharing a video, decide to call a third person (client-3). The initial SSV Web page that comes up on the client-1 end-point when the Share-Video button is clicked, also contains a set of buttons that enable the migration of a point-to-point call to a conference bridge and having the conference bridge call a party to join the conference. Client-1 simply enters the client-3 phone number into the Web page form and clicks the Call-Third-Party button. When client-3 answers the call, the video shared between clients 1 and 2 is automatically shared with client-3 and is synchronized with the others. Also, the Web page that now appears on the client-1 screen includes conference status information as well as the video sharing information. Conference status includes the number of participants currently connected to the conference and their identities.

Several variations on these scenarios are also possible such as when two or more clients are already connected to the conference bridge and decide to call another party. This function is similar to the second scenario except that the migration step is not required. The system also can function as a voice-conferencing facility only, independent of its video-sharing capabilities.

The notable point in these scenarios is the central role of the Web browser and SSV Web pages. A single button (Share-Video) invokes the SIP interface. All further interactions are performed through the browser and SSV Web pages. This splitting of the functionality is crucial to the rapid deployment of such applications.

It is also noted that a current example method of enhancing the IMS is to use a SCIM called the service broker [8] — that makes coordination and control of such multi-standard applications tractable by functioning as a mediator and blender between the various entities. Since it sits between application services and the S-CSCF (serving-call session control function), it enables seamless blending without modifying the individual applications involved in a session. The service broker is analogous to the application policy function of Fig. 2. The SSV does not include a service broker at this time.

MULTIMEDIA PROXY PROTOTYPE

Multimedia proxy (MMP) prototype [5] architecture blends SIP-IMS network with HTTP/RTSP proxy-based network into a converged platform. The MMP leverages the IMS functions for SIP session establishment, subscriber profile management, charging, and QoS policy management. Thus, it uses IMS CSCFs, HSS, charging functions, and PDF/PEF as illustrated in Fig. 6.

MMP has two major functions: the proxy service (MMP-PS) and the application function (MMP-AF).

MMP-PS performs the analogous functions of the signaling function and the QoS/media function of the extended IMS architecture depicted in Fig. 4. The MMP-PS is comprised of an integrated RTSP proxy, HTTP proxy, SIP proxy, and back to back user agent (B2BUA). The B2BUA mediates SIP sessions between IMS and HTTP/RTSP based clients. The integrated proxy interfaces with a common charging function, subscriber profile function, and service management function.

MMP-AF performs functions similar to that of the application policy function of the extended IMS architecture depicted in Fig. 4. The MMP-AF provides a central point that filters and applies policies associated with coordination, dispatch and orchestration between services, enablers, and applications. Some examples of these services are access to media authorization, throughput management, subscriber privacy polices, and billing management for the delivery of blended applications in a multi-network environment. The MMP-AF is comprised of an authorization engine, a real-time transformation engine, a policy creation and management engine, and a service creation and dispatch platform. The authorization engine manages access to content and services. The real-time transformation engine converts multimedia content and session requests required by client or application services. The policy creation and management engine enables the creation of services and enablers and manages orchestration policies that can apply to an individual subscriber, an entire network, or a market segment. This engine can be used to augment polices stored in the HSS if required. The service enabler platform provides an open and extensible architecture for the creation of a suite of services and enablers that can enhance and manage blended services. Examples of such services and enablers are the management of market segmentation, content categorization, virus prevention, or other third party services. The MMP-AF is placed between SIP-IMS based multi-media services and HTTP/RTSP-based services. The MMP-AF interfaces with HSS (through the MMP-PS) to access subscriber policy and subscription information and the S-SCSF for service orchestration performing the function of the IMS SCIM.

By combining the proxy functions that currently exist in different networks into the IMS-SIP controlled domain with MMP, it is possible to create many multimedia services within the IMS environment without the need for separate networks or proxy management. Further, it is possible to apply the same set of enhancements, such as transcoding and access management from a client outside the IMS domain. The MMP architecture also improves the reliability and management of the system by providing a single point of maintenance. By consolidating services into a converged platform, the MMP also ensures a consistent user experience under different network conditions by providing the same set of services for delivery of HTTP and RTSP services in the IMS.

Through the use of the B2BUA concept, the MMP architecture considers the requirement to manage access to these types of services when subscribers share these applications across their community, which creates a need to reach them from a non-IMS client (e.g., lifestyle or community services), or to access them from a non-IMS domain. This means that the MMP must have a solid foothold in both domains. For example, consider that a video streaming service in arbitrary format is located on the open Internet on a server that cannot directly support SIP or SDP-based session management. Further, consider that the subscribers involved are teenagers (either in a group or point to point session), communicating via an instant messaging (IM) service and want to share or view a video during the session. When user 1 requests to share a video stream with the group, the MMP can determine that one or more of the members of the group are not IMS capable and bridge them into the IMS environment through a B2BUA. Since the group can consist of IMS capable and non-IMS capable clients in the transaction, there is a new set of policy issues that must be considered. These include transformation, identity management, and potentially, age verification. All of these must be applied and enforced by both the signaling function and application policy function, some specifically to the IMS domain and some specifically to the existing HTTP/RTSP domain. For example, each of the clients involved may support different content formats and protocols and thus, different policies. Consolidation of the policy management into a converged platform using MMP enables them to be seamlessly applied and managed from within the IMS domain.

By combining the proxy functions that currently exist in different networks into the IMS-SIP controlled domain with MMP, it is possible to create many multimedia services within the IMS environment without the need for separate networks or proxy management.

Because the MMP prototype sits directly in the data path, it is easily integrated into the existing IMS without significant change to the current standard. However, the MMP architecture requires several touch points in both the existing and IMS domains that need to be considered.

Conclusions and Future Work

We presented an extended IMS architecture and an enhanced IMS architecture that harmonize SIP, HTTP, and RTSP proxy-based networks. Both of these architectures may include SCIM [6]. In an evolution towards an enhanced IMS architecture, we portrayed two current blended prototypes — shared streaming video (SSV) application [4] and multimedia proxy (MMP) [5] architecture.

SSV is a prototype application that provides rich user experience by blending SIP-IMS, RTSP, and HTTP centric applications. The MMP prototype is a way of enhancing the IMS architecture to harmonize HTTP and RTSP services in both the existing and IMS domains. Either enables the creation of many multimedia services that can be shared within the entire subscriber community while avoiding the current siloed or disconnected data delivery model.

The MMP architecture provides a central point for the application of policy. It also orchestrates services and enablers in both the IMS and existing domains. These policies, services, and enablers are common to the delivery of all types of blended services while ensuring a consistent user experience. Because the MMP prototype sits directly in the data path, it is easily integrated into the existing IMS without significant change to the current standard. However, the MMP architecture requires several touch points in both the existing and IMS domains that need to be considered:

- Diameter interface to the HSS subscriber service and billing management.
- Harmonization of the subscriber database in the HSS with that of the existing domain to provide a single point of policy management and provisioning.
- Extension of the HSS to include a flexible, application independent, directory model to define service, subscription, and subscriber policy. This includes the need to have and manage arbitrary service specific-schemas.
- Interaction with the IMS call session control functions to extend the orchestration of services and enablers that exist beyond the IMS network to enable blending with existing applications, services, client technology, and user preferences.

There are numerous benefits and several possible pitfalls to enhancing IMS. Some possible pitfalls of an enhanced IMS are:

- Updating standards can be a time consuming process and applications created before the standard is complete may require re-implementation.
- Waiting until the standards are complete could delay the deployment and development of new applications.
- The expense of converting to a new protocol.
- Incompatibility problems while industries retain the original protocols.

Examples of the benefits of an enhanced IMS are:

- Enables a consistent, seamless application mobility experience of SIP, RTSP, and HTTP-based applications.
- Extends the value of existing services into the IMS domain and IMS services into the existing domain, providing accelerated return on investment for the IMS investments.
- Scales the IMS offering to applications and services available on the Web today.
- Avoids maintaining parallel networks mode by creating a consistent, centrally managed, policy enforcement point that is applicable to segment management, regardless of domain. Eliminates the requirement of duplicate functions, for example, one subscriber profile function can be used for three networks — SIP, HTTP, and RTSP based networks.

References

[1] M. Poikselka *et al.*, *The IMS: IP Multimedia Concepts and Services in the Mobile Domain*, England: Wiley, 2006.

[2] Gonzalo Camarillo *et al.*, *The 3G IP Multimedia Subsystem*, England: Wiley, 2006.

[3] "Are There Issues with the Technology?", IMS Guide, Light Reading, Mar. 2005; http://www.lightreading.com

[4] R. Gaglianello and D. Liu, "IMS Shared Streaming Video," *Bell Labs Tech. J.*, vol. 10, no. 4, 2006, pp. 71–75.

[5] OpenWave, "Multimedia Proxy (MMP) Prototype," http://www.openwave.com

[6] 3GPP, "IP Multimedia (IM) Session Handling; IM Call Model, Stage 2," TS 23.218, V6.3.0, Mar. 2005.

[7] 3GPP, "Service Requirements for the IP Multimedia Core Network Subsystem," TS 23.228 Stage 2, 2006.

[8] Alcatel-Lucent, "Service Broker," http://www.belllabs.com/enablingtech/pdf/service_broker.pdf

Biographies

ROBERT GAGLIANELLO (rdg@alcatel-lucent.com) received a B.S. degree in electrical engineering and a B.S. degree in physics from Syracuse University. His graduate work was in the area of electrical engineering, and he received his Master's and Ph.D. degrees in electrical engineering from Ohio State University, Columbus. He is a technical manager in the Multimedia Networking Research Department of the Convergence, Software, and Computer Science Laboratory at Bell Labs, Holmdel, New Jersey. He has been active in the areas of SIP/IMS applications, broadband access (Ethernet passive optical networks), multipoint video compression/decompression, and recently in the area of SIP and multimedia. He was awarded five patents in multipoint video compression/decompression.

SOHEL Q. KHAN [M] (sohel_khan777@yahoo.com) obtained his Ph.D., M.S., and B.S. with distinction in electrical engineering from the University of Kansas. He is a principal product engineer II at Comcast Cable Communication, Philadelphia, Pennsylvania. He was a principal technology strategist at Sprint-Nextel until December 29h, 2006, where he had been since 1998. His research area includes strategic game theory applications in Internet QoS-based traffic engineering and providers' profit optimization, emerging Internet signaling and routing protocols, Internet security, Internet multimedia subsystem, and next-generation networks. He is a member of Tau Beta Pi and Eta Kappa Nu. He is an active contributor to Internet related standards development in IETF and ATIS. He is also involved in standards development with 3GPP/3GPP2, WiMAX Forum, and CableLab.

MICHAEL LUNA (michael.luna@openwave.com) is chief technology officer at Openwave Systems. He is a mobile Internet pioneer and has co-authored industry standards and held industry leadership positions in OMA, WAP Forum, CDG, TIA, and PCMIA. He also developed the first WAP 2 deployment in Japan. In 2002 he was awarded the title of Distinguished Engineer for his exceptional innovation in the telecommunications industry. He has been awarded nine patents in the mobile telecommunications field.