Literature Survey: Investigation of billing principles and infrastructures for next generation services

Moses Thizwilondi Nkhumeleni
Supervisors: Professor Alfredo Terzoli and Mr Mosiuoa Tsietsi
Department of Computer Science, Rhodes University

10 May 2010

Abstract

This paper highlights challenges posed to billing and a variety of billing concepts around next generation platforms. Pricing strategies for services and different business models are also discussed. The paper considers the environment for next generation applications by discussing potential tools that can be used to develop and test billing systems such as the mobicents platform and the Ericsson Charging SDK. The offline and online charging architecture is explored in detail. The paper concludes that billing for next generation application will be more complex and there will be a number of factors to consider when developing billing systems.

1 Introduction

The convergence of voice, video and data allows developers to deliver rich services to the user. These services add additional strain to the current billing approaches. Factors such as variety of services, service composition, and the quality of service offered are driving the search for innovative billing approaches. Billing across telecommunication includes a number of concepts such as online and offline billing, prepaid and post-paid accounts. The charging architecture (online charging architecture and offline charging architecture) is important as it affects the manner in which charging can be conducted in applications. The Mobicents platform is an open source platform that allows developers to develop next generation applications. The Mobicents platform contains a number of wrappers known as resource adapters. An important resource adapter for billing is the Diameter resource adapter which implements the Diameter Protocol. The Diameter Protocol is a fundamental protocol for billing as it performs Authentication, Authorisation, and Accounting [6]. The paper starts by exploring factors that affect billing and concepts around billing. We proceed to discuss pricing strategies and charging models. Examples to implement important concepts are discussed, followed by a discussion around the environment or platforms that can be used to develop billing systems. The last section shows a contrast between the online and offline charging architectures.

2 Factors affecting billing

With the introduction of third generation (3G) networks, the main idea has been to merge the cellular networks and IP networks. Third generation networks operate on packet switched domain and therefore allow the users to access the internet via IP protocol. However, the IP protocol offers best effort delivery. There is no bandwidth guarantee that is offered with data transfer. This means that the services are available without QoS (Quality of Service) guarantee. This presents additional concerns when offering real-time multimedia services. IMS (IP Multimedia Subsystem) allows for synchronisation of service establishment with QoS and hence this allows the users to predict the quality of service [7]. In this section we summarise factors affecting billing and charging in 3G networks.

2.1 Variety of services

IMS promises to deliver a number of different services to the users such as voice, data, video, conferencing and e-commerce services. These newly developed services impose a challenge to the operators current charging model. Traditional services used a time-based charging model. Traditional time-based charging models are unsatisfactory and hence operators need to investigate new ways of billing for their services.

2.2 Service composition

With the introduction of IMS, service composition has become attractive in order to deliver rich applications that combine different media such as voice, data and video. There are a number of additional complications that are associated with services composition. Namely, an application could make use of different application servers and third parties maybe involved in delivering of services to the user [19].

These additional complexities translate to a more complex charging system. For example if we have an application that consumes n different services from different services provider as shown in Figure 1, each service is associated with a certain quality of service and charge price. The overall quality of service will be dependent on the QoS of different services composed in the bundle. Additional complications are added to the charging model since all services are not being used at the same time. It can be seen that during the applications execution the service bundles will change and the QoS will vary. Services in an application will have different level of importance and therefore, an application can also execute without some services [19]. If we have a charging approach that depends on

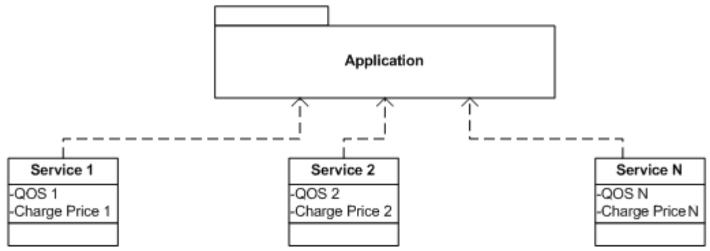

Figure 1: Application composed of different services, adapted from [19]

QoS, charging will vary with the change in time due to the factors mentioned above.

2.3 Quality of Service

Quality of service (QoS) can be defined as the quality of transmission in a network that the user receives based on various factors such as transmission rates and packet loss. When providing services across IMS, the issue of QoS becomes apparent. The quality of service should therefore be considered in the charging model. The users should be able to select the quality of service that they prefer. It is important that different QoS should allow for different rates. Services differentiation is the ability to treat different classes of traffic differently based on their QoS [10].

2.4 Different service providers

IMS allows developers to combine a number of different services to create a single application. These services will be consumed from different service providers. For example an application can use weather services from one service provider and news services from another provider. This poses challenges to the charging model, since the service providers will require their share of the amount that the user is billed for.

2.5 Flexible charging

In order for network providers to remain competitive their pricing strategies will need to be revised. The new charging scheme will be required to be flexible and cater for QoS. Flexibility means that that pricing strategies should be able to change based on the context of the user. For instance different locations would results in user being charged with a different price. Quality of service and having flexible charging of services are fundamental with the introduction of IMS [4].

2.6 User preferences

Traditionally, network operators have proposed a billing model whereby a tariff is set and rolled out to all the customers. This means that all the customers were charged in the same manner and there was no personalised pricing strategy for a customer. The traditional billing strategy did not provide a charging approach which catered for the customers preferences. The billing strategy did not differentiate among the users. Future billing models will need to consider users preference when it comes to billing for different services [3]. So each user should be billed differently based on the users preferences. This means that network operators will need to reverse the traditional billing approach and therefore, consider users preferences or different parameters associated with the user. For example a user can have a tailored billing scheme based on their preferences.

3 Billing concepts

This section defines and highlights major concepts in billing. This section is also used to introduce the terminology that will be used throughout the document.

3.1 Offline and online Charging

In offline charging, billing information is sent to an external billing system. The function of this external billing system is not part of the 3rd Generation Partnership Project (3GPP) standard nor is the transferring of data from the network to a billing system [1]. Offline charging is conducted after resources are consumed. For example, if a user initiates a call; the user would be billed after the call. This means that the charging information gathered does not affect the actual service. Given the above motivation offline charging would be suitable for subscription-based charging model.

With online charging, the user account is checked to see if there is enough credit for the user to consume the service. To clarify this point further, online can be thought of as an approach that has real time interaction with the service that is being offered. Online charging is done making use of an online charging system that conforms to the 3GPP standard [1]. Direct debiting and Unit Reservation are two scenarios for online charging. Direct debiting the account is debited immediately after a charging event occurs. With unit reservation the users has reserved units on the account. When the user completes using the resources the remaining credits are returned to the users account.

The charging architecture will change based on whether we perform online or offline charging. The two charging architectures are described in the last section. In both online and offline charging additional charging approaches are highlighted.

3.1.1 Event-based charging

With event-based charging an event is raised with a single transaction performed by a user. For example, the user sending a multimedia message. Event-based charging results in a creation of a single CDR (Charging Data Record) [1]. CDRs are collection of charging information about an event.

3.1.2 session-based charging

A characteristic of session-based charging is the establishment of a user session, such as a phone call. Session-based charging results in multiple charging events that could result in multiple CDRs [1].

3.2 Account types

There are two primary types of accounts that are used by users, namely pre-paid and post-paid accounts. A contractual agreement is required between the network operator and the customer for a post-paid account. With a post-paid account the user receives a bill after a certain period listing the charges that have been incurred. The user can therefore use the service as much as they want. This allows for a simple billing system to be implemented. Prepaid account requires no contract; however before using services offered by the network provider the user must have credits in the account. A trend has shown that there has been an increase in the number of prepaid users [14]. The major reason for this is that pre-paid account require less administration compared to post-paid account that usually requires a contract to be signed between the user and the service provider. Network operators offering prepaid to customers have experienced huge growth with the number of customers and the revenue [12]. Prepaid function is based on alarms, whereby the user is able to use services up until a certain threshold [14]. For instance the number of credits the user has would translate to the number of seconds that can be used. If the numbers of credits run-out; an alarm is raised.

3.3 Business Models

A business model is defined as the relationship between a business and the service it provides [5]. There are three main business models in the telecommunication industry, namely the content providers, service providers and network operators. Content providers focus mainly on delivering downloadable content such as music. Service providers deliver services to the users, and network operators provide access to the network. Businesses may combine two or more of these business models. Network operators often take the role of service providers or content providers [5]. With the convergence of telephony and IP networks, in order for network operators to remain competitive, they will be required to move towards service delivery models [22].

3.4 Different charges

There are number of different charges that telecommunication primarily focus on, mainly service charges, tariff charges, and content charges. Service charges are for using particular services. Tariff charges are for accessing the network, and finally content charges are charges based on the content downloaded by the customers.

3.5 Real time charging

In real-time billing; information about billing is generated, processed and transported in less than 1 second [1]. With real-time billing the service offered is valued throughout the service, therefore in real-time charging information, changes based on the service that the user receives in a session.

4 Pricing Strategies

There are a number of pricing strategies available. Pricing strategies are used by service providers to bill the users for the service provided or resources consumed. [20] identifies a number of these pricing strategies namely: Flat-rate pricing, Paris-metro pricing, Responsive pricing, and Priority pricing.

4.1 Flat-rate pricing

Flat-rate pricing strategy is based on a fixed subscription amount the customer pays for a service. Flat-rate allows the user to use the services or resources over an agreed period of time. In telecommunication networks flat-rate pricing would allow users to use resources in an unlimited manner; however it is the simplest pricing strategy to implement [20].

4.2 Responsive pricing

Responsive pricing is primarily based on the network traffic. If there is high network traffic the price is adjusted to a higher price and when there is low network traffic the price is reduced. This pricing strategy is effective at controlling the amount of traffic on the network, since the price sensitive users will delay their calls to times when the price is low. Responsive pricing strategy requires more network resources due to signalling information [20].

4.3 Paris-metro pricing

In Paris Metro Pricing, the network is partitioned into separate channels each with a different price [18]. Even though there is no bandwidth guarantee and the network still offers best effort delivery, it is expected that the channel with higher price will have less traffic and channels with lower prices will have more traffic. Similar to Responsive pricing there is a trade off between low quality of service and the price that the users pays for the service.

4.4 Expected capacity pricing

With expected capacity pricing, the prices vary with the amount of capacity that is expected on the network. The user is not limited to consume more resources when the network is congested but does so at higher price [8]. This pricing creates incentives for user to consume resources when the network is not congested by reducing the price. Additionally this pricing scheme directly translates to the cost that the provider incurs.

4.5 Priority pricing

This pricing strategy focuses on the priority that the user receives on the network structure. A higher priority would translate to higher price that the user will be charged. Priority pricing can be used to create QoS profiles since it relates directly to traffic control [20].

5 Charging Models

Charging models are used by the service providers in order to recover costs and to generate revenue. It is imperative to have a suitable charging model for the service or content provided. Users will select a services based on the charging model used.

5.1 Subscription-based

With the subscription-based charging model, the users pay a fixed amount over a period of time. Subscription-based model would encourage more of resource usage since the customer pays a fixed amount which is independent of the amount of resources used. Subscription-based model is most popular in charging for internet access, the users pay a fixed amount based on the access channel they prefer [8]. The main advantage with subscription-based model is that users can predict their charges.

5.2 Event-based

Event-based charging is based on chargeable events that occur. A chargeable event is a single transaction, e.g. sending an MMS. Furthermore, event based charging uses a single CDR [1].

Application	*QoS*	*Weight*	*Price*
app1	a	1	x
app2	b	2	y
...	...	...	...
appN	n	n	z

Table 1: Diagram shows a table containing charging information

5.3 Volume-based

In the volume-based charging model, the user is billed according to the amount of data transferred. The advantage of this model is that the users are only charged for the amount of content that they have received or sent.

5.4 Time-based

Time-base charging is based on the units of time the user consumes. This is a popular traditional billing approach used by telecommunication companies. With the introduction of IMS this charging model will be unable to cater for the variety of services that users will be able to access.

5.5 Reward-based

Reward-based charging model is based on the usage criteria of the user. This model is useful for creating brand loyalty. For example the more the user consumes a services, the price is reduced as they continue to consume that particular service.

6 Examples

In this section we discuss two proposed solutions to billing for services in third generations networks.

6.1 Multi-Definition of services in IMS

In order to overcome the complications associated with service composition as described above, [19] proposes a multi-definition of a service. A table containing QoS, weight, prices of each service is stored in a charging module as shown in Table 1. The table is updated in real time. When charging for an application, the table is queried in order to give a charging amount.

6.2 Call differentiation Approach

In order to address issues with billing for services with QoS, [4] proposed a system which charges users differently based on the QoS of the call. Furthermore, the user could subscribe to one of the following classes: silver, gold, or platinum.

These classes represent the Quality of service that the user has opted. Higher QoS translates to a higher cost. Additionally, the system cateres for emergency calls. Emergency calls do not have peremption because of the critical nature of the calls [4].

7 Environment

In this section we outline the tools or platform that will be used in the research project. The proposed system is based on the Mobicents platform with a strong focus on the Diameter protocol. The Ericsson charging SDK is used as an emulator and primarily focuses as a test bed for the proposed system.

7.1 Mobicents

The Mobicents platform has been proposed as a development platform for next generation applications. Mobicents is an open source platform that allows developers to create applications that combine voice, video and data. Therefore Mobicents facilitates this type of convergence when developing next generation applications. The Mobicents platform is certified as JAIN SLEE compliant. SLEE(Service Logic Execution Environment) is a popular standard in the telecommunications industry. SLEE is similar to EJB(Enteprise Java Beans) but has different components [9]. JAIN SLEE is java standard for SLEE. Mobicents JAIN SLEE is an open source platform and built on top of the JBOSS application server as seen from Figure 2.

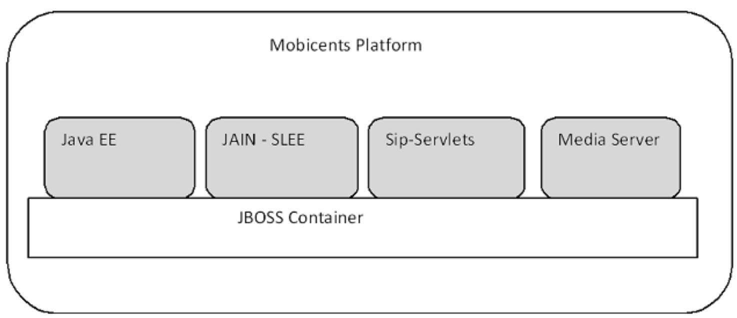

Figure 2: Diagram shows the Mobicents plaform, adapted from [9]

Mobicents JAIN SLEE compliments J2EE(Java enterprise) container features [9]. Mobicents facilitates the composition of different SBBs(Services Building Blocks). The different SBBs available are: billing, call control, user-provisioning, administration and presence. The use of resource adapters allow developers not to be concerned with the underlying network communications. Resource adapters act as wrappers that allow SLEE applications to communicate with external resources using standard protocols [15]. Of particular interest to billing is the Mobicents Diameter Resource Adapter. The Diameter Resource Adapter provides a development API for applications that implement billing. The Diameter Resource Adapter is based on the Diameter protocol, which is described in the next section.

7.2 Diameter Protocol

The Diameter Protocol has been created to perform AAA (Authentication, Authorisation and Accounting). Authentication is verifying the claimed identity of a subject, whilst authorization is focused primarily on access that the subject has. Accounting is focused on collecting information with regards to the amount of resources being used. Therefore accounting allows for billing mechanisms to be implemented. Authentication, Authorisation and Accounting are linked and hence in order to account for the usage of resources all tasks must be performed [7]. IETF proposed Diameter as a replacement of former RADIUS protocol. RADIUS runs on unreliable UDP protocol. Diameter runs on TCP which offers reliability and congestion control. There are a number of entities defined by diameter such as Diameter Client, Diameter server, realm, etc [7]. Diameter client is a located at the edge of the network and mainly performs access control and Diameter server is primarily used for authentication, authorisation, and accounting [7].

Diameter is a peer to peer protocol, therefore peers can send requests to other peers [7]. Diameter messages contain header information and a number of AVPs(Attribute value pairs). AVPs can be considered as containers of data. AVPs can be used to transport authentication information, in order for the user to be authenticated. AVPs can also be used to transport resource usage information which is required for accounting. The Diameter protocol can also be extended by defining new AVPs values, or by creating new AVPs [6]. Diameter makes use of either request or answer commands. There are a number of request and answer commands used by diameter, e.g. ACR (Accounting Request) and ACA (Accounting answer) [6].

7.3 Charging SDK

Ericsson have released as charging SDK which developers can use to quickly develop applications that have charging incorporated into them. The Diameter SDK allows developers to modify account balances based on different charging models such as time, event, and volume. The major advantage of the Charging SDK is that individuals can develop and test applications without access to the live system. The Diameter Charging SDK consists of the Diameter Charging API, Charging Emulator, and the Charging Client.

7.3.1 The Diameter Charging API

The Diameter Charging API provides an interface that allows the developer to abstract away from the underlying details of the Diameter protocol such as the

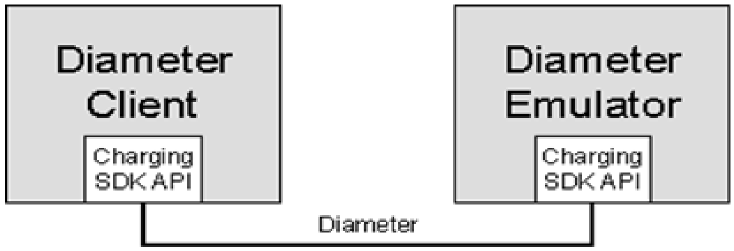

Figure 3: The diameter Client and Emulator, adapted from [17]

transportation of messages. This Charging API is included or imported into the clients application that the user is developing.

7.3.2 The Diameter Client

The client is a test application and is used mainly to set up connections to the server. The Diameter Client also consists of the Charging API as shown in Figure 3. The provided client can be used as an example of how a client can be developed. The Diameter Client allows the user to setup a number of configurations such as the client configurations (Peer ID, Realm, Vender ID, and IP Address). The server configurations that are required to connect to the server (PeerID, Realm) can also be configured. Additionally the client allows the users to observe activities that are currently taking place. The client also displays information about the number of sent requests, and the number of received answers and the number of failed requests.

7.3.3 The Diameter Emulator

The emulator acts as a server application. The emulator emulates a potential prepaid system. Therefore the emulator would respond to the requests from the client. The emulator allows the developers to test the system that they have developed. The emulator contains a database with all the accounts, currency, and tariff records [17]. The emulation has an option menu that allows users to set log files and additional stack options. The user can configure the currency. The users can create and edit account information such as balance, currency and the account holders name. Tariffs can also be defined as to how the transactions will be charged. Similar to the client, the emulator also has a number of counters such as received and sent requestions. Unique to the emulator are the denied and faulty requests.

8 Charging Architecture

Depending on whether we are implementing an online or offline billing system, the charging architecture will change accordingly. This section describes online charging and offline charging architectures.

8.1 Offline charging Architecture

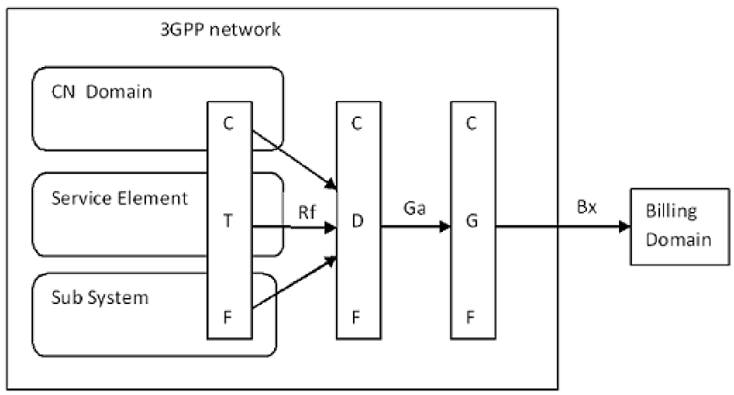

Figure 4: The offline charging Architecture, adapted from [1]

8.1.1 Charging Trigger Function

The charging trigger function(CTF) will create charging events based on the network resource usage [1] . The CTF gathers information about charging and packages them into charging events, which are sent to the next function (Charging Data Function). The CTF consists of functions:

- Accounting Metrics Collection
  This functions simply monitors call signals, events, or sessions that users establish. This function would produce information that identifies the user and the amount of resources consumed.
- Accounting Data Forwarding
  This function receives the information and attempts to identify chargeable events from the received information. The events are assembled and forwarded to the Charging Data Function(CDF)

8.1.2 Charging Data Function

This function will receive charging events from CTF via Rf reference point as shown in the diagram. These events are used to create CDRs (Charging Data Record). CDR will contain information such as time of call, duration of call, amount of data transferred, etc.

8.1.3 Charging Gateway Function

The CDRs produced in the Charging Data Function (CDF) are transferred to the Charging Gateway function (CGF). The CGF therefore links the 3GPP network to the Billing Domain. Remembering that the Billing Domain is not part of the internal 3GPP network structure. The CGF will receive CDR, validates and format the CDR and transfer CDR to the billing domain.

8.2 Online charging architecture

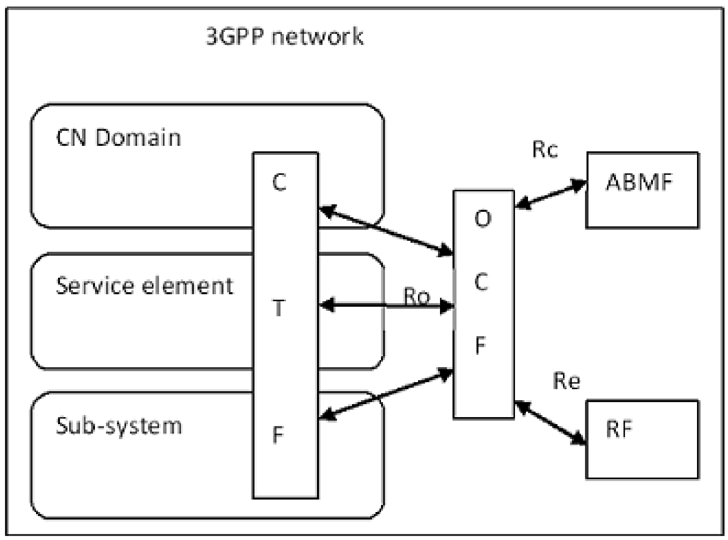

Figure 5: The online charging Architecture, adapted from [1]

8.2.1 Charging Trigger Function

This process is similar to CTF defined for offline charging, however information gathered is not identical to offline charging model. Additionally since online charging requires permission to be granted to resources before use, the CTF must be able to delay the use of resource until permissions are granted [1]. Another difference with offline charging is that there is bidirectional communication between CTF and OCF. The number of messages sent between the CTF and OCF means there is more overhead as a result of charging [21].

8.2.2 Online Charging Function

OCF contains two functions namely: Session Based Charging Function(SBCF) and Event Based Charging Function(EBCF). SBCF conductions online charging for user sessions such as voice calls and IMS sessions [1]. EBCF is responsible for event-based online charging.

8.2.3 Rating Function

Rating Function(RF) is used to determine the value of the resource using the information in the charging event. The RF output therefore is monetary or non monetary units which are sent back to the OCF.

8.2.4 Account Balance Management Function

This is a storage area containing the users account balance.

9 Conclusion

Billing for next generation applications will be more complex. A number of factors such as quality of service and service composition are considered key to developing next generation billing systems. The choice between online and offline billing will affect the architecture that is used when considering charging into systems. Pricing of services also has a critical role to play in delivering new marketable services. Traditional time-based charging models can not cater for the variety of services being presented to the user. Having emulation systems such as the Ericsson charging SDK will allow developers to quickly test their applications.

References

[1] 3GPP TS 32.240. *Charging Managment, Charging Architecture and Principles*. 3GPP Organizational Partners, 2009.

[2] Alexandre, Mendona, Bartosz, Baranowski, Morgan, Tom, and Wells. The guide to the mobicents diameter adapter, 2009. Available from http://hudson.jboss.org/hudson/job/MobicentsBooks/lastSuccessfulBuild/artifact/diameter/index.html; accessed 18 February 2010.

[3] Alonistioti, A., Houssos, N., Panagiotakis, S., Koutsopoulou, M., and Gazis, V. Intelligent architectures enabling flexible service provision and adaptability. In *Wireless Design Conference (WDC 2002)* (2002).

[4] Barachi, M., Glitho, R., and Dssouli, R. Charging for Multi-grade Services in the IP Multimedia Subsystem. In *Next Generation Mobile Applications, Services and Technologies, 2008. NGMAST'08. The Second International Conference on* (2008), pp. 10–17.

[5] Bormann, F., Flake, S., and Tacken, J. Business models for local mobile services enabled by convergent online charging. *Mobile and Wireless Communications Summit, 2007. 16th IST 16* (2007), pp. 1–5.

[6] Calhoun, P., Loughney, J., Guttman, E., Zorn, G., and Arkko, J. Diameter base protocol. In *IETF RFC 3588* (2003).

[7] Camarillo, G., and Garca-Martn, M. *The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds*, 3 ed. John Wiley & Sons, Ltd, 2008.

[8] Clark, D. A model for cost allocation and pricing in the Internet. In *MIT Workshop on Internet Economics* (1995), vol. 49, p. 13.

[9] Deruelle, J. Mobicents communications platform, 2008. Available from http://people.redhat.com/vralev/blog/JavaOne 2008 Presentation - Jean Deruelle.pdf; accessed 01 June 2010.

[10] El Barachi, M., Glitho, R., and Dssouli, R. Context-aware signaling for call differentiation in IMS-based 3G networks. In *12th IEEE Symposium on Computers and Communications, 2007. ISCC 2007* (2007), pp. 789–796.

[11] Ericsson. Charging sdk 2.0, 2008. Available from http://www.ericsson.com/developer/sub/open/technologies/charging _solutions/ tools/charging _sdk _2; accessed 18 February 2010.

[12] Hakala, H., Mattila, L., Koskinen, J., Stura, M., and Loughney, J. Diameter credit-control application. In *IETF RFC 4006* (August 2005).

[13] Kurtansky, P., and Stiller, B. State of the art prepaid charging for IP services. *Wired/Wireless Internet Communications 3970* (2006), 143–154.

[14] Kurtansky, P., Stiller, B., and Series, T. Prepaid charging for QoS-enabled IP services based on time intervals. In *First International Workshop, WINE* (2005), pp. 15–17.

[15] Mendona, A. Mobicents JAIN SLEE Diameter Base Resource Adaptor User Guide, 2010. Available from http://www.mobicents.org/slee/docs/resources/diameter-ro/2.0.0.GA/en-US/pdf/Mobicents _SLEE _RA _DIAMETER_RO _User _Guide.pdf; accessed 10 May 2010.

[16] Niemoller, J., Quinet, R., Levenshteyn, R., and Fikouras, I. Cost Control in Service Composition Environments. In *Proceedings of the 2008 The Second International Conference on Next Generation Mobile Applications, Services, and Technologies* (2008), IEEE Computer Society, pp. 25–32.

[17] Nilsson, M. Diameter Charging SDK Practicalities, 2008. Available from http://a178.g.akamai.net/7/178/8211/0001/ ericsson-com.download.akamai.com/8211 /high1/mainframe _low.htm; accessed 10 May 2010.

[18] Odlyzko, A. Paris metro pricing for the internet. In *Proceedings of the 1st ACM conference on Electronic commerce* (1998), pp. 140–147.

[19] Oumina, H., and Ranc, D. Towards a real time charging framework for complex applications in 3GPP IP Multimedia System (IMS) environment. In *Next Generation Mobile Applications, Services and Technologies, 2007. NGMAST'07. The 2007 International Conference on* (2007), pp. 145–150.

[20] Ozianyi, V., Ventura, N., and Golovins, E. A novel pricing approach to support QoS in 3G networks. *Computer Networks 52*, 7 (2008), 1433–1450.

[21] Ozianyi, V. G., and Ventura, N. Service outsourcing and billing in inter-domain ims scenarios. In *Southern African Telecommunication Networks and Applications Conference (SATNAC 2009)* (August 2009).

[22] Panagiotakis, S., Koutsopoulou, M., and Alonistioti, A. Business models and revenue streams in 3g market. In *14th IST Mobile and Wireless Communication Summit* (2005).

[23] SUBSYSTEM, I. Charging in the IP Multimedia Subsystem: A Tutorial. *IEEE Communications Magazine 45* (2007), 93.