AN OPEN STANDARDS-BASED FRAMEWORK INTEGRATING IMS AND CLOUD COMPUTING
Wei Zhang, Weimin Lei, Xiao Chen and Shaowei Liu
Institute of Communication and Information System
College of Information Science and Engineering, Northeastern University, Shenyang, China
zhangwei1@ise.neu.edu.cn

Abstract
As typical Internet-based applications, cloud computing services cannot guarantee the Quality of Service (QoS), which has actually become technical barriers for cloud service providers to overcome. On the other hand, as the core signaling architecture of Next Generation Networking (NGN), IP Multimedia Subsystem (IMS) is facing the problem of the lack of innovative value-added services. This paper presents an open standards-based framework to integrate IMS and cloud computing. In the target framework, cloud services are regarded as the general IMS applications and then cloud clients are allowed to access cloud services under the control of Session Initiation Protocol (SIP) signaling and QoS mechanism of IMS. This paper highlights the benefits of such integration, introduces architecture overview and cloud service relevant functional components, and mainly discusses several key issues including cloud notification mechanism, standard cloud interface protocol, QoS and charging control of IMS cloud computing services. The work in this paper provides a reference for integrating IMS and cloud computing.

Keywords: Cloud Computing, IP Multimedia Subsystem, Session Initiation Protocol, Session Description Protocol, Quality of Service, Charging

1. INTRODUCTION
Cloud computing is changing the way of developing, deploying and managing applications. It provides the powerful computing power and nearly unlimited storage space by sharing hardware and software resources, which facilitates the rapid deployment of a variety of innovative services in a cost effective manner.

Meanwhile, as the applications delivered as services over the Internet, cloud computing is still facing significant technical challenges in meeting the customers' requirements. No matter what type of cloud services provided, cloud computing systems put forward a high demand for network bandwidth. As Mell and Grance (2011) stated in the NIST definition, one of key characteristics of cloud computing is broad network access. For individual users and small businesses using public cloud services, the quality of the service is determined in large part by their own broadband connections over which public cloud service provider has no control. However, Internet's best effort strategy, which is only appropriate for the early internet services, is difficult to guarantee QoS for cloud customers, thus diminishes the quality of user experience.

In addition, cloud computing lacks a complete signaling control mechanism. Almost all of the cloud platforms extend to the client via web service interfaces. However, the web-based access mode cannot effectively support service access control, various pricing models and other mechanisms. Therefore, it is difficult for telecom operators to provide commercial public cloud services.

Telecom operators could play an important role in the development of cloud computing by making the most of their natural advantages as network operators. The IP multimedia subsystem (IMS), initially introduced by 3GPP, now is evolved to be the core signaling architecture of Next Generation Networking (NGN) for multimedia services and has been widely deployed by telecom operators throughout the world. Its IP infrastructure and highly componentized nature get well with cloud services. IMS has the following promising features:

1) It has open and standardized signaling control mechanisms.
2) It can provide robust multimedia services with guaranteed QoS, charging model, reliability and security.

Even though there have been tremendous efforts to develop the architecture and protocols to support advanced Internet-based services, IMS has not been used on a wide scale due to the lack of innovative services yet. Thus, the development of application servers supporting value-added services would be a priority in IMS.

This paper presents a framework to integrate IMS and cloud computing. Figure 1 shows the general implementation of such integration. As shown in Fig. 1, IMS provides an open and standard service platform on which service providers will build various services. Cloud services are regarded as the general IMS applications. Combined with the characteristics of IMS and cloud computing, such integration will bring them both explosive growths. Based on cloud computing where most applications run in the cloud, user terminals only need capacity such as internet access, audio/video decoding and interactive processing. Thus, cloud computing technology will lead to rapid developments of IMS value-added services.

http://www.hipore.com/ijcc
doi: 10.29268/stcc.2014.2.3.3
Meanwhile, IMS provides the most significant opportunities for cloud computing:

1) Open and standardized signaling control. IMS provides perfect signaling control mechanisms which can implement fine-grained service access control such as digital copyright protection, charging, security.

2) Optional, negotiable and differential QoS control. Interacting with the network elements that transport application flows, IMS can offer negotiable QoS for IP multimedia sessions not only at the time of establishment, but also during the session. And IMS can provide different qualities depending on user profile, location, access network and devices, etc.

3) Service Reusability. The existing IMS services including presence, group management, authentication, and capability negotiation could be exposed via standardized interfaces to the cloud services. In addition, some basic cloud services should also be made available to other complicated cloud services.

4) Promoting standardization of cloud interface protocols. To be deployed based on IMS architecture, all cloud services should follow the uniformly standard interfaces, which will promote standardization work of cloud computing.

The remainder of the paper is organized as follows. Section II describes the background and related work in IMS and cloud computing. Section III presents the architecture and principle of the integration framework of IMS and cloud computing. Section IV discusses several key issues of cloud notification service, cloud interface protocol, QoS and charging control for cloud services integrated in IMS and section V concludes the paper and discusses future work.

2. BACKGROUND AND RELATED WORK

IMS was introduced in the 3GPP architecture release 5 to support IP multimedia services. According to Agrawal, Yeh, Chen and Zhang (2008), the collaboration between IETF, 3GPP, and 3GPP2 is essential to the development of IMS for next-generation networks (p. 138). Now IMS is evolved to support IP multimedia services in the integrated all-IP networks. IMS enables faster integration of subsequent value-added services that take advantage of the current deployed infrastructure and all the core services it provides. IMS interacts with the network elements that transport application flows to achieve QoS for the flows of the IMS controlled applications. IMS provides the capacity of correlating charging at network and service level and supports a variety of charging modes, such as online and offline charging, time based and volume based charging, and so could meet various charging requirements. To obtain openness and interoperability, IMS exploits open standard IETF protocols wherever possible. In particular, IMS uses the Session Initiation Protocol (SIP) for signaling and session management. SIP is an application layer signaling protocol for establishing and tearing down multimedia sessions such as Internet conferencing, telephony, presence, event notification and instant messaging.
Cloud computing has been attracting more and more attention from both academia and industry. According to Armbrust and et al. (2009), cloud computing providers offer their services based on three fundamental models: Infrastructure as a Service (IaaS), such as Amazon EC2 and GoGrid, where users are able to deploy and run arbitrary softwares including operating systems and applications; Platform as a Service (PaaS), such as Salesforce, where users are provided with remote software platforms to run their services; and Software as a Service (SaaS), such as Google App Engine, where users can use the provider's applications running on a cloud infrastructure.

There are more and more public clouds and as well as private clouds built by institutions and companies. Unfortunately, these clouds are often incompatible from each other, which increases the complication of porting applications between clouds as well as introduces potential vendor lock-in. Most recently, many efforts have been made to develop open standards for cloud computing, mostly notable by Open Grid Forum (OGF) with Open Cloud Computing Interface (OCCI), DMTF’s Open Virtualization Format (OVF) and etc.

In the area of the cooperation between NGN and cloud computing, the ETSI Technical Committee (TC) GRID (2009) considered interoperability gaps between grid/cloud technologies and the NGN architecture in an integrated environment and proposes solutions to the identified gaps. Gouveia et al. (2009) discussed the possibility to integrate NGN/IMS and Cloud computing, which offers two interconnection scenarios for combining Cloud and NGN applications in a unified architecture. Chen et al. (2011) introduced an architecture where IaaS is the development platform used for building the IMS system. Bellavista et al. (2012) studied the elastic provisioning of IMS-based telecommunication services on top of multiple cloud platforms.

This paper explains the feasibility of integrating various cloud computing services into the IMS architecture as general applications. The challenges currently faced by cloud computing, such as best-effort QoS, weak charging and interoperability barrier, are expected to be solved or relieved via the signaling control and policy control mechanism of IMS. Meanwhile various innovative value-added services are brought to IMS, which shall contribute the development of both IMS and cloud computing. According to our best knowledge, this paper is the first on reporting the framework that integrates IMS and cloud computing.

3. Framework Scheme
3.1 Architecture Overview

This paper presents a framework integrating cloud computing and IMS. In this framework, IMS provides an open and standard service platform and performs the uniform service control with the subscriber service profile. Various cloud services can be used as common IMS-enabled services. To enhance the feasibility of the presented framework, we takes advantage of existing functional components and protocols in IMS as well as other standard works about cloud computing. IMS terminals only need a few modifications to access various cloud services.

Figure 2 illustrates the functional architecture of cloud computing integrated into IMS, which extends the current IMS specification with the required functionality to meet additional requirements of cloud services. The functional architecture contains main functions and reference points. The main functions associated to cloud services can be divided into two parts: cloud service functions and cloud interaction functions. The core IMS forwards the complete SIP signaling for session management and service notification of cloud services. The data flows of actual interaction between IMS User Equipment (UE) and cloud platform do not traverse the core IMS. This architecture supports the deployment of cloud services in a multi-provider environment.

The UE communicates with the cloud platform over multiple interfaces for different purposes, namely, over a Gm interface via the IMS core for the session management and service perception purpose, over a Ut interface for user profile configuration purpose, or over a Xd interface for actual interaction with cloud platform purpose. These interfaces are compatible with 3GPP IMS specifications.

User data involved in providing cloud services can be classified into two categories: IMS profile and cloud specific profile. IMS profile includes all information required to establish IMS sessions and access cloud services hosted in application servers. Cloud specific profile encompasses all information required to operate a cloud service, such as the list of subscribed cloud services. IMS profile information is located in Home Subscribed Server (HSS). Cloud specific profile information can be located in the dedicated databases, the application servers hosting cloud services, or HSS. In our framework, we use the term HSS for its simplicity. HSS communicates with cloud service functions at Sh reference points and communicates with the core IMS at the Cx reference point. When multiple instances of HSS occur, the core IMS and the cloud service functions may use the services of Subscription Locator Function (SLF) to fetch the address of the HSS at the Dh and Dx reference points respectively.

3.2 Service Notification, Profile Configuration and Service Control Functions

http://www.hipore.com/ijcc
Cloud Service Notification Function (SNF) is designed for service information maintenance and management. Cloud customers require the information about all of the currently available cloud services and subscribed cloud services, and detailed information about each cloud service. Therefore, a means to collect and provide the information of cloud services provided by cloud platforms for cloud clients needs to be offered. The Cloud SNF is responsible for providing information of accessible cloud services. For each cloud service, the following data should be collected and provided:

1) A list of URIs or IP addresses of Profile Configuration Functions and Service Control Functions.

2) The set of network parameters that may be required by the UE to activate and establish cloud interaction channels.

3) User readable data related to the cloud service such as the information about charging policy and vendor.

The UE can obtain the hostname or IP address of the Cloud SNF by using manual configuration or during network attachment. Once a user has been successfully registered, the UE sends a SIP request which is routed to the Cloud SNF through the IMS core. The Cloud SNF derives service information, and then sends downstream to the UE.

Cloud Service Control Function (SCF) handles cloud-related requests and executes service and session control for all cloud services. The UE communicates with the Cloud SCF via the core IMS for the purpose of session management. In other words, before the UE can interact with the cloud platform, a session initiation procedure shall take place between UE and Cloud SCF. The Cloud SCF uses the IMS service control (ISC) interface to communicate with the IMS core. General tasks of the Cloud SCF are summarized as follows:

1) Service authorization and validation for the purpose of granting or denying a user request for the initialization of a new session or the modification of an existing one.

2) Selecting a relevant cloud interaction server.

3) Charging control with IMS offline or online charging system.

3.3 CLOUD INTERACTION FUNCTIONS

Cloud interaction functions include Interaction Control Functions (ICF) and Interaction Process Functions (IPF). In a large cloud platform, there may be hundreds of thousands of or even millions of machines to provide cloud services. An important design principle for interaction functions is to realize a high-concurrency and distributed process architecture. A Cloud ICF controls multiple Cloud IPFs. The Xp reference point between Cloud ICF and Cloud IPF is used to control media session setup and modification. The y2 reference point between S-CSCF and Cloud ICF carries cloud service control signaling originating from Cloud SCF.
to control Cloud ICF. The main tasks of Cloud ICF are summarized as follows:

1) Handling signaling control received from one or multiple Cloud SCFs.
2) Selecting an appropriate server with Cloud IPF.
3) Collecting statistical information and generating charging information about service usages, e.g. for end-user charging based on the viewed content.

Cloud IPFs are responsible mainly for the actual interaction process with UE via the Xd reference point. Most of cloud platforms offer cloud interfaces for users to manage cloud services. For example, Amazon EC2 provides three categories of cloud interfaces: graphical portals, command line clients and web services. A command “ec2-

run-instances ami-xxx -k ec2-

keypair” will create a virtual machine. If users want to connect to the machine, they first use command “ec2-describe-instance” to find the public IP address or public machine name of the virtual machine, and then use SSH command to connect to the machine using the IP address. Therefore the actual interaction between cloud client and cloud platform can be divided into two categories: cloud management interaction performed by Management Interaction Process Function (MIPF) and cloud service interaction performed by Service Interaction Process Function (SIPF).

When an interaction process between UE and cloud platform is involved in cloud service, a Cloud IPF will be selected during the establishment process of SIP session. The criteria for this selection differ greatly for different cloud service types. The criteria may include the UE’s location, load balance policy and so on. The selection process can be divided into two steps: the selection of the Cloud ICF by the Cloud SCF and the selection of the Cloud IPF by the Cloud ICF. In some cases, the selection task may become very simple. For example, if the cloud customer wants to interact with the assigned virtual machine in IaaS, the Request-URI of INVITE message which is responsible for session establishment and modification in SIP protocol may include the IP address of the virtual machine which corresponds to just one server with Cloud MIPF.

3.4 DEPLOYMENT SCENARIO

Figure 1 shows a typical deployment scenario of the cloud computing integrated with IMS. Cloud services are available to the IMS core network and end-users in the same way of other general IMS services. As a uniform service platform, IMS offer unified management and control for all services. Existing standardized reference points and functional elements are used to facilitate the interaction. The existing IMS common service capabilities, such as presence and group management, could be exposed via standardized interfaces to the cloud services. In the same way, some basic cloud services, such as Amazon EC2 and S3, can also be reused to create complicated cloud services. The end-users only need one authentication for access to all authorized IMS services including various cloud services, and obtain the contracted quality of experience.

Cloud Application Servers (or ASs for short, with SCFs, SNFs and/or PCFs) manage and control the interaction between cloud clients and cloud platform. From the perspective of IMS core network, the Cloud ASs are the portal to the corresponding cloud platform. The deployment of Cloud ASs makes the actual interaction between cloud clients and cloud platform under the unified control of standardized signaling mechanisms of IMS and thus provides possibility for the use of IMS QoS and charging mechanism.

Cloud SCF ASs and Cloud ICF Interaction Servers (or ISs for short), need to distinguish between two categories of cloud interactions. There are several approaches that can be used. One alternative way is to extend a new header parameter or a new private header in INVITE message. The extended parameter or header is used to indicate the category of cloud interaction to be established. For cloud service interaction, the INVITE message also generally needs to carry the address information of the Cloud SIPF IS which may have been assigned to the cloud client. In IaaS, for example, Cloud SIPF IS could be a virtual machine created by the cloud client. The information of Cloud SIPF IS that cloud client has access to can be learned during the process of cloud management interaction.

4. SEVERAL KEY ISSUES

4.1 CLOUD NOTIFICATION SERVICE

The SIP events framework defines general mechanisms for subscribing to, and receiving notifications of, events with SIP systems. This paper defines a new SIP event package for cloud notification service which is for subscription and notification of cloud service information. The new event package is similar to the existing event package such as user presence and call state.

After successful registration process with IMS core, UE sends a SUBSCRIBE message with cloud event package to Cloud SNF AS. In a multi-provider environment, a common Cloud Notification Application Server can be deployed. As shown in Figure 3, the cloud clients can SUBSCRIBE to a common domain URI such as “sip: domain.com” or a dedicated URI such as “sip:cloudservice@domain.com”, and be connected to the cloud notification service provided by the common Cloud Notification AS. Cloud SNF AS associated to a cloud platform makes use of PUBLISH messages to perform the state publication of all cloud services provided by the associated cloud platform. All subscriptions to, and publications of, cloud event package are routed to the

http://www.hipore.com/ijcc

33
common Cloud Notification AS. The common Cloud Notification AS combines the separate cloud event partial states published by multiple Cloud SNF ASs into a whole cloud event state and acts as the notifier distributing cloud state to the interested cloud clients in the form of SIP NOTIFY messages with an XML payload. When a new cloud service provided or an existing cloud service removed, this state changes and is reported through the notification service. Through this mechanism, cloud users will learn all information of available cloud services offered by the whole domain and select the appropriate cloud service according to the state information received and their own demands, such as charging policy, the preferred service vendor and QoS requirements.

This paper defines Cloud Information Data Format (CIDF) as a common cloud service data format and also defines a new media type "application/cidf+xml" to represent the XML MIME entity for CIDF. In cloud event package, the body of the PUBLISH/NOTIFY requests contains a cloud service information document which is formatted into "application/cidf+xml" media type.

Figure 4 gives the non-normative diagram of the overall hierarchy of CIDF. A cloud service information document begins with a root element tag <cloudservices> which is comprised of zero, one or more <cloudservice> child elements. Each <cloudservice> element describes a single cloud service and is comprised of <service-description> and <service-state> child elements. The <service-description> element describes the static information of a cloud service such as service type (IaaS, PaaS or SaaS), service vendor, service URI (be used in order to access the cloud service), the charging policy and so on. The <service-state> element indicates the dynamic information of a cloud service such as the current user count and whether the cloud service is currently active.

4.2 Uniformly Standard Cloud Interface Protocol

In order to solve the interoperability problem between cloud platforms, it is necessary to define a common and standardized cloud interface protocol that hides the private implementation details of each cloud platform while meeting all the requirements needed by present and future cloud services. However, cloud computing offers are generally categorized in several types which possess very different process modes from each other type. So, it is a big challenge to develop a uniform cloud interface protocol supporting various kinds of clouds offered by different vendors. How to quickly deal with this issue is not by re-inventing the wheel, but rather by reusing existing research work. As stated before, the OGF OCCI group is developing a clean, open protocol for cloud computing.

OCCI is a boundary protocol and API that acts as a service front-end to a provider’s internal management framework, hiding the details of how to deploy and manage the applications for each specific cloud (Nyren, 2011). Initially, OCCI was developed for IaaS. Now its newest release is also compatible with the two other cloud service models, PaaS and SaaS. In order to be modular and extensible the current OCCI specification is delivered as a document series. The documents are divided into three categories consisting of the OCCI core, the OCCI renderings and the OCCI extensions. Each of the documents focuses on a certain topic and can be independently used from the others. The work done within the OCCI group has gained a significant impact in the Cloud community. Several implementations have been released and are available for use. OpenNebula provided the first reference implementations of the OCCI specifications. rOCCI is an implementation of the OGF OCCI which has been successfully mapped and implemented upon the Amazon
4.3 QoS Control for Cloud Service Interaction

3GPP introduced the Policy and Charging Control (PCC) architecture that integrates flow based charging, QoS and policy control since release 7. As shown in Figure 5, the PCC architecture contains three main elements: the Application Function (AF), the Policy and Charging Rules Function (PCRF) and the Policy and Charging Enforcement Function (PCEF) (3GPP TS 23.203 v11.6.0, 2012). The AF maps the application specific information into the appropriate AVPs and sends this information to the PCRF via the Rx interface. In IMS, the AF is performed by the P-CSCF. The PCRF authorizes the session, performs the policy and QoS control and flow based charging, and communicates these decisions to the PCEF. The PCEF encompasses service data flow detection, policy enforcement and flow based charging functions, which is located at the gateway, e.g. GGSN in 3G networks. The Subscription Profile Repository (SPR) is the database storing the subscriber information and provides this information to the PCRF including subscriber’s allowed services, allowed QoS, charging related information, etc.

Although IMS works at the application signaling level, thanks to the PCC architecture, IMS achieves negotiating the transport level parameters with the network operator infrastructure. The P-CSCF generates the service information from the application specific media description. The most notable media description protocol is Session Description Protocol (SDP) which is responsible for the description of media and session abilities. The four critical components for mapping an SDP description into a QoS authorization are the media announcements ("m="), the connection data ("c="), the attributes ("a=") and the bandwidth ("b=").

In order to establish the transport channel for cloud interaction, the SDP body in signaling process of SIP session for the cloud interaction needs to be extended to support the media negotiation of new media types and formats. SDP is extensible easily to support new media types, formats and attributes, and has so far defined audio, video, text, application, and message media types. In addition, to provide negotiable and differential QoS for packet transmission of cloud services, the SDP body in the signaling process should also carry the related QoS requirements determined by the cloud interaction so that sufficient network resources can be reserved in advance.

As stated above, cloud management interactions with various cloud platforms adopt the uniform standardized cloud interface protocol such as OCCI. Cloud management interaction has the following transmission features: high reliability, low data rate, insensitivity to delay, interactivity, etc. This paper gives an example of SDP extension for the cloud management interaction using OCCI as the standardized cloud interface protocol. As described in the OCCI HTTP Rendering document, the query interface must be implemented by all OCCI implementations, which is located in the root of the OCCI implementation. And three content types (including text/ccsi, text/plain and text/uri-list) are specified to represent the data being transferred. Correspondingly, we extend two new media level attributes, “location” and “accept-types” in SDP. The location attribute indicates the path of the root of the OCCI implementation. The accept-types attribute indicates the content type of the data. Meanwhile, the b attribute is used to determine the maximum authorized data rate. And the Connection-Oriented Media (COMEDIA) specification is used to set up and maintain reliable connections as part of the negotiation mechanism. An example of part of a SDP body offered by Cloud AS SCF in success response to INVITE message is shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1. An Example of Part of SDP Body in 200 OK Response to INVITE Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>c=IN IP6 cloud.example.com</td>
</tr>
<tr>
<td>m=application 8080 TCP/HTTP *</td>
</tr>
<tr>
<td>b=TIAS:32</td>
</tr>
<tr>
<td>a=accept-types/plain text/ccsi/text/plain text/uri-list</td>
</tr>
<tr>
<td>a=location:<a href="http://cloudas.example.com:8080/users">http://cloudas.example.com:8080/users</a></td>
</tr>
<tr>
<td>a=setup:active</td>
</tr>
<tr>
<td>a=connection:new</td>
</tr>
</tbody>
</table>

The different types of cloud service interactions, such as file transfer and online movies, require different network...
status parameters including transmission rate, delay, error rate, etc. It is seemingly impossible to formulate a unified SDP description format for all types of cloud service interactions. But from another point of view, each type of cloud service interaction can be classified into one of media types (conversational voice/video, interactive gaming, TCP-based, etc) defined by the 3GPP PCC architecture according to the transmission requirements of resource type (guaranteed bitrate or Non-guaranteed bitrate), priority, packet delay budget and packet error loss rate. And in PCC architecture, the PCRF maps the service information received from P-CSCF into common IP QoS parameters (e.g. the QoS Class Identifier, the data rates) for all types of media sessions, including audio/video conversation as well as typically TCP-based services and applications. Therefore, SDP only needs to be extended to carry related information for the PCRF to be able to identify the appropriate QoS class identifier, the data rates and other IP QoS parameters. The UE can retrieve necessary network parameters from cloud service information received from Cloud SNF AS and take them as the QoS parameters of the corresponding cloud service.

QoS requirements may change in the course of cloud interaction. In SIP, the UPDATE message allows a client to update parameters of a session (such as the set of media streams and their codecs) but has no impact on the state of a dialog. Compared with re-INVITE, the UPDATE message needs to be answered immediately, ruling out the possibility of user approval. This paper tends to adopt UPDATE message to carry new QoS description information during the session. When receiving updated session description information, the P-CSCF sends an update for service information to the PCRF. The PCRF decides if a new QoS authorization is needed and updates the authorization for the session accordingly. The P-CSCF optionally provides the PCRF with the information of cloud service provider to enable specific PCC decisions.

### 4.4 Charging for Cloud Service Usage

IMS provides a well defined architecture for charging multimedia calls and services. The IMS network elements interact with the charging platform through interfaces using DIAMETER protocol. The IMS charging platform is subdivided into two major components: the Charging Data Function (CDF) and the Online Charging System (OCS). The CDF is responsible for receiving offline charging triggers, while the OCS is responsible for receiving online charging triggers. The Charging Detail Records (CDRs) are collected and co-related at the Charging Gateway Function (CGF) which acts as a gateway to the Billing System.

According to the requirements of IMS charging system, Cloud SCF ASs and Cloud ICF ISs need to implement a Charging Trigger Function (CTF) and act as the charging client. The CTF is responsible for monitoring service usage and generating charging events based on it. The charging architecture of cloud services integrated in IMS is shown in Figure 6. The charging system triggers the charging process and sends charging requests to the offline or online charging functions. The Cloud SCF ASs and Cloud ICF ISs are able to distinguish whether to send charging information over the Rf interface to the CDF or over the Ro interface to the OCS based on the information received in the SIP signaling and the system configuration provisioned by the operator.

Cloud clients manage their own cloud services through the Cloud MIPF IS. Some of these management operations may trigger charging events. For example, in Amazon EC2, clients use the ‘runinstance’ command to launch a virtual machine instance and use the ‘stopinstance’ to stop it. Pricing is per instance-hour consumed for each on-demand instance from the time an instance is launched and until the instance is terminated. So, the two API actions above correspond respectively to the start and end of charging for a cloud service usage. Therefore, in addition to reporting accounting information upon reception of the general session-related SIP messages (INVITE/200OK/ACK, UPDATE, BYE, etc), Cloud SCF ASs and Cloud ICF ISs also need report specific accounting events for cloud service usages.

However, the cloud management interaction between cloud client and Cloud MIPF IS is along the media plane channel, which does not traverse SIP signaling intermediates. The IMS network entities (P/I/S-CSCF, Cloud AS, etc.) involved in the SIP session for cloud management interaction cannot perceive these charging points. Therefore, it is necessary to reflect the management operations involving charging into SIP signaling level. After the successful execution of the
management operations involving charging, cloud client will notify the management operation information to Cloud SCF AS and Cloud ICF IS by sending a SIP message. In addition, the charging of cloud service usage does not involve the charging correlation between the visited network and home network. So, all IMS nodes except for Cloud SCF ASs and Cloud ICF ISs need not report the charging information of cloud service usage.

As an extended SIP method, the INFO message is used to carry application level information between communication entities using the SIP dialog path. And the Info Package mechanism allows an INFO message carrying information with different contents and semantics. Although INFO method is not among the SIP methods which can trigger accounting request messages, IMS charging allows operators to use their own configured rules locally on the ASs (3GPP TS 32.260 v11.4.0, 2012). So, we adopt INFO message to carry the management operation information. As shown in Figure 7, cloud client sends an INFO message to Cloud SCF AS and Cloud ICF IS after successfully executing one cloud interface operation involving charging.

A new Info Package should be extended to support the function mentioned. The management operation information carried by INFO messages should include all of the parameters which can affect charging modes or charging rates, such as service type, resource type, vendor and action. In addition, some proprietary parameters about specific cloud service also need be included. For example, the type of virtual machines (small instance, medium instance, large instance, etc) should be included for Amazon EC2. After receiving an INFO message, Cloud SCF AS and Cloud ICF IS exact and send accounting information to CDF or OCS based on the configured rules by operators. Other SIP network entities involved only simply forward the INFO message.

Apart from the SIP session establishment, the number of SIP messages exchanged during the lifetime of a normal SIP session is rather small. In our scenario, the rate of INFO messages is low and the size of INFO messages is relatively

Figure 7. Message Sequence Chart for Cloud Service Usage

http://www.hipore.com/ijcc
5. IMPLEMENTATION OF IMS-BASED CLOUD COMPUTING PROTOTYPE AND PERFORMANCE ANALYSIS

5.1 IMS-BASED CLOUD COMPUTING PROTOTYPE

Figure 8 shows our prototypical implementation of the IMS-based cloud computing platform. OpenIMSCore of Fraunhofer FOKUS was used as the core IMS component. The OpenIMSCore is an open source implementation of IMS CSCFs, a lightweight HSS and Application Servers (ASs). Each core IMS component, including P-CSCF, I-CSCF, S-CSCF and HSS, runs on a Linux Debian server. The Cloud SCF AS was also implemented based on OpenIMSCore. The Cloud SNF AS has been developed based on OpenSIPS project. Several models were extended to support new-defined cloud event packet. The Cloud PCF AS has been developed based on OpenXCAP, which is an open source XCAP server. XCAP protocol allows a client to read, write, and modify application configuration data stored in XML format on a server using HTTP protocol.

The cloud client, consisting of the IMS signaling part and cloud service interaction part, has been developed based on Doubango project. Doubango is an open source 3GPP IMS client for both embedded and desktop systems. Boghe, one client-side component of Doubango can be used as IMS client for Windows XP. Doubango was extended to support internet explorer to allow the IMS client to interact with cloud platform. The UE refers to the end devices running Windows XP or Windows Mobile systems, as well as IMS-based cloud client software.

A SaaS cloud platform was built using Feng Office, which is an open source Web Office platform developed and supported by Feng Office community. It is a fully featured online office suite with a similar set of features as other online office suites, like Google Apps. Based on it, cloud interaction control function which interact with cloud platform, and SIP-based y2 interface additionally were implemented. Another SaaS cloud platform was built using OpenStack and Asterisk. Openstack is a ubiquitous open source cloud computing platform for public and private clouds. It is simple to implement and massively scalable, and aims to deliver solutions for all types of clouds. Asterisk is a free and open source framework for building communication applications. It includes all the building blocks needed to create communication applications, such as PBX, IVR and conference system. Asterisk was installed on top of the IaaS cloud platform which was built using OpenStack.

5.2 PROCESSING STEPS OF PERFORMING ONE CLOUD SERVICE

Figure 9 shows a step-by-step procedure of the service discovery, service management, session establishment, and service interaction performed in a typical use scenario. When a UE starts up, it performs network attachment through the Network Attachment Subsystem (NASS) to obtain network parameters (such as an IP address, P-CSCF address, etc). Then, the UE initiates registration process with the IMS core. The registered IMS UE is then able to initiate the service discovery process by subscribing to a public SIP identifier, which corresponds to the Cloud SNF AS. Through service discovery process, the registered IMS users get the information about all of the currently available cloud services and subscribed cloud services from Cloud SNF AS. From the obtained cloud service information, the IMS UE can retrieve a list of SIP URIs or IP addresses of the Cloud PCF AS and the Cloud SCF AS. The IMS user is then able to perform the service management process by configuring user profile parameters through the Cloud PCF AS.

After selecting the desired cloud service, the IMS UE first establish a service control session by exchanging the extended SIP signaling with Cloud SCF AS before the
actual interaction with the cloud platform. The SIP URI of Cloud SCF AS is obtained from the service discovery process. The new session will be established as a usual IMS session via the Gm between the UE and the core IMS. The core IMS can initiate a resource reservation process for network resources that are required by the incoming cloud interaction process. The reservation is performed using standardized functions of NGN RACS and NASS subsystems connected to the core IMS. After the successful session initiation, the Cloud SCF AS informs the cloud interaction platform (containing interaction control and process functions) via core IMS using the y2 interface to initiate cloud interaction procedure. After the establishment of the initial data stream, the UE communicates with the cloud platform over the Xd interface for actual cloud interaction.

5.3 PERFORMANCE ANALYSIS

One of IMS’s main functions is to provide QoS control function for services. IMS explicitly defines QoS parameters for negotiating between two UEs before establishing the session. The purpose is to ensure sufficient available resources between two UEs.

In this section, we conducted some experiments to verify QoS policy control ability of IMS. We extended SDP protocol to support new QoS parameters for cloud services, which is transmitted between communication parties through SIP messages. We defined three QoS policies including application type, codec type and QoS level rules. One open source XDMS server was used as the policy repository to store QoS policy files. The PCRF of IMS QoS framework can obtain policy files through XCAP protocol.

In the experiments, the cloud clients can access three cloud service types including online office, audio conference and web browsing. By setting application type rule, we artificially set priorities for these three cloud service types, audio conference service with the highest priority and web browsing with the lowest priority. The experiment system manually limit ed the maximum resource capacity and set a threshold for resource usage. When system resource usage is below the threshold, the PCRF would authorize all types of the authenticated cloud services.
When system resource usage is over the threshold, the PCRF would selectively authorize part of the authenticated cloud service sessions with higher priority and allocate resources according to service priority.

Figure 10 shows the system throughput of different cloud services including online office, audio conference and web browsing. In the early stages of experiment, all types of cloud service requests are authorized and obtain sufficient resources. And system throughputs of three cloud service types steadily grow in this stage. With the increase of system resource utilization, part of the web browsing service and online office service are suppressed and sacrificed to ensure the audio conference service priority over other services. In the later stage, audio conference throughput is growing. Web browsing throughput is reduced. And online office throughput is growing slowly.

6. CONCLUSIONS

This paper presents a framework for the integration between IMS and Cloud Computing based on open standards. Various cloud services could be offered by different vendors and regarded as the general IMS applications. IMS terminals only need a few modifications to access various cloud services with negotiable QoS over Internet and widely available wireless networks. SIP protocol is adopted to establish and manage the transport channel for cloud interaction between cloud clients and cloud platforms. A new event package is presented to perform the cloud notification service through which cloud clients know the information of all cloud services provided by IMS system. The uniform standardized cloud interface protocol, such as OCCI, is adopted to provide interoperability between cloud platforms. QoS and charging control of cloud computing services integrated in IMS are also discussed in detail on the basis of the PCC architecture of IMS. In the future, many potential challenging issues still need to be researched, such as scalability issues of IMS architecture integrated with cloud computing. It is expected that the work in this paper will trigger more research efforts in this emerging area.

7. REFERENCES


3GPP TS 23.228 v11.9.0 (2013, September), “IP Multimedia Subsystem (IMS)”.


Authors

Wei Zhang is a lecturer of Institute of Communication and Information System, College of Information Science and Engineering, Northeastern University, Shenyang, China. She received her B.S. degree in computer science and technology from Ocean University of China in 2002, and M.S. degree from Institute of Computing Technology, Chinese Academy of Sciences in 2005. She is currently working toward Ph.D. degree in Institute of Communication and Information System, Northeastern University, Shenyang, China. Her research interests include multimedia communication, protocols and services in next generation network, and network architecture of cloud computing. She has published more than five research papers in IEEE journals and international conferences. She is also the inventor of three Chinese patents.

Weimin Lei (M’05) is a full professor in Institute of Communication and Information System, College of Information Science and Engineering, Northeastern University, Shenyang, China. He received his B.S. degree from Nankai University, Tianjin, China, in 1992, M.S. degree from Institute of Computing Technology, Chinese Academy of Sciences in 1995, and Ph.D. degree from Dalian University of Technology, China, in 1995.
His research interests include: protocols and services in IP multimedia system, multi-path transmission, overlay network and ad-hoc network. He has published more than forty research papers in IEEE journals and international conferences. He is the inventor of several Chinese patents and pending applications.

Xiao Chen received his B.S. degree and M.S degree in computer science and technology from Northeastern University, Shenyang, China, in 2004 and 2006 respectively. He is currently working toward Ph.D. degree in Institute of Communication and Information System, Northeastern University, Shenyang, China. He is currently a lecturer in College of Software, Shenyang Normal University, Shenyang, China. His research interests include protocols and services in next generation network, and multi-path transmission in overlay network. He is an author of over five research papers in IEEE journals and international conferences and has applied two Chinese patents.

Shaowei Liu received his B.S. degree in communication engineering from North University of China, Shanxi, China, in 2010, and M.S degree in communication and information System from Northeastern University, Shenyang, China, in 2012. He is currently working toward Ph.D. degree in Institute of Communication and Information System, Northeastern University, Shenyang, China. His research interest is IP multimedia communication and system.