L-UNIVERSITA` TA` MALTA
Msida – Malta
DIPARTIMENT TA' L-INGINERIJA
TAL-KOMUNIKAZZJONI U KOMPJUTER

UNIVERSITY OF MALTA
Msida – Malta
DEPARTMENT OF COMMUNICATIONS
AND COMPUTER ENGINEERING

20th November 2008

CCE 4101 – Communications Systems

Assignment

(a) IP Multimedia Subsystems (IMS) Networks

IMS systems are a promising solution to harmonize various services which currently utilize different signaling networks to function.

Discuss the contents of the paper "Experiences with Blending HTTP, RTSP, and IMS" attached and suggest methods that can be used to achieve the goals in converging services on an IMS network.

OR

(b) Cooperative Networks

Convergence and cooperative networks can be a useful tool to make more efficient use of the available spectrum.

Discuss the contents of the paper "Cooperative Networks for the Future Wireless World" attached and suggest methods that can implemented to achieve seamless communications across different network technologies.

Use any published material to sustain your arguments.

The submitted report should follow A4 IEEE double column format with single-spaced, twelve-point font in the text. The maximum report length is four (4) pages. Reports in excess of four pages will not be read and a zero mark will be assigned. All figures, tables, references, etc. are included in the page limit.
A template in Word or Latex can be downloaded from the website: http://www.ieee.org/go/conferencepublishing/templates.

Hard deadline for the submission of the assignment: 15th January 2009 at 12:00.
No Assignment will be accepted after this date and time. Assignment can be submitted in groups of not more than two students.

Experiences with Blending HTTP, RTSP, and IMS

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

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.

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.

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.

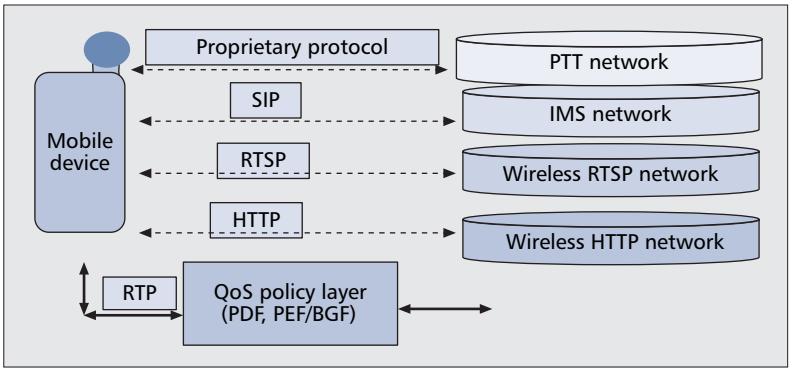


Figure 1. *Current parallel networks.*

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.

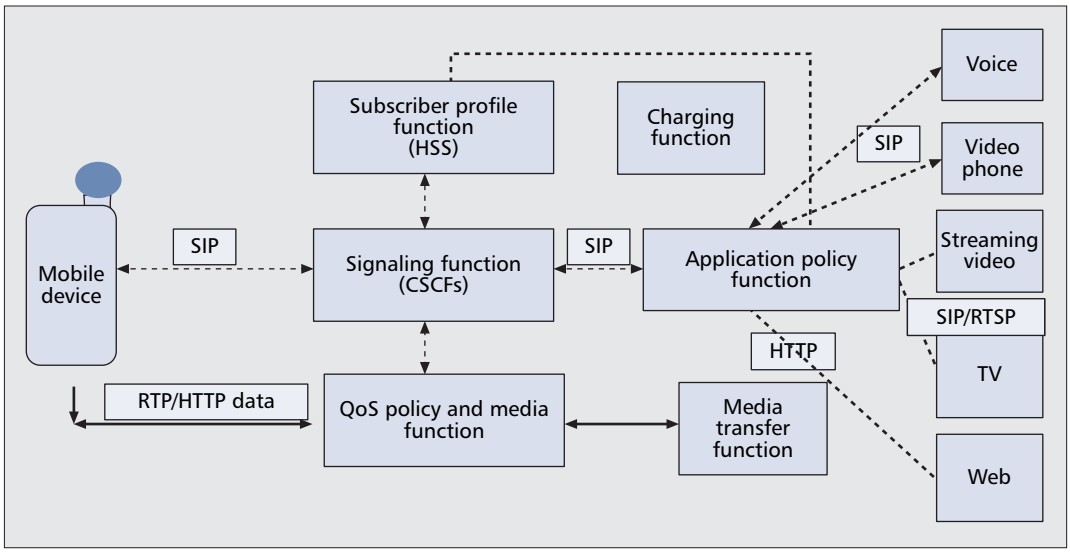


Figure 2. *Extended IMS architecture.*

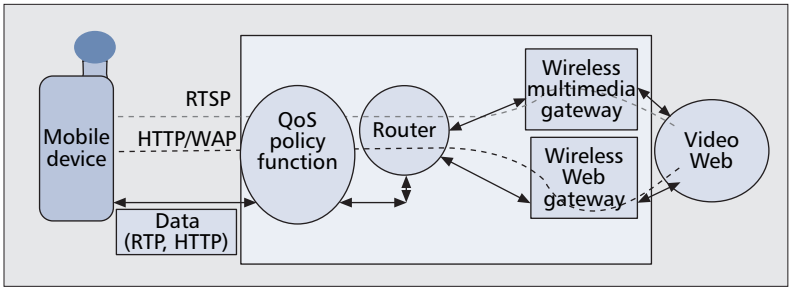


Figure 3. *Current wireless/wireline video streaming and wireless/wireline web architecture.*

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].

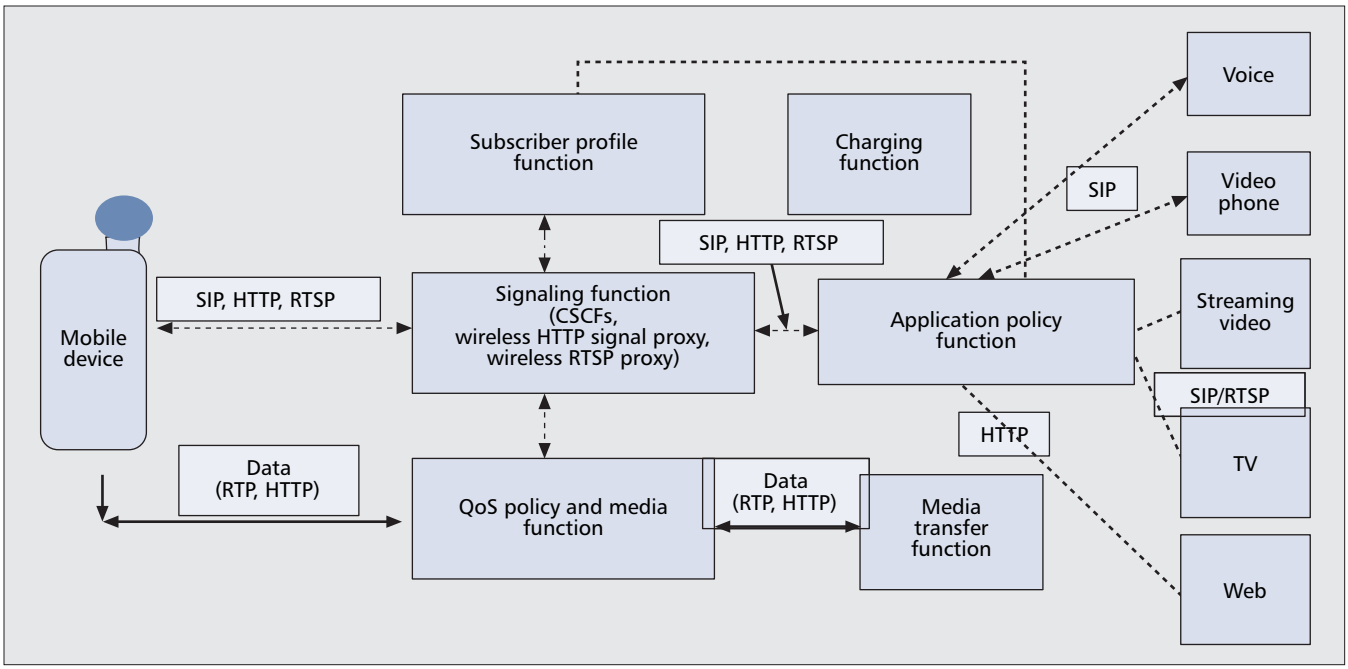
This synthesis enhances service creation, seamless service continuity, and concurrent service availability. It enables wireless and wireline providers to deliver coordinated "IMS ready content," Web, and video streaming content from a single platform; thus, eliminating duplicate components in the network. 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.*

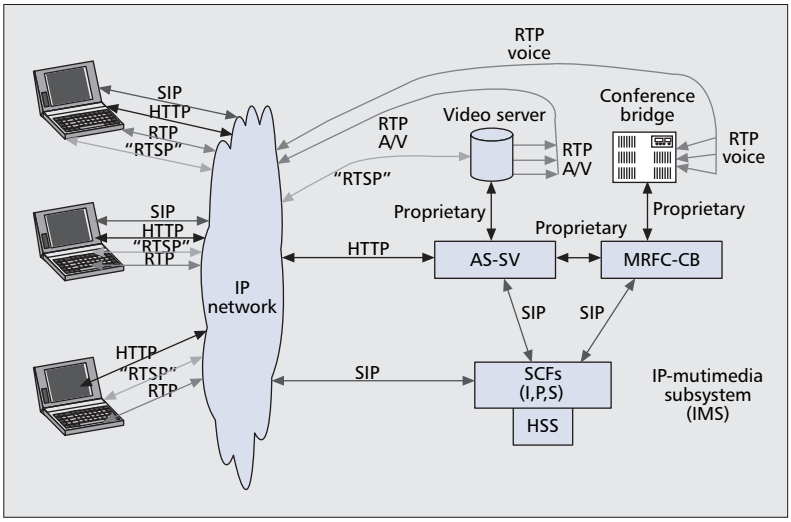
phone call by clicking a phone number located within a browsed page. Other extensions include creating a SIP based technique for general purpose sharing, enabling applications such as the sharing of Web pages, slide presentations, and documents.

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

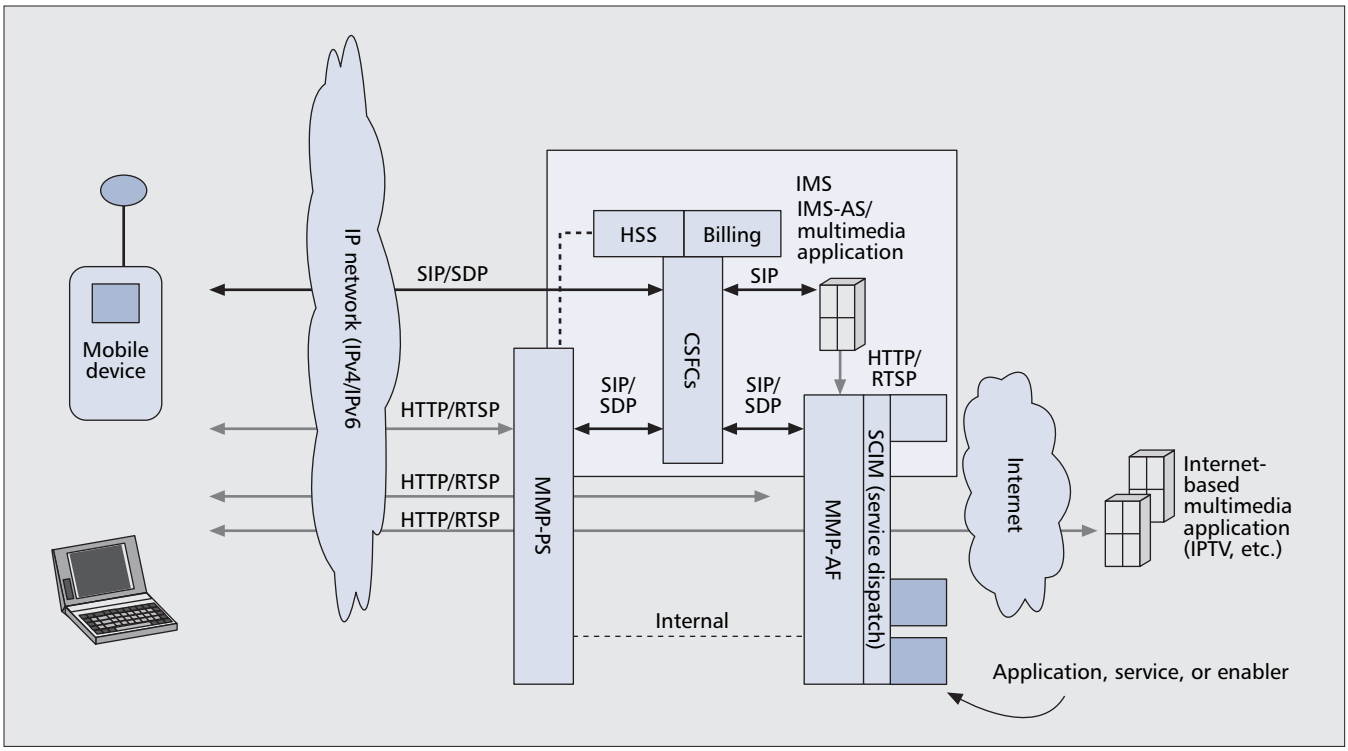


Figure 6. *MMP prototype architecture.*

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.

Cooperative Networks for the Future Wireless World

Christos Politis, University of Surrey, U.K.; Toshikane Oda, Nippon Ericsson, Japan; Sudhir Dixit, Nokia, USA; Andreas Schieder, Ericsson Eurolab GmbH, Germany; Hong-Yon Lach, Motorola Labs, France; Michael I. Smirnov, Fraunhofer FOKUS, Germany; Sami Uskela, Nokia, Finland; Rahim Tafazolli, University of Surrey, U.K.

ABSTRACT

Beyond-3G (B3G) systems have been envisaged as an evolution and convergence of mobile/wireless communication systems and IP technologies to offer a multitude of services over a variety of access technologies. To fulfill the vision, it is necessary to understand the requirements with respect to the support of heterogeneity in network accesses, communication services, mobility, user devices, and so on. Besides, it is equally important to promote the necessary research in networking technology by providing a guiding framework of research areas and technical issues with priority. The new architectures and technologies will have to address the fundamental assumptions and requirements that govern the design. All these issues are being tackled by the Cooperative Network working group (CoNet) of WWRF; the group is working on a series of white papers outlining B3G visions and roadmap, architectural principles, research challenges, and candidate approaches. This article outlines the CoNet concept, architectural principles, and guidelines for research into cooperative networks assuming that the B3G systems will be built over generic IP networking technologies. The article also presents the key research challenges, research framework, and major network components and technologies. The key points are that the system should be layered on demand, encourage reuse of independent modularized functional blocks, support multiple services and service creation, ensure consistent end-to-end connectivity across different access technologies, and cooperate in terms of network control, operations, and maintenance. The architecture shall include the endpoints of communications as part of the communications system, and should provide a secure and trusted environment in which network functions are performed; the network should self-organize dynamically. Additionally, the article presents the IST WSI architecture proposed to CoNet as a reference model along with some approaches to the outlined research challenges. Finally, this study selects three important network components and technologies (i.e., mobility management, multiple access, and moving networks) in order to provide with answers and possible solutions the research challenges presented in earlier sections.

INTRODUCTION

Beyond-third-generation (B3G) systems have been envisaged as an evolution and convergence of mobile communications systems and IP technologies to offer a multitude of services over a variety of access technologies. To fulfill the vision, it is necessary to understand the requirements with respect to the support of heterogeneity in network accesses, communication services, mobility, user devices, and so on. It is equally important to promote the necessary research in networking technology by providing a guiding framework of research areas and technical issues with priority. The new architectures and technologies will have to address the fundamental assumptions and requirements that govern the design. All these are being tackled by the Cooperative Network working group (CoNet) of the Wireless World Research Forum (WWRF), which is working on a series of white papers outlining B3G visions and roadmap, architectural principles, research challenges, and candidate approaches. In the remainder of this article CoNet's main objectives are outlined, while we examine the most important elements of CoNet's architectural principles. Additionally, we address the flamboyant research challenges in cooperative networks and try to give a glimpse of what we think would govern their design and encompass their futuristic form by introducing the main network components and technologies. This is mainly accomplished by focusing on three "hot concepts": mobility management, multiple access, and moving networks.

0163-6804/04/$20.00 © 2004 IEEE

CoNet Objectives

The CoNet concept will enable seamless communication activities on mobile devices operating in networks composed of heterogeneous technologies by enhancing the secure networking capability of B3G systems in terminal devices as well as in the infrastructure. It is envisaged that B3G networking by 2010 will support [1]:

- Coexistence and convergence of different legacy (wireless, wired, broadcast, etc.) and new networks, including moving and ad hoc networks
- Evolutionary deployment of secure, extensible, scalable, reconfigurable, and manageable B3G systems
- Independent evolution of network technologies and services
- Controlled service access by users and devices anywhere anytime by different means
- Pervasive and seamless mobile multiparty multimedia communications and access to all kinds of services to minimize adverse impact on user experience in the changing networking environment
- Value-added interface to upper-layer applications to enable and enhance their location, context, and quality of service (QoS) awareness
- Interactions with access systems for optimal mobility management and better use of scarce radio resources
- Unrestricted and innovative business models

The desire to seamlessly deliver application services across heterogeneous access networks in a B3G system leads to an obvious conclusion that B3G systems would best be built on IP networking technologies. However, it is also obvious that existing IP networking technologies do not provide all the necessary features to enable a B3G system and must be extended.

Therefore, it is inappropriate to set a single roadmap for technology development to achieve the vision, due to different concerns and priorities with respect to technology maturity, technical difficulties, market requirements, business models, and more. There will be different paths pursued, learning and leveraging from each another.

The following are some possible goals B3G research efforts could aim for to achieve the vision outlined above [1]:

- A system in which a user can move seamlessly between cellular, wireless, and wired access networks
- Interdomain administration and management support to allow seamless mobility between administrative domains
- Support of multiparty multimedia (including QoS-sensitive) sessions between networks and administrative domains

Architectural Principles

This section describes architectural principles and basic requirements commonly regarded as the cornerstone for studying related technologies and designing future IP-based cooperative systems. These features mainly include functional blocks, multiple services and their creation, connectivity issues, end-to-end concept, security, control, operation and maintenance, and self-organization [2].

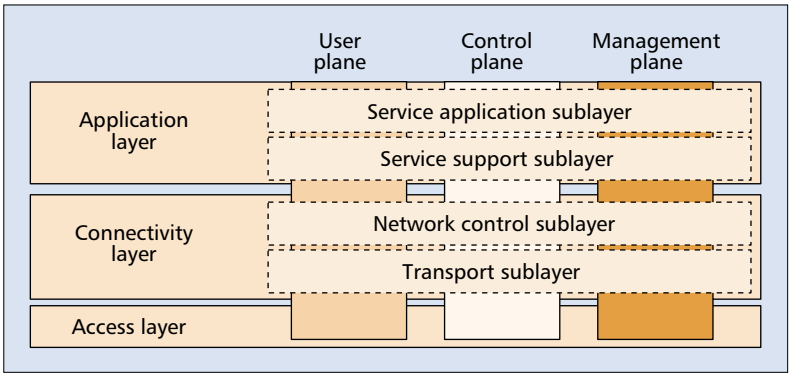


Figure 1. *Layers and planes.*

In CoNet, a *layered* approach is globally supported. This means that the functionality of the system is grouped into distinct layers that form logically separate subsystems. Any layered model has at least three layers: application, connectivity, and access. The application layer can be further divided into service application and service support sublayers. The connectivity layer can be further divided into network control and transport (IP) sublayers. And the access layer may contain several independent access networks that can also function simultaneously.

The layers should have well defined interfaces and be functionally independent of each other. Such an approach is required to ensure easy adaptation of heterogeneous access technologies, related technology changes, and flexible support for rapid service innovation.

Because of the need for cross-layer optimization (e.g., for power management, QoS support, etc), the layered approach in CoNet is considered as a design principle rather than as a universal design pattern, meaning that every B3G system may potentially have its own layered architecture.

Independent Functional Blocks, Modularization, and Reuse

The CoNet architecture should define each function as an independent functional block. This promotes component-based modular architecture, where different building blocks can be combined as needed to realize complex systems. This in turn requires standardized definition of syntax and semantics of these interfaces. For instance, functionalities required to process and route user data, handle control signaling, and deal with network management can be separated into different functional blocks in the user, control, and management planes, respectively. It should be noted that each layer of the architecture has its own separate set of functions for each plane; hence, nine different functional blocks can be identified, as shown in Fig. 1.

Realization of the functionalities as independent building blocks allows the introduction of new functions when needed without changing the whole architecture. The building blocks should be reused, when applicable, with different access technologies and realization of various

Besides protecting end users and networks from each other through authentication and authorization, the components of the CoNet concept should also be protected from any intrusion or malicious attack.

services. Therefore, implementing the same or very similar functionality multiple times should be avoided.

Cooperative Connectivity

The CoNet architecture should ensure connectivity between all the entities of a network in a consistent manner across all access technologies for any service. This requires consistent support for device mobility, QoS, authentication, authorization, and accounting (AAA),[1] and so on. The connectivity layer provides cooperation across various realizations of networks, called *cooperative connectivity*, and shall be independent of the various transport technologies used to link the nodes of the network together.

By separating access and transport the CoNet architecture makes it transparent to the common and standardized transport infrastructure and hides the technology from the end user, while facilitating the most efficient usage of spectrum resources. The CoNet architecture shall support both session continuation (i.e., handover when the access technology used is changed) and simultaneous usage of several access networks. Finally, the user should be able to seamlessly roam across different access technologies and administrative domains without any manual user intervention. This implies that the CoNet architecture should also support connection of subscribers to private IP networks through Network Address Translators (NATs) and Simple Traversal UDP through NATs (STUNs).

An End-to-End Approach

The CoNet architecture should include the endpoints (i.e., terminals) as part of the communication system, and support end-to-end (E2E) negotiation and fulfillment of QoS parameters, security settings, and so on. This E2E approach does not mandate that all functionality should be located in the endpoints. On the contrary, the functionality can also be provided in hop-by-hop, and/or edge-to-edge manner subject to proper and lawful termination of transport connections. The CoNet architecture should ensure that the interoperability of (and communication between) heterogeneous endpoints is maintained.

Security

The CoNet architecture should provide a securely protected environment in which network elements are deployed and interact. Besides protecting end users and networks from each other through authentication and authorization, the components of the CoNet concept should also be protected from any intrusion or malicious attack. The architecture should provide a security framework to enable protection of private information and data. Furthermore, the CoNet architecture should provide accounting capabilities and further enhance the AAA paradigm. The architecture should be distributed and consist of security components able to communicate through well defined interfaces. Underlying mechanisms such as DIAMETER [3] would provide the necessary secure communication framework where multiple networks cooperate.

To accomplish these, we would be well advised to rely on already existing security algorithms that are stable, resilient, and well established. The creation of new mechanisms and protocols is outside the scope of CoNet. Rather, its objective is to build the trust models and security associations between the various components of the distributed CoNet architecture with scalable access control to distributed components and resources.

Self-Organizing CoNets

The CoNet architecture is expected to comprise a number of networks exhibiting different capabilities in terms of coverage, capacity, transmission rates, transmission delay, and transmission cost. Networks might be in a cooperative or competitive relationship with each another.

Part of the access domain in the wireless world may not be centrally organized in the future and may even provide infrastructureless connectivity. Nodes may come and go and be loosely associated with each other, forming alliances whenever and wherever. Such nodes that could be part of the personal, the local, and even the global sphere would temporarily cooperate to provide connectivity in an ad hoc manner or share application resources. A network that is a member of the CoNet has to build relations with other networks to provide expected global connectivity and support the demanded access versatility. Relations between cooperative networks are supposed to be established dynamically. The creation of these dynamic relationships is a first aspect of self-organization.

The dynamic organization of relations between networks and individual network elements is supposed to result in a CoNet structure that adapts its topology to meet the demands of varying traffic patterns and transmission demands. To ease administration and operation, the network nodes should mostly be self-configuring, and resources should be distributed among them dynamically to cope with varying traffic volumes and traffic characteristics.

The Architectural Framework

An architectural framework supporting the research work on CoNet should aid the structuring of research tasks and identify relevant fields of research, including a reference model[2] to ensure that research toward 4G is carried out in a harmonized and open-ended manner. The Information Society Technology (IST) project Wireless Strategic Initiative (WSI) has proposed its reference model to WWRF.

The WSI Reference Model

The basic structure of the WSI reference model covers all aspects of the wireless world from business models and user issues down to radio interfaces. The reference model describes the grand building blocks of the wireless world and how they interact at reference points, and captures user scenarios and different views. The model also describes what kind of elements can be used to set up future wireless communication systems. It identifies how these elements can interact, the functionality they have to provide, and how a system can be composed out of them.

Following the sphere model, developed by

[1] Within the Internet Engineering Task Force (IETF) AAA combines authentication, authorization, and accounting; within CoNet this concept has to be enhanced with basic support for charging, roaming, auditing, and digital rights management.

[2] The development of the WWRF and its reference model was initiated by the European IST project Wireless Strategic Initiative [4].

the WSI think tank in 2000, the reference model identifies the building blocks of the wireless future. The main achievement is the definition of a *communication element* (a representation of a certain device or node) that acts as a communication entity in the different spheres; a large number of such elements may exist in the wireless world. Depending on their current communication context, they can be logically placed in different spheres (e.g., global, local, and personal) reflecting the proximity of building blocks with respect to a user. Therefore, the personal sphere can be seen as some kind of body area network, the local sphere serves as local networking infrastructure, and the global sphere is responsible for global connectivity.

The functionalities integrated in the communication elements are provided by different building blocks, assuming that the reference model should separate *content processing*, *control*, and *management* functions (see the subblocks inside a communication element) into their own E2E planes and subsystems. The building blocks that form communication elements are connected by vertical and horizontal reference points. The vertical reference points define interfaces between the building blocks of the communication elements, and horizontal reference points define interfaces between communication elements that reside in different spheres.

Research Challenges

In order to figure out features and characteristics the envisioned network shall have in different aspects, it is necessary to identify the significant requirements for the target network in terms of service provisioning and system capabilities. Those requirements will give an integral part of the guiding framework for investigation of design philosophies, principles and guidelines for reference models, network functional architectures, and solutions for individual functional components.

Furthermore, it is vital to identify and understand technical components that require further research, taking into account the requirements for the target network. We describe overall problem areas and generic technical issues to be addressed to provide guidance for defining the scope and specific targets of respective ongoing and future research activities/projects in the field of networking technologies to realize the CoNet vision [5].

Multi-Access Capability

In order to enable a system for multiple access network technologies, a number of different issues need to be investigated. Here we discuss some of these issues.

The device needs to be provided with access-technology-independent capabilities (application programming interfaces, APIs, and protocols) to detect which different access networks are currently available and learn their characteristics. Similarly, mobile terminals should be able to scan in a specific environment to discover candidate available access networks and register some policy issues.

Furthermore, there are different ways to enable access for a user and/or device over multiple access network technologies. Either the user could log on to each access network separately, or the authentication and authorization mechanisms for the different access networks could be connected to allow the user/device to move between different access networks without needing to log on multiple times. Preferably, the user should not have to close down and/or restart applications when moving between different access networks. To enable this, we might need a solution for mobility management that is tied to the logon procedures for the different access networks. Concurrently, APIs toward the applications should be designed so that an application can communicate requirements for access selection to the network or device, and also obtain information about the access network characteristics currently used. There is a need for a multi-access paging mechanism. It should be possible to page a terminal with multiple accesses even if some of the interfaces are switched off. Since B3G systems will potentially have a wireless link in the E2E connection that is either inherently unreliable and/or deal with handovers, solutions will be needed to support E2E session-level reliability, where a session goes into hibernation when disconnection happens and is resumed when a signal reappears after a period of disconnection.

The building blocks that form communication elements are connected by vertical and horizontal reference points. The vertical reference points define interfaces between the building blocks of the communication elements, and horizontal reference points define interfaces between communication elements that reside in different spheres.

Mobility Management

The architectural changes in the networking infrastructure are mainly dictated by the increasing levels of heterogeneity in modern networks, their interconnectivity, and their transport capabilities. Single access shifts toward multiple simultaneously used access networks, also known as multihoming. Single address space shifts to multiple coexistent address spaces — IPv4, IPv6. Single device usage shifts toward multiple device usage: PDA, smart phones, laptops, digital cameras, and more, used interchangeably or at the same time for different purposes. Additionally, all networking entities become mobile: users changing devices, devices changing access networks, access networks allocating network (IP) addresses dynamically and from different address spaces. Personal area networks, becoming reality in the immediate future, are inherently mobile. The use of vehicular and transport related networking will also increase with completely new applications, for example, in the telemetric area, which put new requirements on capabilities for mobility support. In essence, devices and gadgets without networking capabilities today are moving toward being networked as they serve as sinks and sources of digitally formatted information and content as well as running applications with networking needs.

Considering the above background, CoNet mobility management primarily consists of two major functional components: mobility routing management, including handover, and location management, providing functionality of location updating and paging.

A large number of technical issues need to be addressed in future research in the area of mobility management in order to achieve the required mobility routing capability in future networks. The following lists some of the issues.

- Mobility architecture (mobility support on a number of levels, layers, and protocols)

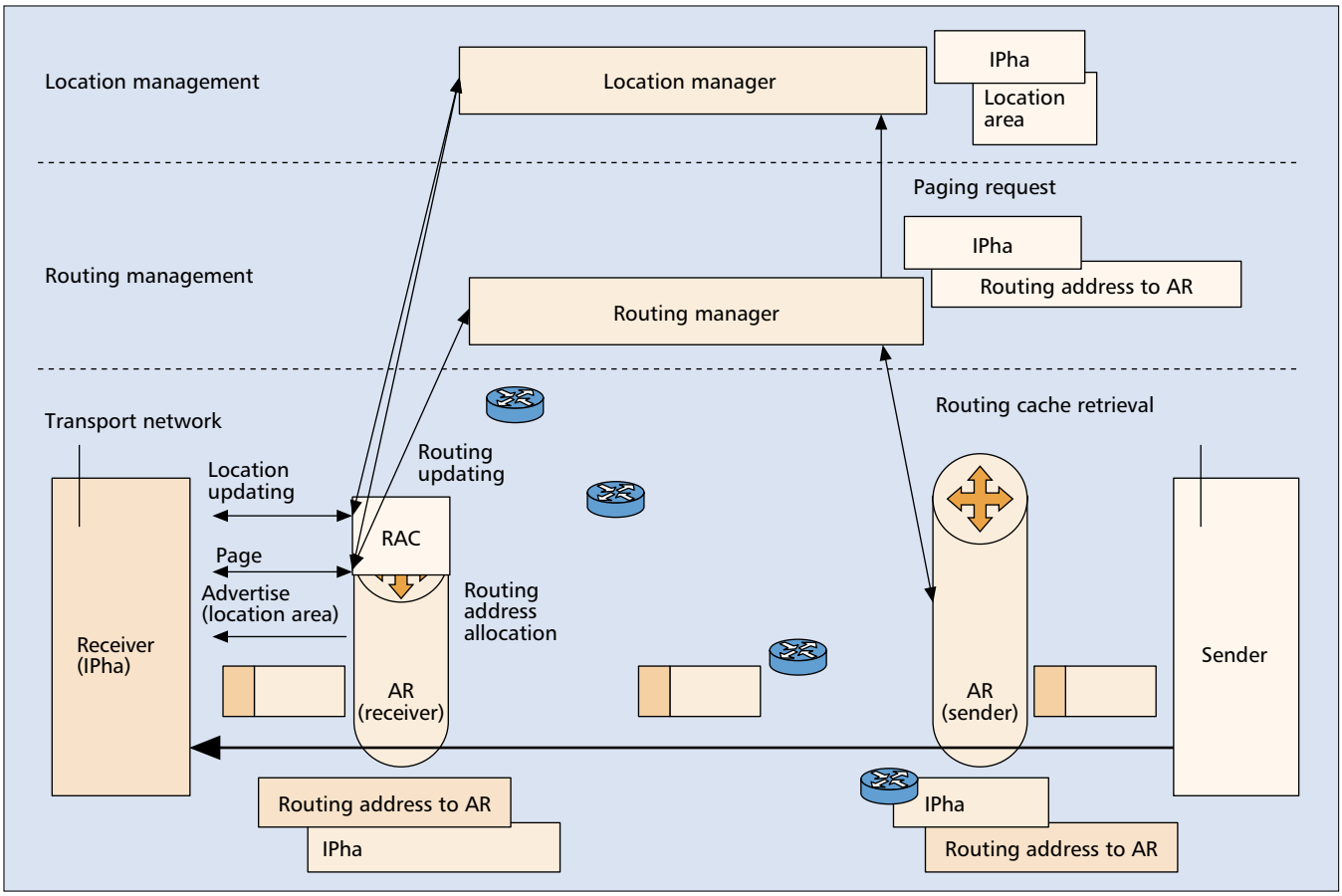


Figure 2. *The general architecture of IP² MM.*

- Group mobility (formation and mobility of networks, aggregation of data flows, group mobility support on several levels, support for multicast mobility)
- Session continuity (seamless support with efficient, timely, and predictable behavior, harmonized with transport protocols, supporting context and/or state transfer)
- Naming and addressing, mobility triggering, proactive mobility support

Moving Networks

A moving network (NEMO) consists of one or more mobile routers (MRs) with a number of devices connected to them. These networks can change their point of attachment to other networks because the NEMO physically moves or changes in topology.

NEMOs appear in different sizes, from small groups with a few devices attached to an MR up to NEMOs on ships and trains that might have several hundred or more mobile nodes (MNs) attached to them. The administrative characteristics will differ between NEMOs as well, whether the MRs and MNs are administered and owned by the same or different entities. In small NEMOs the MR and MNs are typically owned by the same entity, while in large NEMOs the MR(s) and MNs are typically owned by different entities. This affects, among other things, the level of trust between the different nodes. Different types of MNs have different requirements on location updates and connections to external networks (home operators) when the MNs are not actively participating in a session. A laptop might not require instantaneous location detection to a peer in another network, while a cell phone has to be reached immediately when someone is calling it. This puts different requirements on what and how much update traffic will be sent through the MRs.

Here are some research areas that could be addressed:

- Addressing and scalability.
- Speed: Some NEMOs may move at high speeds, which puts requirements on handovers and location updates, as well as frequency of these handovers and location updates.
- Mobility management, nested mobility, access control, and security.
- Multihoming: A mobile node can be attached to multiple ISPs/operators.
- Aggregation of signaling and user data traffic between MRs and external networks.

Network Components and Technologies

The objective of this section is to provide several possible solutions for research challenges, such as mobility management, multiple access IP QoS, and moving networks,[3] that are identified in [5] and discussed in previous sections. The possible solutions should take into consideration

[3] *These three research topics/challenges were picked due to their "great weight" in the global research community.*

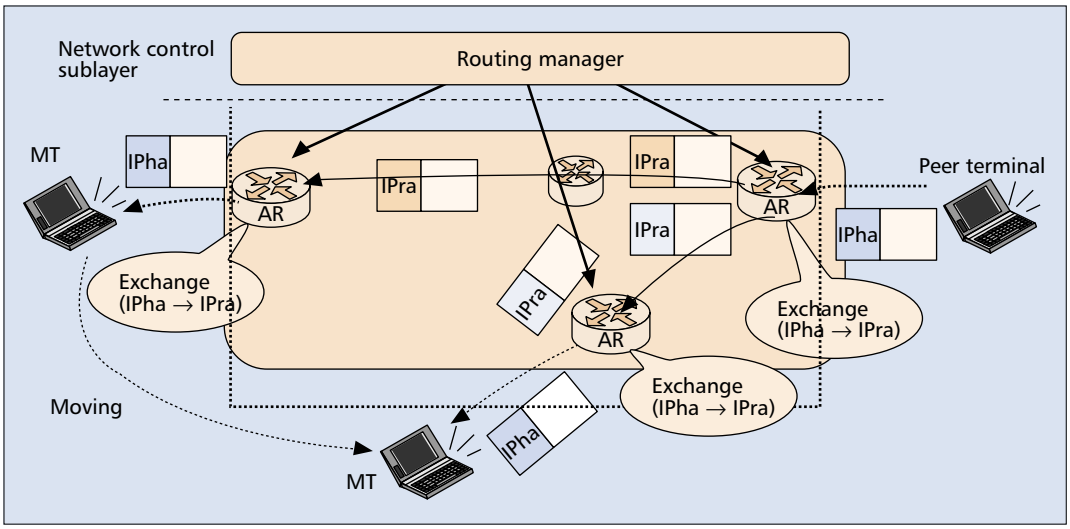


Figure 3. *The concept model of separation between IPha and IPra.*

The managers are triggered from an access layer to initiate access of mobile terminals and perform handover and location updates and other functions. The managers are not to be directly accessed by the mobile terminals.

the requirements identified in [5] and the architectural guidelines proposed in [2] and further described here.

This section also provides evaluations of the possible solutions and shows how these solutions can be quantitatively effective with respect to the requirements and architectural guidelines.

Mobility Management

One mobility management concept is based on IP^2, which is described in [6]. The solution assumes a routing manager and a location manager to reside in the network control sublayer controlling the access layer. Both management functions treat routing information and location information separately. The managers are triggered from an access layer to initiate access of mobile terminals (MTs), and perform handover and location updates and other functions. The managers are not to be directly accessed by the MTs. This enables the managers to be protected from malicious attempts and may evolve independent of the access layer.

The routing manager creates, updates, and deletes the routing cache in the transport sublayer (i.e., in the routers) for the MT it is currently managing. The location manager manages the MT location information and pages the MT when necessary. The routers forward packets and perform address exchange procedures for packet routing and rerouting. This enables MTs to communicate at any time without revealing their locations or movements to the peer terminal(s). In addition, this mechanism can also realize efficient utilization of radio frequencies (resources), MT power consumption, and use of network resources. Figure 2 shows a generic architecture of the proposed solution.

According to the vision of CoNet, the IP host address (IPha) of an MT should not have to be changed in order to maintain communication, even when the IP routing address (IPra) changes due to its movement.

The IPha is used only to identify an MT, while the IPra is used only to transport packets within networks and thus for location information. On the other hand, MTs are not aware of the IPra by exchanging the IPha and IPra at the edge router (router between MT and network), and packets sent to MTs are not encapsulated. Therefore, the proposed routing mechanism can prevent the location information of the MT being disclosed to the peer terminal. In addition, it can reduce the packet header overhead caused by encapsulation.

Figure 3 shows the separation between IPha and IPra. The access router (AR) of a peer terminal receives the packet of an MT (i.e., the destination address of the packet, which is the IPha). It then changes the destination addresses from IPha to IPra according to the routing cache, where the IPha and IPra are mapped; this whole process is managed by the routing manager. Then the AR forwards the packet to the MT according to its routing table. The packets can be routed to the AR to which the MT is attached, since the AR advertises the route in the network prefix of the IPra. The AR to which the MT belongs receives the packet and exchanges the destination addresses according to the routing cache from IPra to IPha. Then the AR forwards the packets to the MT with the IPha as the destination address.

If the AR of a peer terminal does not have the routing cache for the MT IPha, the AR requests the routing manager of the MT to set the routing cache for address resolution.

When the MT moves from an AR to another AR, the routing manager is triggered by the routing access controller (RAC) in the access layer, as previously mentioned (Fig. 2). Then it updates the routing cache in the AR of the peer terminal and creates a new one in the new AR to which the MT is attaching.

Before the update and creation steps, the new AR allocates the MT with a new IPra and notifies the routing manager of this. The access layer or routing manager triggers the new AR to allocate the IPra.

Multi-Access IP QoS

It is assumed that the E2E packet delivery service is provided at the IP layer. The quality of this delivery service determines the E2E QoS that is available for the applications. From the

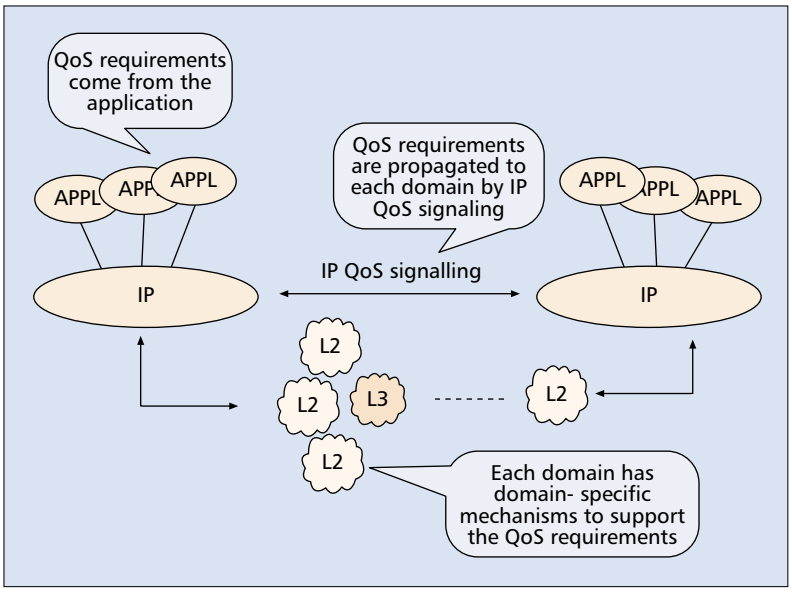


Figure 4. *A high-level view of a multi-access multidomain scenario.*

applications' perspective, the key elements of E2E QoS control are related to the following aspects (Fig. 4):

- The mechanisms and interfaces that allow the application (and application-level entities) to specify the QoS requirements. For instance, the application may be associated with a configuration or management tool that allows the specification of default (qualitative or quantitative) QoS parameters assigned to the application. Alternatively, a configuration tool may provide default QoS settings for a group of applications.
- The QoS requirements need to be transported to each domain along the E2E path. A domain may use this information to configure domain-specific functions and resources such that the required QoS can be provided within the domain. When a new domain becomes part of the E2E connection (due to mobility or, e.g., access selection) QoS related information may have to be transported to this new domain.
- Each domain may have specific mechanisms to support QoS requirements. Examples of such mechanisms include per-flow admission control and resource reservation (e.g., in a cellular wireless network), scheduling and prioritization mechanisms (e.g., in a differentiated services IP domain), or resource overdimensioning (e.g. in a best effort IP network).

The E2E QoS architecture of Fig. 5 builds on the high-level view of Fig. 4 and takes into account the above three aspects of E2E QoS provisioning. A communication session typically begins with exchanging application and session layer information. Such information exchange allows the involved parties to negotiate service capabilities and application-specific attributes such as encoding formats and version numbers. The communicating parties may proceed by requesting an appropriate bearer service from the network. This phase requires that QoS information be distributed to all domains of the E2E path. Specifically, the requirements on the E2E QoS information distribution mechanism should include:

- Applicability over various QoS technologies.
- Clear separation from the control information being transported.
- Independence from the application.
- Independence of the flow identification (state association with a user flow and flow identifier must be independent).
- Support for both uni- and bidirectional reservations
- Efficient service re-establishment after handover. In particular, the distribution mechanism must be able to seamlessly inter-work with other mobility and handover protocols.
- Admission control and resource reservation in domains along the E2E path must enable a domain to communicate the result of an admission decision to involved parties.

Once the QoS information is distributed along the E2E path, each domain may configure its own respective mechanisms to provide the necessary QoS for the user data flows.

Moving Network

To support moving network mobility, various techniques are proposed in the IETF NEMO WG [7]. However, all these proposals are based on Mobile IP; thus, there are several open issues, which are the same as those of Mobile IP from the mobile carriers' points of view regarding triangulation and route optimization techniques [8]. For instance, because of the packet overhead due to encapsulation, we cannot achieve route optimization and location confidentiality at the same time. In addition, these proposals assume that the prefix inside a NEMO is constant even if handoff occurs. Therefore, it makes it possible to hide the mobility of a NEMO from MTs; however, it causes more packet overhead due to several encapsulations, or using the routing header option to support route optimization, in comparison to host mobility.

As an alternative solution, four concepts are proposed: *care-of prefix (CoP)*, *aggregate router (AGR)*, *concatenated routing management*, and *hierarchical address management* to achieve the same quality as host mobility. However, location confidentiality and route optimization cannot be kept at the same time, which is an open issue in Mobile IP. In addition, the Mobility Management for IP-Based IMT Network Platform (IP^2) is proposed as one of the possible approaches for the next-generation mobile communications network beyond 3G systems [6, 8] to resolve this open issue.

According to the NEMO's assumption, the IP packet prefix in the NEMO is constant, and the packet header size gets bigger than the host. To resolve this problem, it is necessary to identify the care-of address (CoA) of an MT in the NEMO as well as the MT that can be anywhere in the core network. The CoP enables this. More concisely, the CoP changes the prefix inside the NEMO every time the NEMO moves to another access router, and gets the correct topological address (CoA) of the MT in the NEMO. When the NEMO connects to an AR, the AR assigns a new prefix, which shows that the NEMO has moved. Each MT in the NEMO uses this prefix to generate the CoA. In this way, an MT's CoA can be identified from anywhere in the core network. Furthermore, the allocated prefix is original for each NEMO in order to avoid generating duplicate CoAs even

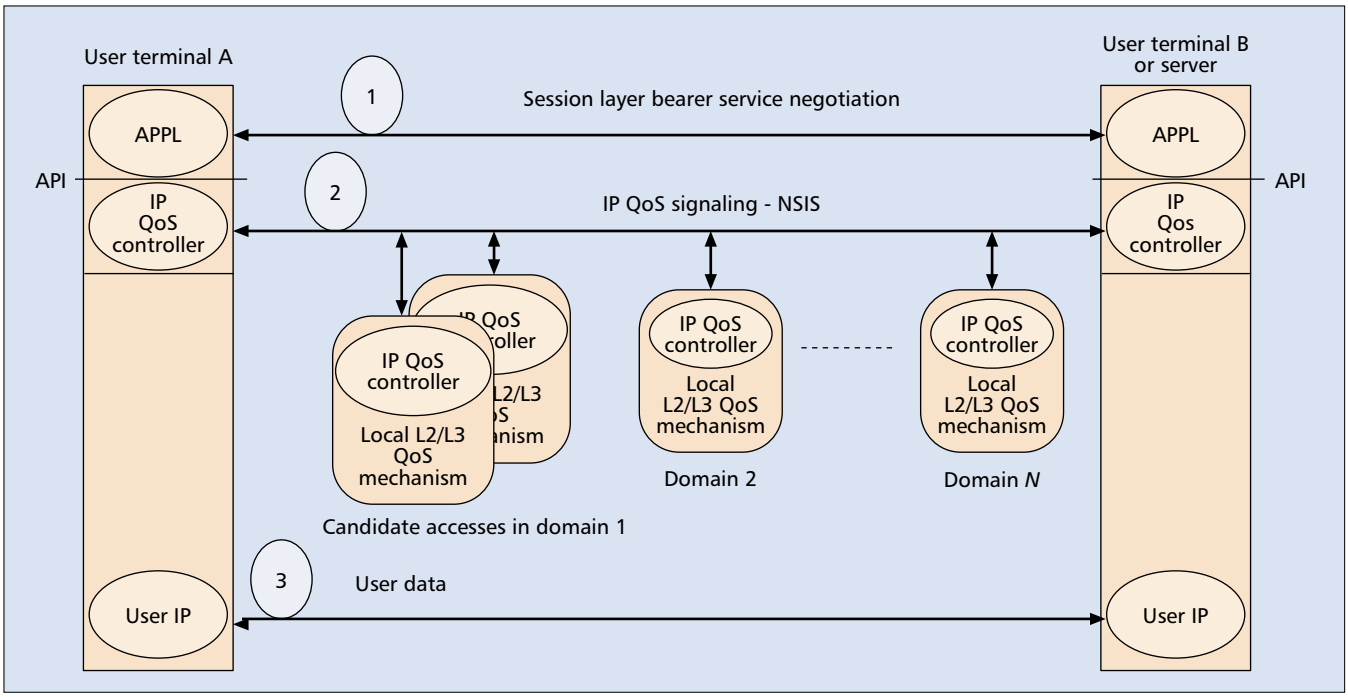


Figure 5. *Setting up an end-to-end IP QoS bearer by means of a QoS bearer specification mechanism.*

when handoff occurs. In the CoP, all CoAs of all MTs in the NEMO are changed when handoff occurs, so the number of binding update signals to the home agents (of all MTs and all core networks), in order to realize route optimization, are sent according to the characteristics of the NEMO, as shown in Fig. 6. The aggregate router and concatenated routing management can solve this problem and are described in the following sections.

Aggregate Router

The aggregate router (AGR) handles the mobility management of a NEMO. This is a router like the modified mobile anchor point (MAP) in Hierarchical Mobile IPv6 (HMIPv6) [9]. In addition, the AGR aggregates the update signals (U-Plane) to the ARs of all CNs. This reduces the number of binding update signals toward CNs with which the NEMO is communicating. Furthermore, the packets destined to the NEMO are always sent via the AGR. Therefore, AGR should be optimally located, considering route optimization for each communication path. Moreover, if the Moving Network moves too far, we should relocate to a different AGR. In addition, there are trade-off relationships between route optimization, handoff delay, and AGR relocation frequency. For instance, if the AGR is located near the NEMO, it may achieve route optimization and reduced handoff delay, but AGR relocation may occur frequently. On the other hand, if located far from the core network, an AGR cannot achieve route optimization and reduced handoff delay; but in this case, AGR relocation rarely occurs.

Concatenated Routing Management (Concatenated Home Agents)

In this concept, each home agent (HA) only holds information of an MT's whereabouts/attachments (e.g., with mobile router 1, MR_1), while the MR_1 HA maintains all the binding information for home address (HoA) and CoA. In this way, we can realize NEMO handoffs only by updating the MR_1 HA (Fig. 7).

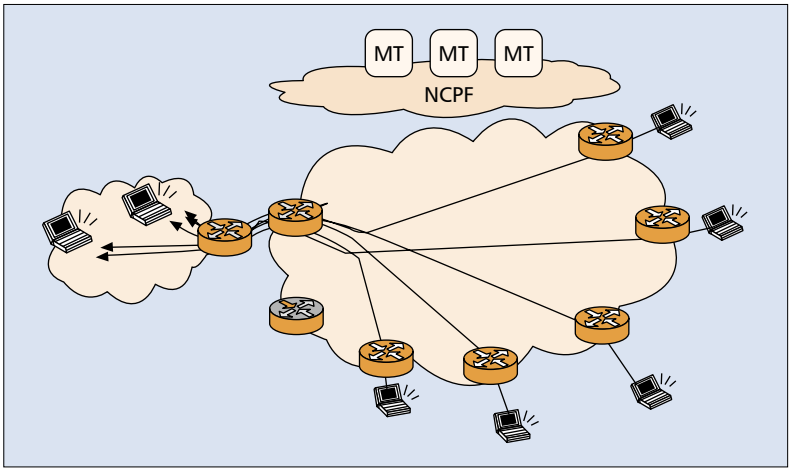


Figure 6. *Characteristics of moving networks.*

Hierarchical Address Management

When the NEMO moves, it is necessary to inform the AGR and MR_1 HA of the updated CoA, since these are changed. Therefore, the data volume of binding update signals is very large. The proposed address management concept resolves this issue. Concretely, in this concept the CoA of each MT in the NEMO consists of shared information and individual information. Shared information indicates the location of the NEMO, which changes whenever a handoff occurs. On the other hand, individual information provides the location of each MT in the NEMO and does not change even if handoff occurs. The CoP enables this address manage-

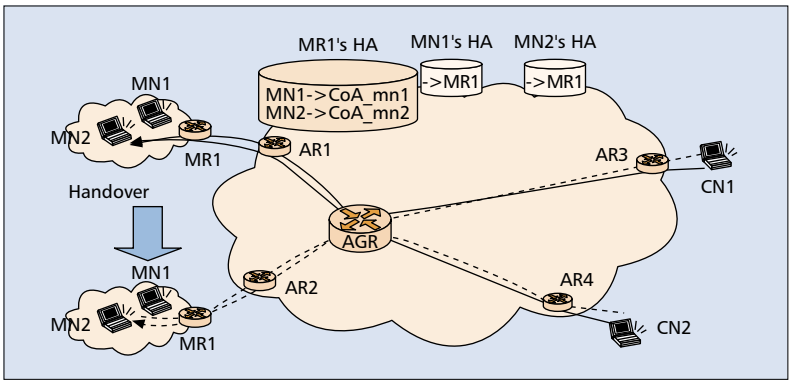


Figure 7. *AGR and concatenated HAs.*

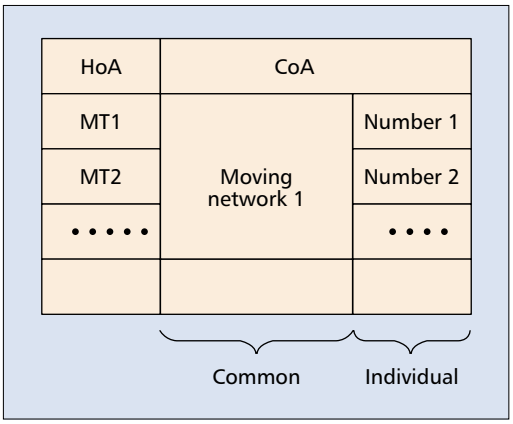


Figure 8. *Hierarchical address management.*

ment because the AR delegates the individual prefix to the NEMO, in order to avoid generating duplicate CoAs. In more detail, MR_1, AGR, and MR_1 HA manage the binding information as mentioned above, so we can realize a handoff of the NEMO by updating only the shared (common) information (Fig. 8).

Summing up, hierarchical address management including an aggregate router and concatenated routing management makes it possible to support frequent handoffs.

IP² Routing for the Moving Network

IP² resolves open issues in Mobile IP, as mentioned earlier [8]. Therefore, by applying the four key technologies for IP² routing management as discussed in the previous subsections, we can get much better results.

The IP² architecture is as shown in Fig. 9, where the aggregate router and concatenated routing management are introduced in the network control sublayer. In this proposed architecture, we can categorize the two routing mechanisms by their location confidentiality (i.e., the location privacy information). Here, one case is the gateway, with IPha and IPra, being the MR. In this case, the MR deals with IPra, which indicates the location of the MTs; thus, we cannot keep the location information confidential from the MR. Therefore, it may be easier if the MR is owned by carriers/providers. Another case is that the gateway with IPha and IPra is an AR. In this case, the MR does not deal with IPra, so we can keep the location information confidential from the MR and MTs. Therefore, it is possible that users can own the MR (Fig. 10).

Conclusions

This article outlines the CoNet concept, architectural principles, and guidelines for research into cooperative networks assuming that B3G systems will be built over generic IP networking technologies. The article has also presented the key research challenges, research framework, and major network components and technologies. The key points are that the system should be layered on demand, encourage reuse of independent modularized functional blocks, support multiple services and service creation, should have consistent end-to-end connectivity across different access technologies, and cooperate in terms of network control, operations, and maintenance. The architecture shall include the endpoints of the communications as part of the communication system, should provide a secure and trusted environment in which network functions are performed, and should self-organize dynamically. Additionally, the article has presented the IST WSI architecture proposed to CoNet as a reference model along with some approaches to the outlined research challenges.

Finally, this study has selected three important network components and technologies (i.e., mobility management, multiple access IP QoS, and moving networks) in order to provide answers and possible solutions to the research challenges presented in earlier sections.

Acknowledgments

The authors would like to acknowledge the contributions of their colleagues from the CoNet SIG in WWRF Working Group 3.

References

[1] H. Abramowicz *et al.*, "The Wireless World Initiative: A Framework for Research on Systems Beyond 3G," IST Mobile and Wireless Commun. Summit 2004, Lyon, France, June 27–30, 2004.
[2] S. Dixit *et al.*, "Cooperative Network Architectural Principles, Version 1.07," WWRF, 10th mtg., Oslo, Norway, June 10–11, 2004.
[3] DIAMETER Base Protocol, IETF draft, http://search.ietf.org/internet-drafts/draft-ietf-aaa-diameter-11.txt
[4] IST-1999-12300, "WSI: Wireless Strategic Initiative," http://www.ist-wsi.org
[5] T. Oda *et al.*, "Co-operative Network Research Challenges, Version 1.0," WWRF, 10th mtg., Oslo, Norway, June 10–11, 2004.
[6] H. Yumiba, K. Imai, and M. Yabusaki, "IP-based IMT Network Platform," *IEEE Pers. Commun.*, vol.8, no.5, Oct. 2001, pp. 18–23.
[7] NEMO WG, http://www.nal.motlabs.com/nemo/
[8] C. Perkins and D. Johnson, "Route Optimization in Mobile IP," draft-ietf-mobileip-optim-11.txt, Sept. 2001, work in progress.
[9] H. Soliman *et al.*, "Hierarchical Mobile IPv6 Mobility Management (HMIPv6)," draft-ietf-mobileip-hmipv6-07.txt, Oct. 2002.

Biographies

Christos Politis [M] (c.politis@surrey.ac.uk) received an engineering degree from the Technological University of Athens, Greece, in 1996, an M.Sc. in mobile and satellite communications from the University of Surrey (UniS), United Kingdom, in 1999, and his Ph.D. in mobile networking from the Centre for Communication Systems Research (CCSR) at UniS in 2004. Past positions include telecommunications engineer with INTRACOM SA, Athens, IT engineer at AMSAT, and

wireless communicationss engineer at Hellenic Air Force General Staff. He is currently a senior researcher at CCSR, UniS. He has been involved with several IST (FP5) projects and is currently working in the IST MAGNET Integrated Project. He is a patent holder, and has published more than 30 papers at international journals and conferences. His research interests include wireless LANs, mobility management, security and QoS frameworks for IP-based infrastructures, and personal area networks.

TOSHIKANE ODA (toshikane.oda@ericsson.com) has been with Nippon Ericsson K.K., Tokyo, Japan, since 2000, where he is currently manager of the Market Support and New Technology Department. He is vice-chairman of the Working Group on IP-Based IMT Platform in TTC. He worked in R&D on network design and services in KDD, and was also involved in standardization, especially in International Telecommunication Union — Telecommunication Standardization Sector Study Group 2 (ITU-T SG2). He received his M.E. and Ph.D in communication engineering from Osaka University, and an M.S. in computer science from Columbia University.

SUDHIR DIXIT (sudhir.dixit@nokia.com) has been a research fellow at Nokia Research Center, Burlington, Massachusetts, since November 2003, and concentrates on research on next-generation networks. Prior to this, he was a senior research manager from 1996 to October 2003, and managed the Mobile Internet Performance group and its earlier incarnations, specializing in pervasive communications, content networking, and optical networking. Before joining Nokia, he held various engineering and management positions at other major companies, including NYNEX (now Verizon), GTE, Motorola, Wang, Harris, and STL (now Nortel Europe Labs). These are on wireless IP, IP over WDM, and content networking in the mobile Internet. He received a B.E. degree from MANIT, Bhopal, India, an M.E. degree from BITS, Pilani, India, and a Ph.D. degree from the University of Strathclyde, Glasgow, Scotland, all in electrical engineering. He also received an M.B.A. degree from Florida Institute of Technology, Melbourne.

ANDREAS SCHIEDER (andreas.schieder@ericsson.com) received his M.Sc. degree and Ph.D. in electrical engineering from the Technical University of Aachen in 1995 and 2003, respectively. During his occupational career within Ericsson Research he was involved in research on UMTS and GPRS performance evaluation. He concentrated on the support of real-time applications and actively contributed to the standardization of the 3G radio access network GERAN. He is a WG vice chair of the WWRF.

HONG-YON LACH (hong-yon.lach@motorola.com) received a B.S. degree in computer science, with a minor in economics, from the American University of Paris, France, in 1987. He has extensive experience in R&D of OSI and IP communication protocols and applications. In particular, he was a major architect of and contributor to ETSI Project HIPERLAN (Type 1) to design and specify the MAC and CAC services and protocols, and has been contributing to IETF. He has also actively participated in various EU projects in the ESPRIT, ACTS, and FP5 programs. In May 1997 he joined Motorola Labs — Paris, France. Currently, he is manager of the Edge Networking Research Laboratory (ENRL), which researches, develops, and demonstrates innovative technologies and software to enable IP-based mobile multiparty multimedia networking, especially in the context of B3G systems. He also helps lead Motorola Labs' overall Internet research activities, both technically and strategically, in Motorola Labs Global Centre for Networks and Systems Research.

MICHAEL I. SMIRNOV (smirnow@fokus.fraunhofer.de) graduated with honors from St. Petersburg Electrotechnical University in 1977, got his Ph.D. (computer science) in 1984, and a degree of associate professor in 1989; in 1992 he was appointed deputy director (research) of IMICS Research Institute in St. Petersburg, Russia. Since 1994 he is with GMD (now Fraunhofer) FOKUS (Research Institute for Open Communication Systems) in Berlin as a researcher and project manager (1994-96), head of competence centre (1999–2001), Direktorium member (since 2001) responsible for Internet research, standardization activities, and links to academia. He has published over 50 technical papers in Russian, three in German, 38 in English; and has co-edited six books, two journals (special issues), and five IETF drafts. His contributions include routing research, performance analysis, implementation of IP QoS multicast over ATM; distributed computing algorithms and protocols; systems research in networking; QoS; scalability analysis, advanced

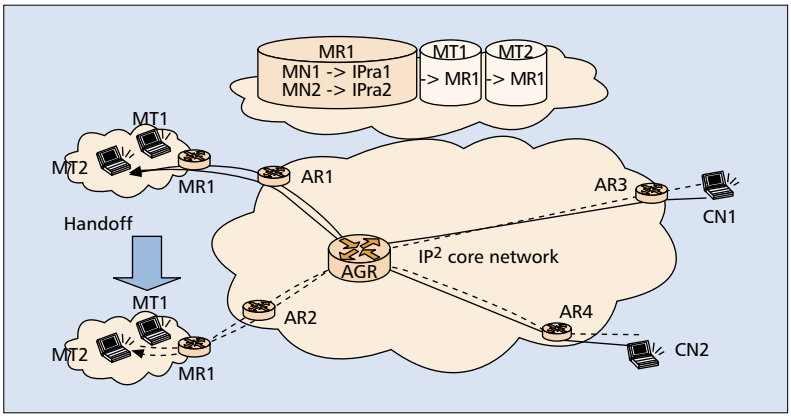


Figure 9. *IP² architecture for a moving network.*

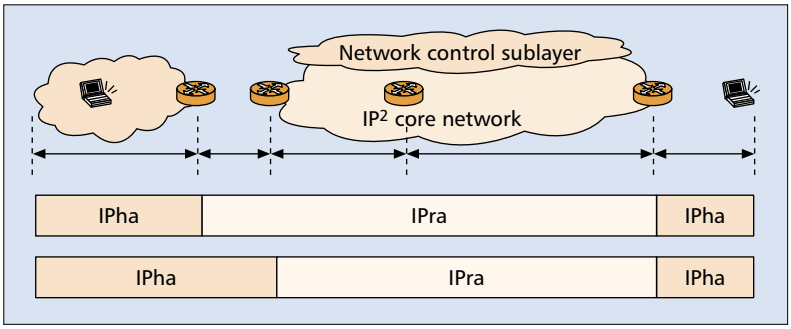


Figure 10. *Categories of routing mechanisms.*

Internet services, policy-based control and programmability, and autonomic communication. He is serving the networking community as COST263 chair (1999–2003); E-NEXT PMC member, as chair of international conferences (1986, 1992, 1995, 2000, 2001-2004); as EU Commission expert (high-performance networking, future and emerging technologies). Since 1997 he is an adjunct professor at TU Berlin in advanced Internet services.

SAMI USKELA (sami.uskela@nokia.com) received his M.Sc. degree in electrical engineering from Helsinki University of Technology in 1999. He has been with Nokia since 1997. He has been involved in research on network architectures and service platforms ever since. He has been actively involved in standardization of the 3G systems: during 1998 he contributed to network architecture and service platform standardization in TTC (Japan) and during 1999–2000 to all-IP standardization in 3GPP. He has also been active in WWRF; he served as editor of WWRF white paper on cooperative networks during 2002–2003 and contributed actively to the discussions in the forum. Since May 2001 he has been responsible for B3G network research within Nokia Networks, currently working as senior research manager. He has published several articles in international conferences and journals, and holds over 20 patents and patent applications covering different areas of network architectures and service platforms.

RAHIM TAFAZOLLI (r.tafazolli@surrey.ac.uk) is the head of the Mobile Communications Research Group in CCSR, UniS. His research activities are on optimization techniques for mobile multimedia networks, mainly on advanced resource management, mobility management, and media access control. He has published more than 300 research papers in refereed journals and international conferences, and as an invited speaker. He currently has more than 15 patents in the field of mobile communications. He is advisor and consultant to a number of mobile companies. He is the founder and past chairman of the International Conference on 3G Mobile Technologies. Regulations on Information Technology and Telecommunications, a member of the WWRF Vision Committee, and past Chairman of New Technologies group of WWRF, and academic coordinator of the U.K. Mobile Virtual Centre of Excellence.