◆ Service Control for Next-Generation Applications in Wireless IP Multimedia Networks

Lynell E. Cannell, Michel L. F. Grech, Anne Y. Lee, and Musa R. Unmehopa

Within the 3rd Generation Partnership Project (3GPP), the thrust of the session initiation protocol (SIP)-based Internet protocol (IP) multimedia subsystem (IMS) is envisaged to allow a swift progression towards the provision of multimedia applications for increasingly demanding end users. The paradigm of service programmability using open network application programming interfaces (APIs), with open service access (OSA) as its main exponent, is helping to drive this development together with the use of SIP. The focal point of this paper will be the multimedia services architecture in the IMS by providing details of the interaction of the IMS and the application servers in the form of the OSA gateway and the SIP application server. The paper aims to assess the value of the IMS service control (ISC) interface on application server interaction in the IMS. The paper will provide an OSA application use case, and will also present the presence server as an example of a SIP application server that fits in with the IMS.

Introduction

For the last few years, the 3rd Generation Partnership Project (3GPP) [1] has been responsible for producing technical specifications for a third-generation mobile system. Initially focused on evolving the Global System for Mobile Communications (GSM) core network, it has now turned its attention to evolving the core network based on an Internet protocol (IP) backbone. In particular, as part of one of the 3GPP releases, a new subsystem is being introduced, referred to as the IP multimedia subsystem (IMS) [5]. The primary aim of the IMS is the delivery of an integrated voice and data infrastructure to support real-time as well as non-real-time multimedia capabilities. This IMS was initially introduced as an overlay to the existing packet domain, whereas in the 3GPP this is provided by the General Packet Radio Service (GPRS). The overlay concept is shown in **Figure 1**, where the underlying packet network could either be GPRS or packet data network (PDN). As part of the harmonization process between the Universal Mobile Telecommunications Network (UMTS) and CDMA2000* networks, the 3rd Generation Partnership Project 2 (3GPP2), the equivalent standards partnership project with the task of evolving code division multiple access (CDMA) networks, aims to introduce an equivalent IMS as part of its own releases.

Bell Labs Technical Journal 8(1), 27–42 (2003) © 2003 Lucent Technologies Inc. Published by Wiley Periodicals, Inc.
Published online in Wiley InterScience (www.interscience.wiley.com). • DOI: 10.1002/bltj.10045

The overlay architecture will evolve in future releases to support access independence, such that subscribers may access the subsystem through other means, such as a wireless local area network (WLAN) connection or a fixed wireline connection. This is subject to additional work chiefly in the area of authentication of the end user without a UMTS subscriber identification module (SIM).

The IMS aims to bring about the union of the Internet and the wireless environment. This synergy would allow for communications with multiple sessions, where mixed media can be used person-to-person or person-to-machine, and where media can be changed during communications. The expectation is that this powerful mix of capabilities will allow for new and richer services than currently available in the wireless environment. From an end-user perspective, service integration can be of significant benefit allowing for adoption of a personalized lifestyle. For example, by introducing a single, global, easy to memorize, and intuitive identifier (not a telephone number, but an operator-administered uniform resource identifier (URI), such as myAddress@MobileProvider.com), a user may be reachable through multiple communications forms such as voice, e-mail, video, or instant messaging using a single communications device.

The IMS introduces a multimedia call control model [4] based on the session initiation protocol (SIP) defined by the Internet Engineering Task Force (IETF) [8]. New network elements are introduced as part of this subsystem that are associated with the session establishment, control, and tear down. These nodes are generally referred to as call session control functions (CSCF) and several variants exist depending on the roles they provide. They maintain state information for the duration of the session. In order to allow interworking with other networks, such as the existing public-switched telephone network (PSTN), signaling and media gateways are introduced as well. Additionally, media resource functions are introduced to support conferencing and announcement capabilities.

Although these nodes provide the basic capability for multimedia connectivity, the feature-rich services anticipated by the end user will be provided by a separate services layer. This services layer can be described briefly as a collection of application servers that can exist in the mobile network operator's domain or can exist outside the operator's traditional domain. In

Panel 1. Abbreviations, Acronyms, and Terms

3G—third generation
3GPP—3rd Generation Partnership Project
3GPP2—3rd Generation Partnership Project 2
API—application programming interface
B2BUA—back-to-back user agent
CDMA2000*—3G evolution of IS-95 standard
CDMA—code division multiple access
CF—call forwarding
CPL—Call Processing Language
CSCF—call session control function
GGSN—gateway GPRS support node
GMLC—gateway mobile location center
GPRS—General Packet Radio Service
GSM—Global Systems for Mobile Communications
HSS—home subscriber server
I-CSCF—interrogating CSCF
ID—identification
IETF—Internet Engineering Task Force
IMSI—International Mobile Subscriber Identity
IMS—IP multimedia subsystem
IP—Internet protocol
ISC—IMS service control
ISDN—integrated services digital network
ISUP—ISDN signaling user part
MRFC—media resource function controller
MRFP—media resource function processor
OSA—open service access
P-CSCF—proxy CSCF
PDN—packet data network
PSTN—public-switched telephone network
S-CSCF—serving CSCF
SDP—session description protocol
SGSN—serving GPRS support node
SIMPLE—SIP for instant messaging and presence leveraging
SIM—subscriber identity module
SIP—session initiation protocol
UE—user equipment
UMTS—Universal Mobile Telecommunications System
URI—uniform resource identifier
URL—uniform resource locator
WLAN—wireless local area network

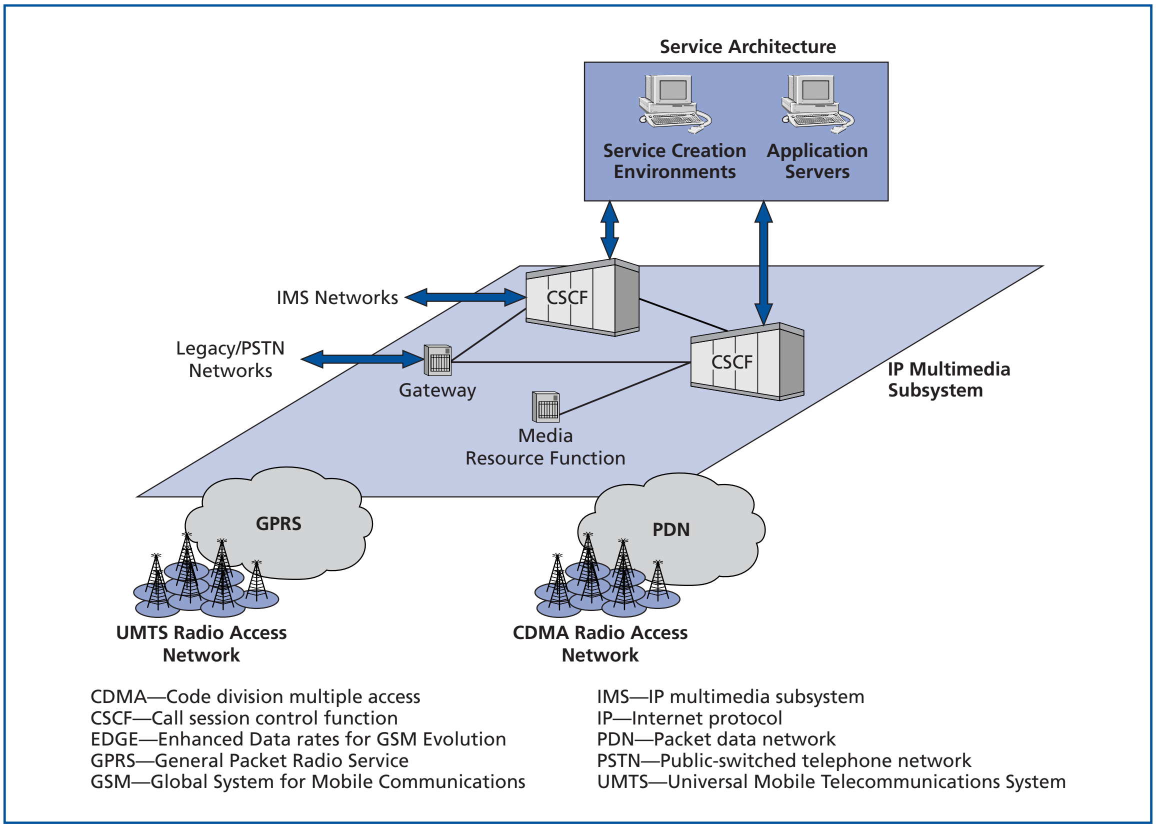


Figure 1.
IMS overlay architecture concept.

the latter case, the service provider and the operator enter a commercial agreement to provide services. A recent report commissioned by the UMTS forum [11] estimates that the revenue from third-generation services is expected to be in the region of $322 billion in the next eight or so years. Any service that enhances the user's experience is likely to generate some revenue. There is no space for complacency. Services will have to be dynamic to meet customer's very high expectations, and a higher service churn is expected.

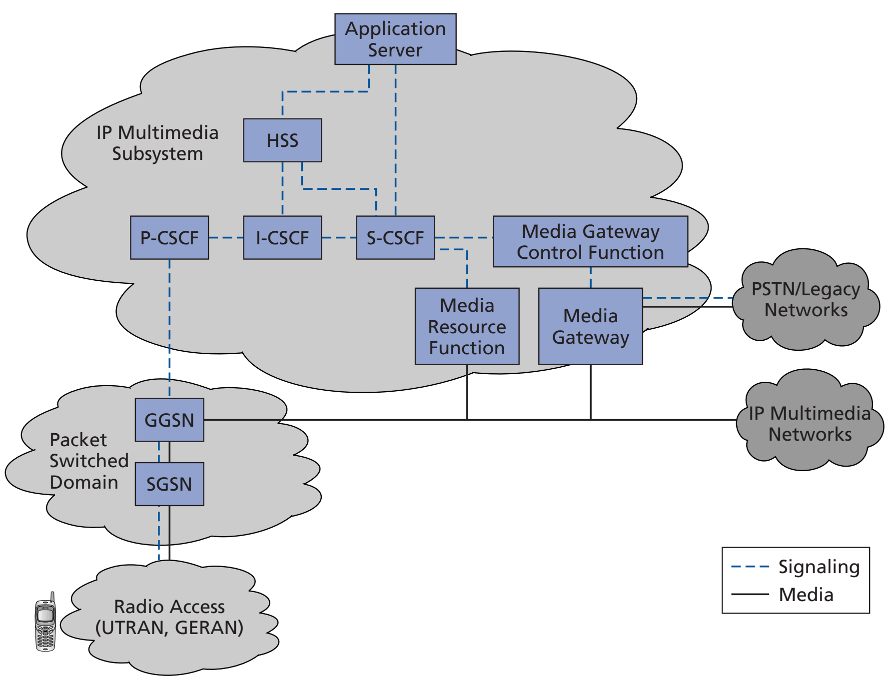
This paper is primarily focused on the service architecture that accompanies the IMS. Within 3GPP, three sets of standardized platforms for service control are being introduced, each one addressing particular needs, extending from legacy service support to lightweight SIP-based support. This paper will focus on what are generally considered the two primary mechanisms for interaction between the IMS and the application service layer, in the form of the open service access (OSA) gateway and the SIP application server. First, the paper will provide an overview of the IMS and the new network elements that are introduced. Next, some of the new procedures and processes required to support services in the IMS are presented. When the stage is set, the IMS service architecture will be introduced, along with an overview of the various service control scenarios. This discussion will be augmented with examples of the two application servers of interest, the OSA gateway and the SIP application server. To illustrate the interactions for service control, the presence service is used as an example in this paper, where we outline how IMS supports presence and how the architecture supports a presence-enabled service.

Overview of the IMS

This section will provide a brief overview of the IMS. A detailed overview of the IMS architecture is provided in [7]. **Figure 2** shows the high-level IMS architecture and the relationship to the underlying packet-switched domain and radio access. Although the figure shows underlying 3GPP GPRS and UMTS radio access architecture, the same IMS high-level architecture also applies when there is 3GPP2 underlying architecture.

Figure 2 and the following description of the IMS are simplified to highlight aspects that are important for understanding how services are provided.

The IMS employs the concept of home control architecture, so that no matter whether a subscriber is in the home network or in a visited network, the subscriber's services are orchestrated by the home network. In this way, the subscriber sees the same subscribed services and subscribed service behavior at all times, regardless of where the subscriber is roaming. The three call session control function types shown in Figure 2 enable home control. The proxy CSCF (P-CSCF) is the contact point into the IMS for the user equipment (UE) and may be in a visited network. The interrogating CSCF (I-CSCF) is the contact point into the UE's home network from other networks. The serving CSCF (S-CSCF) is the SIP session control point for the UE and is always located in the UE's home network. In the IMS model, the S-CSCF does not contain service



CSCF—Call session control function
EDGE—Enhanced Data rates GSM Evolution
GERAN—GSM EDGE radio access network
GGSN—Gateway GPRS support node
GPRS—General Packet Radio Service
GSM—Global System for Mobile Communications
HSS—Home subscriber server
I-CSCF—Interrogating CSCF
IMS—IP multimedia subsystem
IP—Internet protocol
P-CSCF—Proxy CSCF
SGSN—Serving GPRS support node
UMTS—Universal Mobile Telecommunications System
UTRAN—UMTS terrestrial radio access network

Figure 2.
IMS high-level architecture over a GPRS network.

logic—service logic is provided by the application servers.

The home subscriber server (HSS) holds subscriber and service-related data. The I-CSCF accesses the HSS to determine an appropriate S-CSCF for the subscriber, the S-CSCF accesses it in order to be able to enlist appropriate application servers to provide services to the subscriber, and the application server may access it to get user profile and subscription information.

An application server may need to access IMS network media resources, such as an announcement facility. The IMS architecture isolates these resources from the outside by having the application servers access IMS network media resources by going through the S-CSCF, which acts as a proxy, to reach the media resource function controller (MRFC). The MRFC controls the network media resources located in the media resource function processor (MRFP). For simplicity, in Figure 2 the MRFC and MRFP are grouped together and labeled as the media resource function. The MRFC and MRFP may be used both for services involving IMS sessions, as well as for sessions involving calls to or from the PSTN or a legacy mobile network.

The media gateway control function and media gateway are employed when the session involves a call to or from the PSTN or a legacy mobile network. The media gateway provides the bearer path between the packet-switched and the circuit-switched environments and may also provide media resources. The media gateway control function controls the media gateway and also does protocol conversion between SIP and ISDN signaling user part (ISUP).

The CSCFs, media gateway and media resource control functions, and application servers all use SIP-based signaling among themselves. Of particular interest is the SIP-based service control interface referred to as the IMS service control interface between the S-CSCF and the application server.

Procedures and Processing in the IMS

The following subsections go into more detail about service-related aspects of IMS and about how the different IMS entities interact to provide services. Several new procedures and new processing in the IMS are discussed.

Addressing in IMS

As part of the IMS, a new addressing space is introduced that requires a registration process. An IMS subscriber is characterized by a single private identity but may have several public user identities. The *private user identity* is a unique global identity (akin to the international mobile subscriber identity [IMSI] in GSM) associated with a user's subscription, and is predominantly used to uniquely identify a user from a network perspective. Private identities are assigned by the home network operator and are typically used for registration requests and authentication. *Public user identities* are a more human readable form of communications addresses that are used by any user requesting communication with other users. The addressing scheme adopted allows for normal telecommunications numbering in the form of a telephony uniform resource locator (URL) (e.g., +447123456789@myMobileNetwork.com), or the SIP URL format (e.g., sip:Me@myMobileNetwork.com), allowing a user to publish this public user identity to receive and initiate communications requests. Within the same IMS subscription, a user may choose to have multiple public user identities that are associated with the private user identity, for example to distinguish between a work-related subscription (e.g., workAddress@MobileProvider.com) and a personal subscription (e.g., familyAddress@MobileProvider.com). A different profile may be associated with each identity, each with different service subscriptions enabled or different service customization details. In order for a user to originate or indeed receive communications sessions in the IMS, a process of registration needs to take place for each public user identity. The registration is a means of associating a user with an S-CSCF and allows the delivery of subscriber services according to the public user identity being used. It is possible to register multiple public identities, allowing sessions to be made and received to and from the different public user identities. Details of the registration process are covered in the next section.

Registration Process in IMS

In order for a user to originate or receive any multimedia session in the IMS, the user needs to register with the network. This registration process takes place independently from registration at the radio access side with the underlying packet data network. The IMS registration process may be user- originated or automatically triggered from the user's terminal. However, expectations are that this would be a process initiated on "power on" after successful attachment to a packet-switched domain (e.g., the GPRS network).

The registration process effectively is the allocation of a serving CSCF to a public user identity. The allocation of the S-CSCF is dynamic, based on several factors, such as the load on the CSCFs in the network or the services that a user wishes to access, as implied by a particular public identity.

The details of the registration process are described in [7]. **Figure 3** provides an overview of the process, where a user sends a SIP REGISTER message (flow 1) to a P-CSCF, which may be in the home or visited network. The P-CSCF forwards this message to the I-CSCF of the user's home network (flow 2), since the private identity of the user provides this information. The I-CSCF then selects the appropriate S-CSCF and forwards the message (flow 3). As part of the registration process within the S-CSCF, the HSS is queried for the user's profile (flow 5), which includes the set of initial filter criteria giving information about the application servers that need to be involved for the user, under which circumstances each gets involved, and the relative priorities. Filter criteria are explained in the next section. The S-CSCF sends a third-party register request to each application server (flow 6) dictated by the filter criteria, and the application server can then get additional application server-specific or service-specific data from the HSS (such as privacy information), if needed (flow 7).

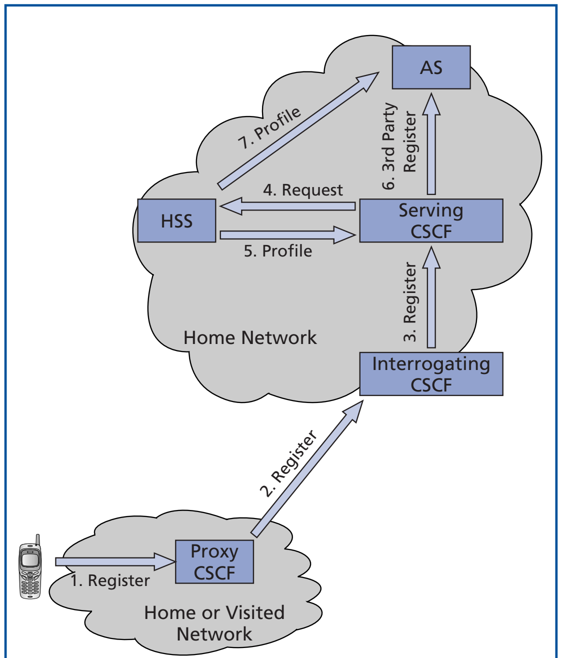


Figure 3.
IMS registration service.

Filtering Criteria and the IMS Service Control Interface

The S-CSCF uses filter criteria to involve the application servers as needed to provide services to subscribers. The filtering is done on SIP request messages such as REGISTER, INVITE, SUBSCRIBE, or BYE, but not on responses to requests. Filtering can be based on a variety of criteria such as the method of a SIP request, on whether the request is received in the originating or terminating case, on whether a particular media type is included in the session description protocol (SDP) of a request, or on the presence or content of a particular SIP header.

A specific user may get services from more than one application server. When a SIP request for a dialog comes in, the highest priority filter criteria are considered by the S-CSCF first, and, if the SIP request is selected by the criteria, it is passed as is on to the application server corresponding to the highest priority filter criteria. The application server performs service logic, may modify the SIP request, and may send the request back to the S-CSCF. The request that is the output of the first application server is subjected to the next highest priority filter criteria and, if it satisfies these criteria, it is the input to the corresponding (second) application server. This process continues until all the different priorities of filter criteria are considered or until the service logic performed in one of the application servers results in a final response to the SIP request. Once it is contacted, an application server can choose to remain in the signaling path or can remove itself from the path for subsequent requests, depending on what services the user is using (or is eligible to use).

An application server can play the role of user agent, of redirect server, of proxy server, or it may employ third-party call control [8]. This flexibility of the application server provides a platform that enables a variety of services. The S-CSCF adds an identifying indication to a request before forwarding it to an application server so that it can identify the message that comes back from the application server even if the application server has performed third-party call control and thus changed the dialog identification.

Flexibility to meet end-user needs is a hallmark of the IMS. The filtering criteria tailor which service platforms and service providers have the ability to control services on a per subscriber basis. The selected application servers, in turn, use service triggering points to apply service logic. The network operator can search everywhere to find the most attractive services to enhance subscriber satisfaction. The IMS service control (ISC) interface to the application server, because it is based on SIP, does not lock in predefined services. Instead, SIP and thus ISC provide primitives from which to build a variety of services.

Service Domain

The introduction of the IMS, the process of registering a public user identity, and the concept of filtering criteria were presented in the previous section. This section will elaborate on the application of the filtering criteria to invoke service interaction. The S-CSCF does not provide any services to the end user as it relies on the application server to provide end user services.

One key aspect of the dynamic service market is to have in place at an early stage a service architecture that is flexible, future proof, and allows a great deal of operator differentiation. This service architecture needs to be able to allow both operator-provided services as well as third-party access, both in the trusted domain (having a very close relationship with the operator) and in the untrusted domain, where security mechanisms have to be put in place by the operator to control or limit access to the network.

The service architecture defined in the IMS allows for three types of application platforms as identified in [6]. Each of these types addresses a particular need for various service deployment scenarios. First, there is the drive for rapid application development and third-party provisioning introduced by the concept of open network application programming interfaces (APIs) such as OSA and Parlay. Second, the increased usage of Internet protocols and concepts in the UMTS architecture allows for more lightweight, SIP-based mechanisms for service control. An example is SIP Call Processing Language (CPL). Third, the requirement to leverage and protect the considerable investments that exist in legacy services is met by providing a means to deliver these legacy services to end users attached to the new IMS domain. This paper's focus is on the first two types of application platforms, i.e., the OSA gateway and the SIP application server.

Application Servers in the IMS

Figure 4 shows how the IMS application servers are represented by the SIP application server and the OSA gateway. Both support the ISC interface to the S-CSCF and a Diameter interface to the HSS. The S-CSCF treats these as generic application servers making no distinction between the two. The filtering criteria determine the application servers invoked to deliver services for a particular subscriber. The SIP application server can be any SIP-based application server that supports the required interfaces. The OSA gateway is described further in the following section.

Overview of OSA

Within 3GPP, OSA is generally considered the service platform of choice for third-party application creation and deployment. OSA defines a set of APIs that allows third-party independent software vendors to use core network service capabilities as building blocks to create applications. For a more elaborate

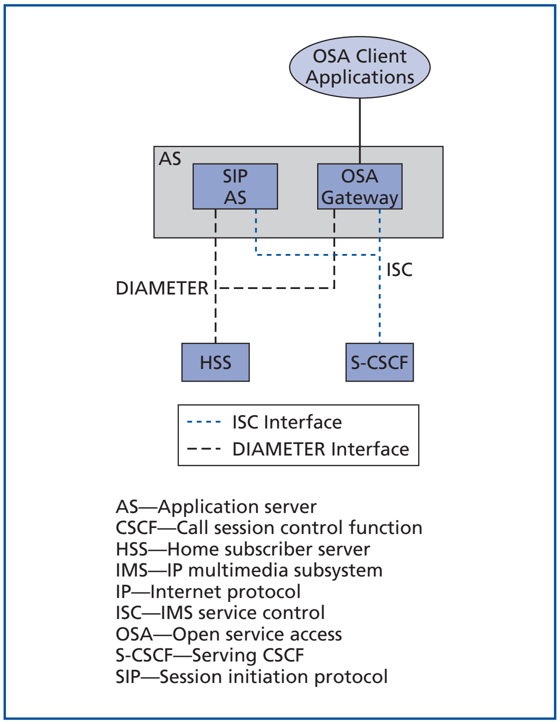


Figure 4.
Application servers in the IMS.

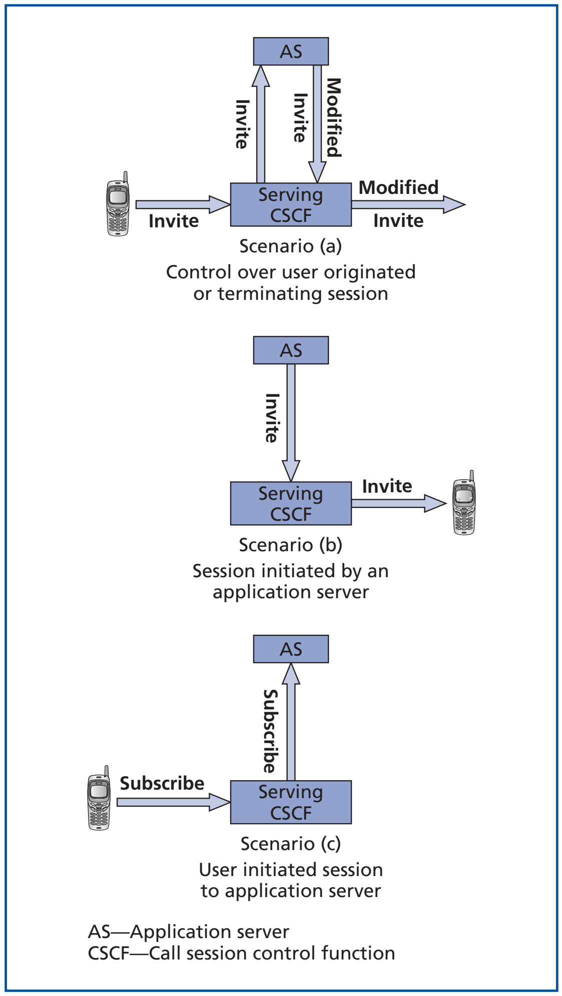


Figure 5.
Service control mechanisms.

introduction to OSA, the reader is referred to the 3GPP specification [2] or to recent papers on the subject [10, 12]. Within the IMS, access to OSA is offered through a gateway, which is seen as a special case of a SIP application server. This OSA gateway communicates with a S-CSCF through ISC, while offering an OSA API to applications.

Service Control Mechanisms

Within the IMS, it is possible to categorize the mechanisms for service interaction into three broad categories, as listed below and as illustrated in **Figure 5**, regardless of the type of application platform.

- *Service control over sessions initiated by or terminated to a user*. This is a well-established mechanism in traditional telecommunications systems, where any session originated by a user, or a session terminating to a user, could be subject to service control. An example of originating session control is a prepaid service in which the application server monitors the session (e.g., duration rated against a type of media) and terminates the session prematurely when the user runs out of credit. Session redirection is a prime example for a terminating session, where an incoming session is redirected to a different user based on predetermined settings of the user. Scenario (a) in Figure 5 shows how a user-initiated session (SIP INVITE message) may be modified by the application server.
- *Application server initiated sessions to a user*. This possibility covers the category where an application server initiates a session directly to a user (see scenario (b) in Figure 5). An example of

this mechanism is a conference service that invites pre-determined conferees to join a videoconference.

- *User initiated session directly to an application server*. This possibility covers the category where a user initiates a session or an association directly to an application server (see scenario (c) in Figure 5). An example of this mechanism is where a user requests access to presence information or a user requests the streaming of a new video clip.

IP Multimedia Service Example

This section describes a new IP multimedia service, while illustrating and highlighting control aspects of the IMS service architecture. The specific service chosen uses *presence* as one of its enablers. Before describing the service itself, a discussion will first be presented on the presence concept and the architecture being defined in 3GPP for presence within the IMS.

Presence Overview

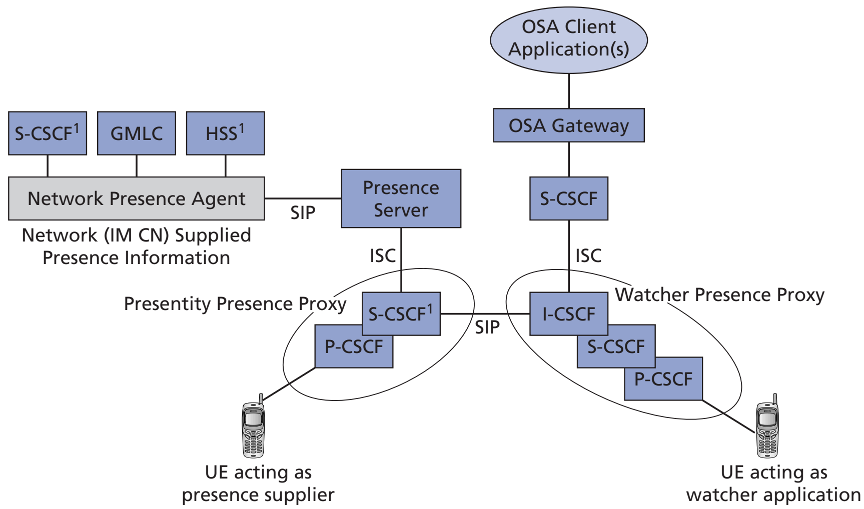
Presence provides the ability to know that a friend or colleague is "on-line" and facilitates the sending and receiving of messages more or less instantaneously. Unlike e-mail, an originating user knows that the destination user has received the message the moment the "send" button is pressed. This gives the ability for an electronic exchange of messages to take place in real time. This electronic dialogue can take many forms, such as text messages, images, files, or video clips. In order for one to be able to determine that other users are on-line in a non-intrusive manner, a mechanism to advertise this status is required together with the ability for other users to use this status information. In theory, this is a simple process where a user publishes dynamic information pertaining to his/her status. The act of publishing such information is referred to as *presence*, and the associated information is called *presence information*. Presence information is dynamic information about the subscriber or any entity that has observable, variable attributes, such as the user's present location or disposition to communicate. 3GPP is currently standardizing a presence service capability with the requirements defined in [3].

Presence entities. As in the IETF, 3GPP has adopted the term *presentity* (from presence entity) to describe an entity that publishes presence information. Generally, a presentity is associated with a 3GPP subscriber, but can be associated with a more abstract entity, such as a weather station. A centralized network node known as a presence server may distribute this information to interested parties. A consumer of presence information is described as a *watcher*. A watcher may be an application residing in an application server or in a user's mobile terminal. Watchers watch presentities, under the restrictions imposed by the privacy rules determined by the latter. The interested reader is referred to [9] for more information on the IETF work on SIP for instant messaging and presence leveraging (SIMPLE).

Presence architecture within the IMS. An overview of the 3GPP presence service architecture is shown in **Figure 6**, where only support within the IMS is shown. Central to this architecture is the presence server, which collects and distributes presence information. In reality, one would anticipate that there would be multiple presence servers for easier management and load sharing. The presence server is considered a SIP application server in the general IMS architecture, where it communicates with S-CSCFs through the ISC interface. Potentially, presence information may be supplied by network nodes (such as the S-CSCF) as well as by direct user input. 3GPP introduces the concept of a network presence agent in order to convert network presence information into a uniform protocol (SIP) to the presence server.

3GPP introduces the concept of watcher presence proxy and presentity presence proxy providing address resolution, routing, identification of the correct network of the presence server associated with presentity, accounting, and interworking with external or noncompliant networks. However, all functionality of these proxies, with the exception of interworking, is provided by combinations of the P-CSCF/I-CSCF/S-CSCF, as shown in Figure 6. Interworking functions can be provided by a dedicated application server, but this is not covered in this paper.

Presence messages and call flows. To amplify the strength of the IMS service architecture and the ISC



[1] Indicates same S-CSCF duplicated for clarity

3GPP—3rd Generation Partnership Project
CN—Core network
CSCF—Call session control function
GMLC—Gateway mobile location center
HSS—Home subscriber server
I-CSCF—Interrogating CSCF
IM—IP Multimedia
IMS—IP multimedia subsystem
IP—Internet protocol
ISC—IMS service control
OSA—Open service access
P-CSCF—Proxy CSCF
S-CSCF—Serving CSCF
SIP—Session initiation protocol
UE—User equipment

Figure 6.
3GPP presence service architecture in IMS.

interface, this section shows two service examples built on the presence service. The examples make use of call flows, where time progresses from top to bottom. The two examples are based on the two types of application servers of interest, i.e., the SIP application server and the OSA gateway. Although the difference goes beyond the choice for application server, these examples will demonstrate that the architecture and ISC interface can very aptly support both. A SIP-based watcher application residing in a UE watches the presence information of a presentity on the end user's behalf. For example, the application can be an instant messaging buddy list. In the OSA application server case, the watcher is a third-party OSA application. The watcher monitors presence information on behalf of all subscribers to the OSA application. The application could be a conference scheduler, e.g., whenever a session initiation trigger is received, the application will check the availability of all conferees and set up the conference accordingly. **Figure 7** shows how a SIP-based watcher application located in a user's terminal (labeled UE) may request to receive presence information about another user (flow 1). S-CSCF(1) represents the S-CSCF with which the watcher is registered, while S-CSCF(2) is the S-CSCF with which the user being watched is registered.

Potentially, S-CSCF(1) and S-CSCF(2) may be located in different networks. The UE's (i.e., the watcher's) initial request to be notified of presence information for a specific presentity (flow 1) is received at the P-CSCF, as it is the UE's contact point into the IMS. The P-CSCF relays the request to the S-CSCF to which the UE is registered, i.e., S-CSCF(1). In turn, S-CSCF(1) forwards the request to the contact

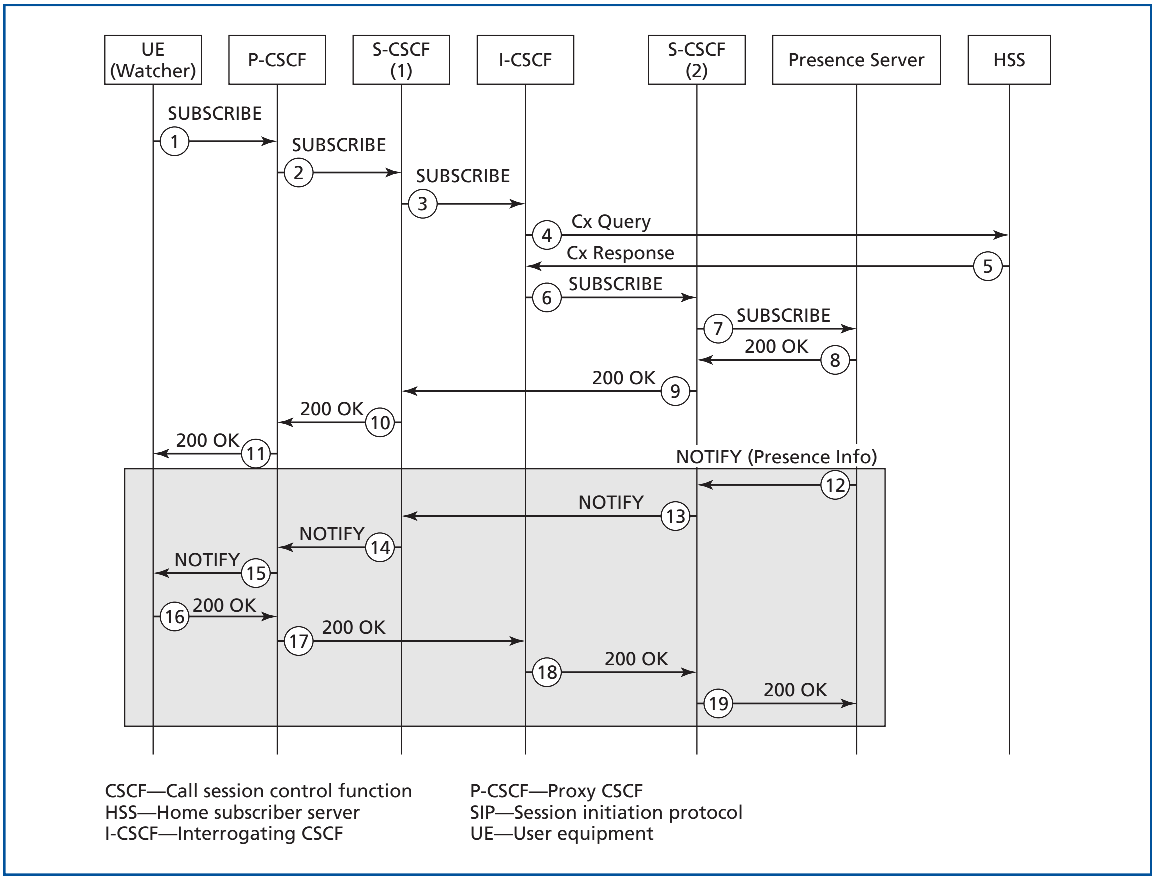


Figure 7.
Flows for SIP-based watcher.

point into the presentity's network, i.e., the I-CSCF. The I-CSCF will interrogate the HSS to obtain the address of the S-CSCF with which the presentity is registered (flows 4 and 5). Once this address is obtained, the I-CSCF can forward the request to S-CSCF(2). Since S-CSCF(2) does not perform any service logic itself, based on its filter criteria, the watcher's request is forwarded to the presence server (flow 7). Subsequently, the presence server will acknowledge the request for notification of presence information. On the return path for the acknowledgement, S-CSCF(2) can bypass the I-CSCF (flow 9), since in this case, the routing information to the S-CSCF(1) has been included in the request by S-CSCF(1) (flow 3). The acknowledgement is passed through the P-CSCF back to the SIP-based watcher application in the UE. Once the presence information for the presentity has been collected, or a change in the presence information has occurred, the presence server will send the information to the UE, through the S-CSCF(2), S-CSCF(1), and P-CSCF (flows 12 to 15). The UE will acknowledge the receipt of this notification back to the presence server (flows 16 to 19). There are no flows to show how a UE updates the presence server with presence information. At the time of writing, IETF and 3GPP had not agreed on a standardized mechanism as to how this is achieved.

Figure 8 shows how a third-party OSA application may request to receive presence information about another user. S-CSCF(1) represents the S-CSCF

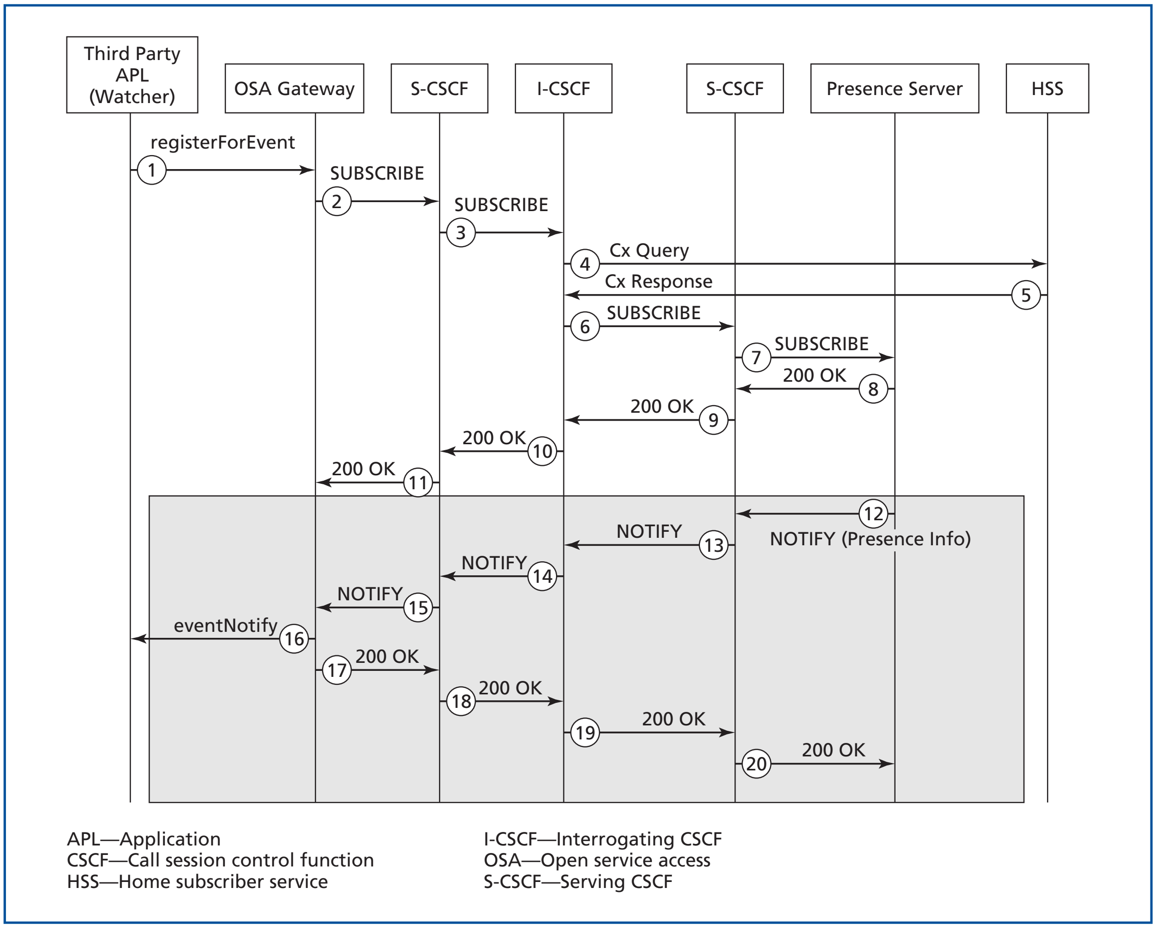


Figure 8.
Flows for an OSA-based watcher.

that the OSA gateway has access to, and is independent of any registered users. As in the previous example, S-CSCF(2) represents the S-CSCF with which the user being watched is registered. The third-party OSA application is not a SIP application. Instead, it communicates through the OSA API interface with the OSA gateway, which, in line with the IMS service architecture, is a SIP server. The third-party OSA application acts as a watcher. As a result, the OSA API methods need to be mapped onto the SIP requests and vice versa. The OSA gateway performs this functionality, as is described in [12]. The OSA API method *registerForEvent* (flow 1) allows an OSA client application to register its interest in a certain presence-related event. An example of an event could be a change in presence information for a certain subscriber on an end user's buddy list. The OSA API method *eventNotify* (flow 14) will report the occurrence of the event in the network up to the third-party OSA application. Details of the OSA API for presence can be found in [6]. Once the OSA gateway has mapped the OSA watcher request (flow 1) onto a SIP watcher request (flow 2), the request is forwarded to the S-CSCF(1) to which the OSA gateway has access. Entirely in line with the IMS architecture, S-CSCF(1) forwards the request to the contact point into the presentity's network, i.e., the I-CSCF (flow 3). From this point onward, the call flow is exactly the same as the one for the SIP-based watcher application. Once the OSA gateway receives the notification from the presence server (flow 13), this needs to be mapped onto the *eventNotify* method.

These call flows presented two quite distinct service examples, i.e., a SIP-based watcher application residing on a UE on behalf of a single subscriber and a third-party OSA-based watcher application residing on an application server outside the operators' trusted domain on behalf of a whole set of subscribers. Nonetheless, the architecture and call flows support both examples fairly effortlessly and straightforwardly. In fact, from the perspective of the IMS network elements, both examples, although supporting completely different business and deployment models, look extraordinarily similar. These examples are unmistakable substantiation of the flexibility and strength of the IMS service architecture and the ISC interface.

Description of a Presence-Enabled IP Multimedia Service

In today's well-known presence-enabled service of instant messaging, presence information of subscribers to the service is provided to other subscribers giving them additional data for making communication decisions. New presence-enabled services of the future may also provide end users with presence information of other end users (or entities), as with the instant messaging service. Alternatively, presence information may be used to provide other enhanced services. In other words, presence information may be made available to application servers as well as to end users. The service to be described here is such a service.

This service example is a presence-enhanced call-forwarding feature. Today's call-forwarding service makes rerouting decisions based on criteria such as time of day, day of the week, immediate, busy, and no answer. The rerouted target endpoint can be another phone number or a message center. Depending on the media type of the calling party, the message center being rerouted to could be a voice message center, video message center, text message center, or unified message center.

With presence information, call forwarding can be enhanced in a variety of ways. The following illustrates some of those ways. Let us say that the called party is present and available as MyWork@MyITSP.com, but not as MyHome@MyITSP.com. Then, although both public identifications (IDs) map to the same private ID (or handset), those calls to MyHome@ITSP.com can be forwarded to a message center. In this way, all public IDs can automatically register with the IMS at power-up of the physical handset, but call management of these multiple IDs is still possible by the end user.

Now, let us say the called party is present but unavailable and "grumpy." This presence information can be provided to the calling party prior to being rerouted to a message center assuming permissions are set up to allow for this. The additional information gives the calling party more data for deciding on future communication actions.

As another enhancement on call forwarding, it may also be possible to allow the called party to specify select calling parties contacting MyHome@MyITSP.com to be "forwarded" to MyWork@MyITSP.com.

An existing SIP twist on call forwarding is the serial and parallel forking features. SIP allows for a proxy server to forward INVITEs to multiple endpoints simultaneously or serially in order to find the called party. The session is set up with the endpoint that answers first. Presence information could be used by the IMS in making routing decisions for a call as an alternative or enhancement to the SIP serial and parallel forking features. With presence data, serial and parallel forking may be reduced in some scenarios.

IMS and Service Architecture Walk-Through

The call flow in **Figure 9** details the messaging through the IMS and presence architecture used in providing the presence-enhanced call forwarding service. In this scenario, the application server with the call forwarding feature is the watcher. It subscribes to the presence server for presence information of the UE subscriber (presentity). When a presence event occurs, such as when the UE subscriber updates his/her presence information, a NOTIFY is sent to the

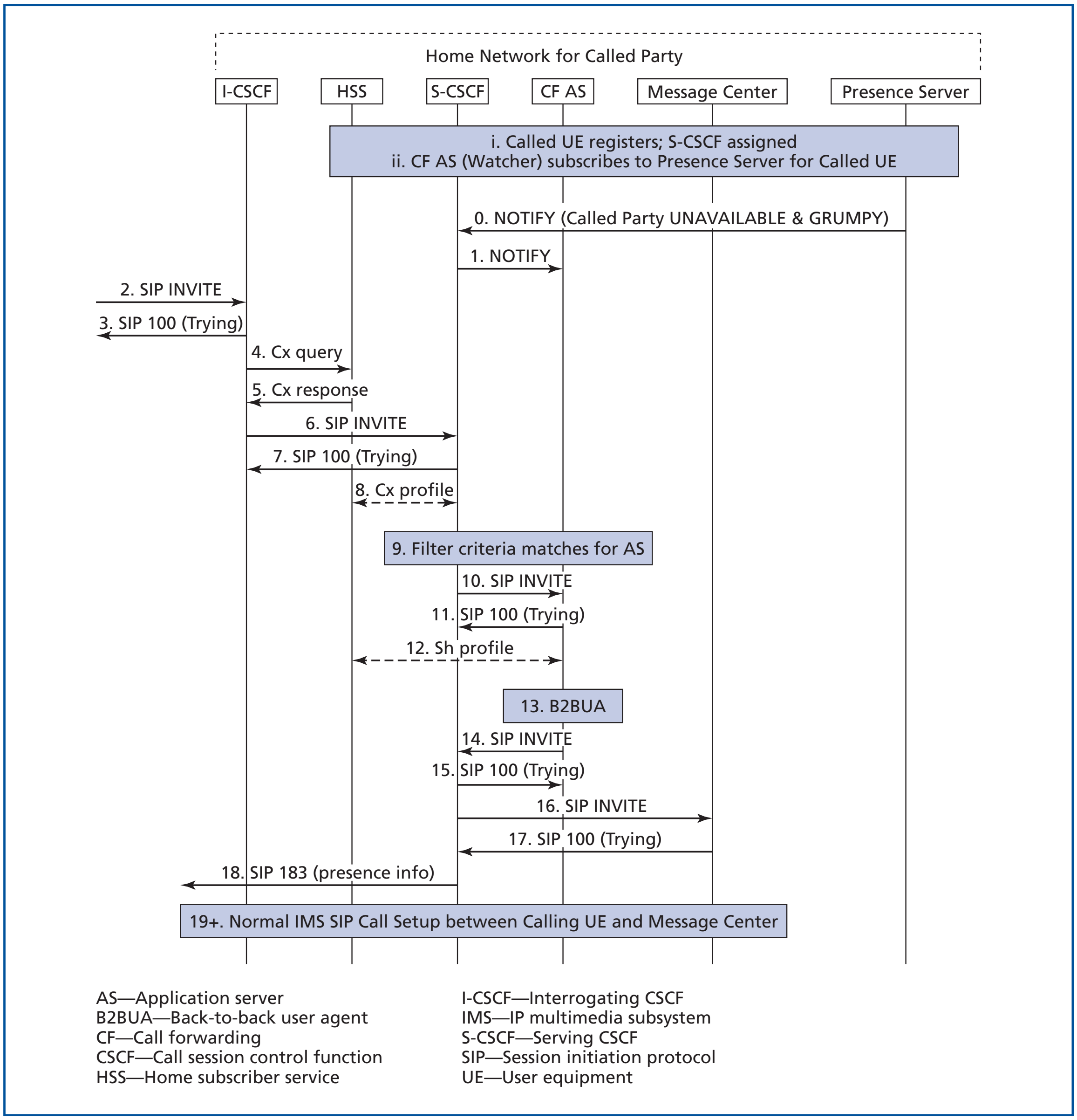


Figure 9.
Presence service example.

call-forwarding (CF) application server. In this example, the UE subscriber updates his/her presence status as "present but unavailable and grumpy." The CF application server, in turn, updates its subscriber profile records appropriately.

When a call comes in for the UE subscriber (steps 2–9), the S-CSCF concludes that it needs to forward the SIP INVITE to the CF application server based on filter criteria that it retrieved from the HSS. The CF application server reviews its stored presence information for the called party. Previously provisioned actions map the presence status of "present but unavailable and grumpy" to "forward calling party to message center." Based on this information, the CF application server forwards the calling party to the UE subscriber's message center. The CF application server acts as a back-to-back user agent (B2BUA) [8] in setting up this session, steps 13–19+.

The CF application server also notes that the UE subscriber gives permission for calling parties to get his/her presence status. Therefore, the CF application server sends back to the calling party, the UE subscriber's presence info (step 18). The presence info is not supplied in the SIP NOTIFY message, as was described in Figures 7 and 8, since in this scenario the calling party does not act as a watcher. For the sake of brevity, the call flow in Figure 9 does not show all SIP response messages, e.g., the 200 OK after the NOTIFY.

Conclusion

The IMS aims to bring mixed media capabilities to wireless users where, through a single identity, the user may initiate or receive real-time as well as non-real-time communications with other users or applications. Through this paper, we have explored the interaction of the IMS and the application servers in the form of OSA gateways and SIP application servers. One key aspect of the architecture is that the ISC interface allows the network nodes in the IMS to be able to interact with different types of application servers in a uniform manner. The example of a presence-enhanced call-forwarding service provided in the paper has demonstrated the flexibility of the IMS service architecture. This flexibility is key to the successful deployment of IMS.

*Trademarks

CDMA2000 is a trademark of the Telecommunications Industry Association.

References

[1] 3rd Generation Partnership Project, <http://www.3gpp.org>.

[2] 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects; Virtual Home Environment/Open Service Access (Release 5), 3GPP TS 23.127, June 2002, <ftp://ftp.3gpp.org/Specs/2002-12/Rel-5/23_series/>.

[3] 3rd Generation Partnership Project, Technical Specification Group Services and Systems Aspects, "Presence Service, Stage 1 (Release 6)," 3GPP TS 23.141, Jan. 2003 <ftp://ftp.3gpp.org/Specs/2002-12/Rel-6/23_series/>.

[4] 3rd Generation Partnership Project, Technical Specification Group Core Network, "IP Multimedia (IM) Session Handling; IP Multimedia (IM) Call Model; Stage 2 (Release 5)," 3GPP TS 23.218, Dec. 2002, <ftp://ftp.3gpp.org/Specs/2002-12/Rel-5/23_series/>.

[5] 3rd Generation Partnership Project, Technical Specification Group Services and Systems Aspects, "IP Multimedia Subsystem (IMS); Stage 2 (Release 5)," 3GPP TS 23.228, Jan. 2003,<ftp://ftp.3gpp.org/Specs/2002-12/Rel-5/23_series/>.

[6] 3rd Generation Partnership Project, Technical Specification Group Core Network, Open Service Access (OSA), Application Programming Interface (API), Part 14: Presence and Availability Management (Release 5), 3GPP TS 29.198, Oct. 2002,<ftp://ftp.3gpp.org/Specs/2002-12/Rel-5/29_series/>.

[7] J. G. Adamek, E. H. Henrikson, H. A. Lassers, A. Y. Lee, and R. B. Martin, "Services and Technical Considerations for the Wireless IP Multimedia Subsystem," Bell Labs Tech. J., 7:2 (2002), 91–104.

[8] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP: Session Initiated Protocol," Internet Engineering Task Force, IETF RFC 3261, June 2002, <http://www.ietf.org/rfc/rfc3261.txt?number=3261>.

[9] J. Rosenberg, "Session Initiation Protocol (SIP) Extensions for Presence," Internet Engineering Task Force, Internet Draft, May 2002, <http://www.ietf.org/internet-drafts/draft-ietf-simple-presence-10.txt>.

[10] R. Stretch, "The OSA API and Other Related Issues," BT Tech. J., 19:1 (2001), 80–87.

[11] UMTS Forum, Report No 20, IMS Service Vision for 3G, Apr. 2002, <http://www.umts-forum.org/reports.html>.

[12] M. R. Unmehopa, M. L. F. Grech, J. A. Dobrowolski, and J. R. Stanaway, "The Support

of Mobile Internet Applications in UMTS Networks Through the Open Service Access," Bell Labs Tech. J., 6:2 (2001), 47–64.

(Manuscript approved March 2003)

LYNELL E. CANNELL is a distinguished member of technical staff in the Circuit and IP Core Systems Engineering Department at Lucent Technologies in Naperville, Illinois. Her current job responsibilities include defining requirements for third-generation CDMA/UMTS wireless networks. She holds a B.S. in the teaching of physics, and an M.S. and a Ph.D. in physics from the University of Illinois at Urbana-Champaign. Dr. Cannell has been awarded one patent.

MICHEL L. F. GRECH is a technical manager in the GSM/UMTS Standards Department in Mobility Solutions at Lucent Technologies in London, United Kingdom. He is responsible for a team driving service-related and regulatory standards encompassing 3GPP, the Open Mobile Alliance, the GSM Association, the UMTS Forum, and various other European regulatory forums. Mr. Grech holds a B.Sc. in electronic engineering from the University of Essex in Colchester, England, and an M.Sc. in telecommunications technology from Aston University in Birmingham, England. A member of IEE and a Chartered Engineer, he holds seven patents and has authored several technical papers.

ANNE Y. LEE is a distinguished member of technical staff in the Circuit and IP Core Systems Engineering Department in Mobility Solutions at Lucent Technologies in Naperville, Illinois. She currently heads a team of systems engineers in defining requirements for a converged 3GPP/3GPP2 IMS system as well as a cross-functional team in developing IMS/SIP service demos and trials. Ms. Lee holds a B.S. in electrical engineering and an M.S. in computer sciences in the Illinois Institute of Technology in Chicago.

MUSA R. UNMEHOPA is a standards consultant in the UMTS Standards Department of the Wireless Advanced Technologies Laboratory at Lucent Technologies in Hilversum, The Netherlands. He is the Lucent standards delegation manager for Service Mediation Standards in 3GPP, Parlay, ETSI, and the Open Mobile Alliance (OMA). He is also the vice-chairman of 3GPP Core Network Group 5 (Open Service Access,) and he leads the Parlay Task Force in the OMA. Mr. Unmehopa holds an M.Sc. in computer science from the University of Twente in the Netherlands. His research interests include service delivery and mediation architectures in next-generation networks. ◆