Charging and Billing in Modern Communications Networks — A Comprehensive Survey of the State of the Art and Future Requirements

Ralph Kühne, George Huitema, and George Carle

***Abstract*—In mobile telecommunication networks the trend for an increasing heterogeneity of access networks, the convergence with fixed networks as well as with the Internet are apparent. The resulting future converged network with an expected wide variety of services and a possibly stiff competition between the different market participants as well as legal issues will bring about requirements for charging systems that demand for more flexibility, scalability and efficiency than is available in today's systems. This article surveys recent developments in charging and billing architectures comprising both standardisation work as well as research projects. The second main contribution of this article is a comparison of key features of these developments thus giving a list of essential charging and billing ingredients for tomorrow's communication and service environments.**

***Index Terms*—charging, billing, accounting, communication network, architectures, requirements**

I. Introduction

CHARGING and billing are essential aspects for any commercially successful network operation and service provisioning. However charging and billing models in use today are relatively simple with time-based and volume-based charging as well as flat pricing being some prominent examples. But with the convergence of communication networks, the increasing integration of different access technologies and the advent of a wide range of new services the requirements for adequate charging and billing will most probably increase. Especially when services are bundled to customer-specific product offers more sophisticated forms of charging and tariff models will have to be supported to realise an advantage over competitors by addressing issues like specific content, quality of service and cross-channel discounts and allowances.

Although a flat pricing approach can mainly reduce the complexities involved in the charging of services, it is not a real solution to the outlined problem. It will only allow for an undifferentiated charging of customers thereby ignoring the need to address individual demand with fitting customer-specific product offers and loyalty programmes. Also 'bit pipe' degradation, customer dissatisfaction and problems with customer protection legislation such as mandatory cost indications may result from too simplified approaches to this topic. Therefore a flexible, scalable, efficient and feature-rich charging solution is to be seen as an important cornerstone for commercial success in future networking scenarios. Regarding the used terminology, there is a certain divergence between the mobile telecommunications [1] and the Internet communities [2] which has already been examined in [3]. In this article we use the term "charging" in a broader sense (cf. Fig.1) that denotes the overall process of identifying and recording chargeable events in a network (Metering), the Collection and formatting of the gathered information (Accounting), its transmission as well as subsequent evaluation (Rating) in order to directly charge the customer's account for his resource consumption in a dialogue-based procedure (Online Charging, which is often used to realise pre-paid billing as payment method) or deferredly at a later point in time (Offline Charging/Billing, often used to implement the post-paid payment method), for instance by a single bill normally at the end of the accounting period (usually one month). This understanding follows to a large degree the definition of the 3rd Generation Partnership Project (3GPP) that also uses "Charging" as an umbrella term for all involved (sub)processes. We note that depending on the respective community, the terms "Accounting" or "Billing" could also be used as overall terms.

In this article we survey recent developments in charging and billing architectures comprising both standardisation work as well as research projects and examine to what extent key features are present. To do so, in Sections II and III, we describe the main key charging and billing features for tomorrow's communication and service environments. Then, in Section IV, we give a detailed overview of the considered architectures and solutions mainly concentrating on the Internet and the mobile telecommunications world (up to an including today's 3rd Generation (3G) mobile networks, such as UMTS (Universal Mobile Telecommunication System) or cdma2000). Also peer-to-peer charging solutions are presented as recent tendencies of decentralised networking approaches and that may serve as access extensions for operator-controlled infrastructure networks or as "networks before the managed network". Finally, in Section V, our findings are wrapped up by a comparison on key features of the surveyed architectures and solutions and concluded with trends that can be derived from this comparison.

II. Changing Charging and Billing Requirements

Despite the apparent differences between the various application areas, certain basic functional charging and billing

Manuscript received 25 June 2010; revised 16 November 2010 and 21 November 2010.

R. Kühne was with University of Tübingen, Tübingen, Germany. He is now with the Hasso Plattner Institute at the University of Potsdam, Germany (e-mail: ralph.kuehne@computer.org).

G. Huitema is with University of Groningen / TNO ICT, Groningen, The Netherlands.

G. Carle is with Technical University of Munich, Munich, Germany.

Digital Object Identifier 10.1109/SURV.2011.122310.000084

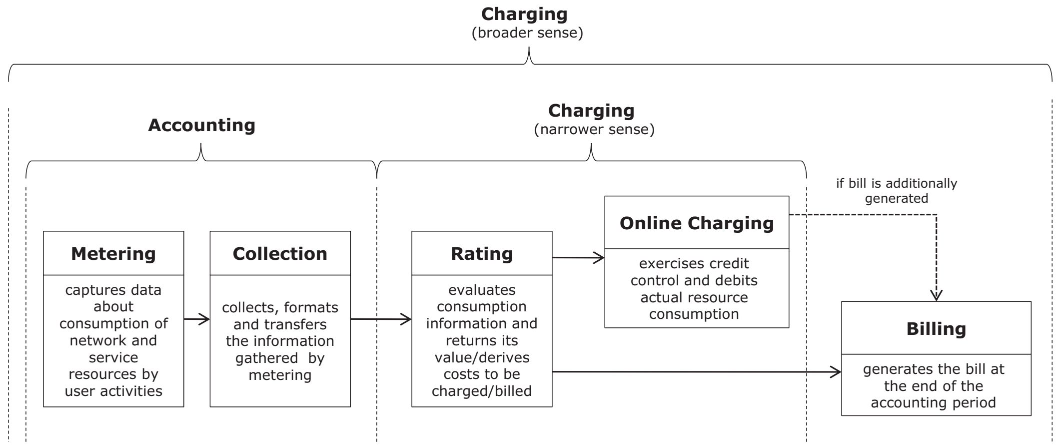

Fig. 1. Charging and Billing Sub-processes and Terminology

requirements can be identified (cf. [4], [5], also [2]) which any charging and billing system has to fulfil in order to carry out its basic task: charging the right user correctly. On the one hand, they refer to collecting and recording information for the relevant user activities (Accounting) and comprise the completeness, correctness and integrity of this information as well as guaranteeing the user's accountability for the resources he or she has consumed (e.g. duration of a phone call, number of sent short messages, volume of transmitted data, downloaded content, clicking a link, etc.; cf. also [6]). On the other hand, this information then needs to be processed in order to derive the value of the consumed resources that the user has to pay for, i.e. a correct and efficient rating and subsequent transactions that transfer the derived value from the user to the party that provided the consumed resources. How such an account management is realised depends on the application area and can reach from classical post-paid or pre-paid accounts and credit card transactions to micro-payments or the exchange of community money. Beyond that, non-functional requirements like scalability, security, extensibility and reliability are important issues. When we look at these functional and non-functional requirements towards charging and billing today, the question arises if and how these requirements will change in future. In order to answer these questions, we have to investigate next-generation charging and billing, i.e. which features will be required. It is of course a challenging task to predict such features, but the trends and developments already mentioned above, i.e. service diversity, access system heterogeneity and a competitive market will serve as good indicators of what charging and billing will have to offer in the future. Besides the usual features of correctness, completeness, security and accountability that are and will remain in the future absolute necessities for conducting business and carrying out charging and billing, we here come to five functional features that we see as prerequisites for charging and billing solutions in future network and service environments: 1) Cost transparency; 2) Online charging capabilities and charging convergence; 3) Easy introduction of new services; 4) Synchronisation of charging processes and 5) Configurability. Below, we will introduce these features and the reasons for their future importance.

A. *Cost Transparency*

Cost transparency refers to informing the user about the charges he or she has already incurred, is incurring or will incur by some network activity or service consumption. Besides a monthly itemised bill, which shows all charges of consumed resources and used services of the operator and third-party providers in this accounting period (one-stop-billing), more up-to-date information for customers is gaining interest. This can be seen for instance by 3GPP's newly started work on revision and extension of older specifications from the time GSM was introduced. The reason for this is that so-called *Advice of Charge* (AoC) carrying this cost information may prove a competitive advantage for service providers to address customers' demand for a higher degree of cost transparency as well as to help to overcome user inhibitions to use new services as the resulting costs are not obvious. Also, e.g. for premium services like downloading ring tones or calling special service numbers, regulatory measures may enforce information towards customers about the price of using certain services. In such cases, the capability to provide AoC before resources are consumed or services are used (predictive AoC), becomes a prerequisite to be provided to. Regarding requirements towards the accounting system, this transparency feature especially emphasises the need for real-timeliness in accounting systems. In particular for predictive AoC, user activities must be captured and directly presented to the user before costs incur and that the user may chose not to continue with his currently started activity.

B. *Online Charging and Online/Offline Convergence*

This key feature relates to real-time capabilities of both accounting and charging/billing. Overall real-timeliness of provisioning and processing accounting data (either by employing an online charging system or by a hot-billing solution on the charging/billing side) in combination with an efficient accounting is a prerequisite for effectively supporting pre-paid customers without exposing the service provider to high credit risks and/or the user to long waiting times (which may prohibit service usage in the worst case). Online charging capabilities are of special importance as pre-paid billing is one of the continued success stories of mobile telecommunications. Additionally, real-time capabilities also enable to enforce self-imposed budget limits for post-paid subscribers. This latter application of online charging can gain momentum in the nearer future by customer demand or, more pressingly, by regulation. Closely connected with providing pre-paid billing, is the call for charging convergence. A convergent charging solution aims at avoiding the double effort that would normally be caused by the operation of two separate systems for online and offline charging/billing, respectively. Both systems in parallel require investments and maintenance and probably limit the possibility to treat pre-paid and post-paid customers alike regarding offered tariffs and employed charging models. It has therefore quite some potential, but also has to deal with difficulties stemming from this convergence. This mainly relates to the hard real-time requirements of the dialogue-based online charging, which will deter a service usage until it has been monetarily authorised, and that therefore demand for a different system layout than the relatively straightforward offline requirements when it comes to processing and rating incoming accounting data. A practical solution is to prioritise accounting data for online-charged users by dynamically assigning the resources of the converged charging system.

C. Introduction of New Services

The third key feature is about the degree of flexibility regarding the administration, extension and maintenance of a charging system. If new services are introduced, high integration costs may arise from the necessity to modify parts of the existing system to accommodate the new service's charging process to the already present ones. A prominent example for such problems is the Short Message Service (SMS) that could not be provided to pre-paid customers at first, because no standardised online charging capability existed in the network for some time. These causes for long time-to-markets will not be acceptable anymore for future network operators or service providers, especially in the highly competitive mobile telecommunications business.

D. Synchronisation of Charging Processes

The fourth key feature is about the synchronisation of charging processes. If accounting data is generated in an uncoordinated fashion at different places and on different levels in the network, the recorded accounting data needs to be correlated and assigned to the actual event or activity that shall be charged. Thus the lack of synchronisation does not only cause load on the respective data generating entities, but also in the network and on the entities participating in the subsequent charging steps. If synchronisation between the different accounting resources can be achieved the correlation effort decreases significantly and network as well as charging resources are freed.

E. Configurability

An important key feature of next-generation charging and billing is that of configurability. In our context it denotes how the charging of a user's resource consumption and the accounting of the corresponding user activities can be initialised and controlled. If implemented efficiently, it helps to easily change and adapt tariffs to address market developments and individual needs of customers as well as to cut administrative costs. In this sense, it goes together with the two previous features – the easy introduction of new services and the synchronisation of charging processes – as the three overlap in their effect and in their consequences. As we will see in the following survey, configurability is a prominent issue for research and standardisation since configuring and administering the resources of accounting, charging and billing systems always is a work-intensive and therefore costly task. This is even more the case, since the mobile telecommunications market is highly competitive and is also reaching saturation, operators and service providers have to intensify efforts to decrease costs in order to stay profitable despite declining average revenue per customer. All measures that help to cut operational and capital expenditures are therefore of importance and searched for. This is especially true, since it is not easily possible for a provider to increase its customer base in such market environments, so it needs to find other means to stay profitable and in business. Besides, technological progress, an increasing diversity of services and convergence tendencies will entail an ever-increasing complexity. Examples, as we already mentioned above are the fixed-mobile convergence efforts, the advent of a wide range of new value-added end-user services combined with the introduction of service creation frameworks as well as ubiquitous computing endeavours. We therefore foresee that simple configurability alone will not allow realising the full potential that is necessary to support these trends and at the same time keep the costs on a reasonable level. Rather self-configuration capabilities of the charging and billing systems will become a necessity in such future environments.

III. State of the Art of Charging and Billing in Standardisation

In the previous chapter, we listed a number of features of future charging and billing systems and explained why these will be important for next generation network and service environments. In this and the next chapter we will survey the state of the art in charging and billing systems based on these features. This chapter contains standardisation work both from an industrial as well as from a research point of view, and in particular compromises the work of the 3GPP, the work of the Internet communities IETF/IRTF. In the previous chapter, we listed a number of features of future charging and billing systems and explained why these will be important for next generation network and service environments. In this and the next chapter we will survey the state of the art in charging and billing systems based on these features. This chapter contains standardisation work both from an industrial as well as from a research point of view, and in particular compromises the work of the 3GPP, the work of the Internet communities IETF/IRTF. Besides these two, another key player in this area is the industry association TM Forum (TeleManagement Forum)[7]. TM Forum focuses on enabling best-in-class IT for service providers in the communications, media, defense and cloud service markets. It provides business-critical industry standards and expertise to enable the creation, delivery and monetization of digital services. TM Forum consequently has a different focus, namely on business processes which results in less detailed technical definitions which could be of interest for our current discussion. Nevertheless, it is interesting to note that TM Forum's work on business processes resulted in a widely accepted integrated business architecture, called *Frameworx*. Key element of this business architecture is the TM Forum's Business Process Framework *eTOM* [8]. It is the industry's common process architecture for both business and functional processes and has been implemented by hundreds of service providers around the world. The eTOM process framework consists of a hierarchy of process elements that capture process detail at various levels. With the eTOM process framework one can drive down operational costs by analyzing all facets of an organization's processes, thereby eliminating duplication, identifying missing process steps, expediting new development, and simplifying procurement. Charging-related topics can be found within eTOM (version 9.0) under the term "Billing & Revenue Management" as a vertical process grouping. This part of eTOM got a boost by the Global Billing Association's (GBA) joining the TM Forum in 2007, such that not only offline charging is addressed but also online charging is now included. Traditionally, the billing system manufacturers, members of the formerly GBA, and now of the TM Forum, were active in those areas that the 3GPP considered out of scope, namely the billing domain. However, they also provided so-called hot billing solutions that were and still are in direct competition with offered online charging products according to 3GPP specifications. However, because of the business process-focus, we will not consider these solution in more detail in the further course of this article. In the following chapter we will then have a look at the work done in a number of recent research projects (MOBIVAS, Moby Dick, Daidalos, 3GET/ScaleNet, MMAPPS) in the field.

A. The 3GPP Charging Standardisation Work

In combination with the further development of mobile telecommunications networks extensive charging standardisation work has been done by the 3GPP [9]. This work is still ongoing. The 3GPP (Third Generation Partner Project) was created in 1998 to specify a third generation mobile telecommunication system fulfilling the requirements of the International Telecommunications Union's (ITU) IMT-2000 (International Mobile Telecommunications-2000) initiative. The objective was an evolutionary further development of the then already existing GSM (Global System for Mobile Communication) core and supported radio access technologies, which resulted in UMTS (Universal Mobile Telecommunications System). Today, this original scope has been largely extended to also include the evolution of third generation and beyond mobile systems and technologies (e.g. HSPA High Speed Packet Access, LTE - Long Term Evolution), the maintenance of GSM and its corresponding technologies (e.g. GPRS - General Packet Radio Service, EDGE - Enhanced Data rates for GSM Evolution) as well as the evolution of the IP Multimedia Subsystem (IMS) in an access-independent manner. For these purposes the 3GPP prepares, approves and maintains globally applicable standards, called Technical Specifications (TS). [10] The organisational partners that form the 3GPP today are the Association of Radio Industries and Business (ARIB) in Japan, the Alliance for Telecommunications Industry Solutions (ATIS) in North America, the China Communications Standards Association (CCSA), the European Telecommunications Standards Institute (ETSI), which pioneered the concept of a partnership project back in 1998 and has only recently integrated its charging work of ETSI TISPAN (Telecommunications and Internet converged Services and Protocols for Advanced Networking) that has played the key role in creating the Next Generation Networks (NGN) specifications into 3GPP, the Telecommunications Technology Association (TTA) in South Korea as well as the Telecommunication Technology Committee (TTC) also in Japan. Besides the 3GPP, there exists a 3GPP2 (Third Generation Partnership Project 2) [11], which is its American-Asian sister organisation. 3GPP2 was also born out of the IM-2000 initiative in 1998, but in contrast to 3GPP it is aiming at the further development of the global specifications for the ANSI-41 Cellular Radiotelecommunication Intersystem Operations network, i.e. the second-generation CDMA (code division multiple access) system, to a third-generation mobile system according to IMT-2000, named CDMA2000. The organisational partners of 3GPP2 are the already mentioned ARIB, the CCSA, the TTA, the TTC and the Northern American Telecommunications Industry Association (TIA). Only the latter one is not partner in 3GPP. Regarding charging standards and terminology, 3GPP2 has in principle taken over 3GPP's work with only minor or no changes and the concepts discussed in the following hold for 3GPP2 as well.

1) Logical Charging Architecture: Up to and including Release 5 the Technical Specifications are self-contained documents, in which all charging functionality necessary for the problem at hand was defined anew, independent of the fact, whether this had already been done in the context of another charging specification or not. With 3GPP Release 6 this lack of a common framework was solved by a harmonisation process leading to a more integrated approach. Since 3GPP Release 6 all common functionalities have been concentrated in a single standard [1] that serves as umbrella specification for all subsequent documents that delve into the specific charging details of the relevant technologies and services. These subsequent specifications are split up into three groups that describe the charging processes of the three different *charging levels* that were identified by the 3GPP, i.e. bearer-level charging (Circuit-Switched, Packet-Switched Domain, and the 3GPP Interworking WLAN), subsystem-level charging (IP Multimedia Subsystem) as well as service-level charging (Multimedia Messaging Service, Push-to-Talk over Cellular, Location Services, Multimedia Broadcast and Multicast Service, and in Release 8 also Short Message Service and MultiMedia Telephony [4]). As a whole they form the so-called "middle tier" specifications. Although the architectural differences between the bearer domains, the services and the subsystem influence the way in which the respective charging functionality is implemented, the charging requirements are, from a functional point of view, nevertheless similar. It was

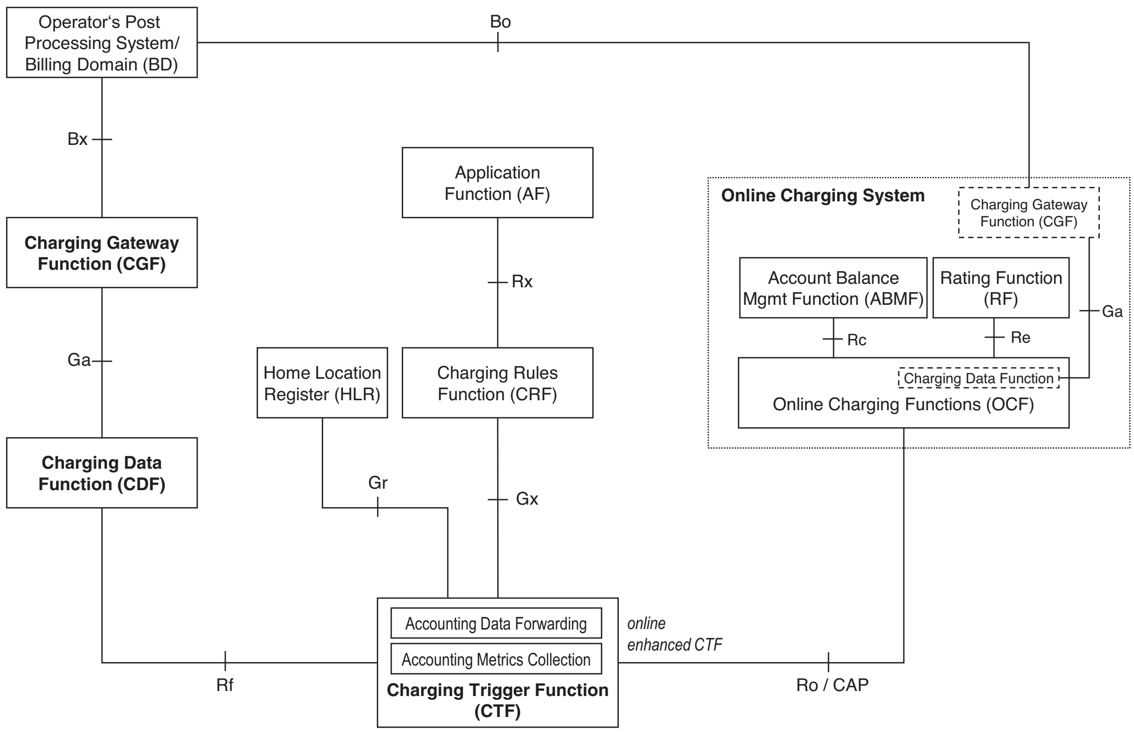

Fig. 2. 3GPP Release 6 Logical Charging Architecture with its components and interfaces

therefore possible for the 3GPP to define common logical charging functions (cf. [1]) that provide the different aspects of the needed functionality for all charging relevant parts of a 3GPP network and combine these into a common logical charging architecture, which is shown in Fig.2 including the interfaces between these functions. Regarding the charging processes in a 3GPP network, two different charging mechanisms can basically be distinguished. The first is *offline charging*, which is primarily carried out by the elements of the left hand side of Fig.2, and the second is *online charging* that involves the right hand side elements. Both will be introduced and described in greater detail in the next section.

2) Architectural Components and their Interactions: **Offline Charging Functionality.** In offline charging, the network, i.e. the charging relevant network and service entities, reports the occurred resource usage to the billing domain via a series of charging functions after this resource usage has taken place. Therefore, there is no direct, real-time interaction between the offline charging process and the service being delivered and the offline charging process consequently works on data about past events or, in other words, about services already delivered. This is the characteristic feature of offline charging. Three logical charging functions are involved in the offline charging process, namely the *Charging Trigger Function* (CTF), the *Charging Data Function* (CDF), and the *Charging Gateway Function* (CGF) as well as the relevant reference points between them and to the billing domain (cf. [4]). According to 3GPP specification, the Charging Trigger Function is an integrated component of any network or service element that provides charging functionality. It monitors the usage of the resources, which it provides or controls, and generates charging information when a so-called *chargeable event occurs*. As we have already discussed in the last chapter, chargeable events may actually be any activity that utilises or consumes resources of the network or of a service for which the CTF has been provided with metrics or filter criteria that allow for the recognition of the occurrence of an activity and of the involved identities of subscribers. The gathered information is then used to create a *charging event*, which is a data record containing a set of charging information gained from exactly one chargeable event, i.e. the one that caused its generation. The CTF itself consists of two functional blocks, the network element dependent *Accounting Metrics Collection* and the *Accounting Data Forwarding*, which provides network element independent functions for the offline charging process. The latter depends on the information made available by the former (like the recognition of chargeable events) and is responsible for generating the charging events and forwarding them via the Rf reference point. The CTF is always an integrated part of the respective network element. Real network examples for Charging Trigger Functions are the GGSN (Gateway GPRS Support Node), the SGSN (Serving GPRS Support Node), the CSCF (Call State Control Function) in all three flavours (Proxy, Interrogating and Serving), the MSC (Mobile Subscriber Server), the IMS Application Servers and many more. The Charging Data Function is the recipient of the charging events sent by the CTF. It makes use of the information contained in the received charging events to generate so-called *Charging Data Records* (CDRs) that have a well-defined content and format as specified in the "middle tier" charging specifications. The CDF may be a separate entity or may be integrated with the CTF into the network element, as it is the case for both GGSN and SGSN, for instance. In such a case, the Rf reference point is not present.

When the CDF completes the generation of a CDR, i.e. it closes the record after having added the information of e.g. several subsequent charging events, the CDR is directly sent to the Charging Gateway Function (CGF) via the Ga reference point. The Charging Gateway Function, acting as gateway of the 3GPP network to the billing domain, receives the CDR and carries out a CDR pre-processing. Finally, the CGF stores it in one of potentially several CDR files it maintains. The CGF may be a separate element or combined with a CDF. In the latter case, the Ga reference point is not present. For the transfer of the CDR files to the billing domain via the Bx reference point two options exist (cf. [12]). In the first one, the CGF transfers the CDR files to the billing domain. This push mode is triggered, for instance, at a certain time of day, at certain intervals or when a CDR file exceeds a certain size limit. In the second option, the billing domain pulls all CDR files that are ready for transfer. Rating then takes place in the billing domain, which is out of scope of 3GPP standardisation work.

Online Charging Functionality. In contrast to offline charging, online charging not only requires the occurrence of chargeable events to be recognised and recorded, but also that the respective resource usage is authorised before the actual resource consumption takes place in the network. This authorisation is obtained from the Online Charging System (OCS) upon request by the network or service elements and depends on the subscriber's account balance as well as on the rating of the chargeable events, i.e. the requested subscriber activities. Additionally, it may be limited in its extent, for instance, regarding the time or the data volume. Thus, a re-authorisation may become necessary as the subscriber's resource usage progresses (cf. [4]). The characteristic feature of online charging is therefore the direct interaction between the charging process and the service being delivered. Online charging should not be confused with pre-paid billing, since the first describes charging's mode of operation regarding the provided services and the latter the chosen method of payment. That is for pre-paid a credit account is used at which the customer has to deposit a certain amount of money in advance. Besides, online charging may also be employed for post-paid customers to enforce customer-defined budget limits. In essence, the difference between the pre-paid and post-paid method is whether the account is a debit or credit account. In contrast, for online and offline charging it is the different interaction with the delivered services, which brings forth this latter distinction. The Charging Trigger Function employed for online charging is quite similar to the one described for offline charging. This especially holds for the network element-dependent part of the CTF, the Accounting Metrics Collection. With the Accounting Data Forwarding it is different. As online charging comprises a *credit control dialogue* between the OCS and the CTF, the Accounting Data Forwarding needs to be functionally extended to support this interaction via the Ro reference point. This extension primarily relates to delaying the requested resource usage until authorisation is acquired from the OCS as well as to the so-called *quota supervision*, i.e. monitoring the resource consumption in relation to the authorised amount (quota). In this sense, the charging events created by a CTF for online charging are first of all authorisation requests asking for the provision of quota, while in the offline case it is primarily a report of resource usage that has already taken place. Additionally, an extended CTF has to be able to suppress or terminate the resource usage if the authorisation is not granted or renewed, something that is not the case for an offline CTF. Examples of network elements that take the role of charging trigger functions for online charging are again the SGSN (as the SGSN-based online charging solution is CAMEL-based the CAP reference point is used), the GGSN and the IMS Application Servers. On the other hand, the Proxy- and the Interrogating CSCF are not included and the Serving-CSCF employs a special gateway for credit control interactions with the OCS. This so-called IMS Gateway Function (IMS GWF) acts as online CTF towards the OCS, but for the S-CSCF it is an Application Server with which the latter interacts based on the Session Initiation Protocol (SIP). By the IMS GWF's translating between SIP service control and Diameter credit control [13], online charging becomes transparent to the S-CSCF. Thereby, it appears as just another service controlled by an IMS AS [14]. The already mentioned Online Charging System (cf. [15]) is the second major function involved in online charging. It comprises three logical functions, the *Online Charging Function* (OCF), the *Rating Function* (RF), and the *Account Balance Management Function* (ABMF), as well as the reference points for the interaction between them. The Online Charging Function is responsible for the correct execution of the charging transactions that take place in online charging. It exercises control over the resource usage in the network on the basis of authorisation requests received from the CTF. Based on the included charging events, it determines the associated costs with the help of the Rating Function, contacts the Account Balance Management Function to check the subscriber's account for cover, carries out direct debiting or unit reservation and returns its authorisation decision to the CTF (cf. [1], [15]). The Rating Function assesses the requested resource usage based on the information contained in the charging events. It must support the determination of monetary and non-monetary units (i.e. the rating) for the usage of network resources and of external services and applications as well as the possibility to define special discounts, bonuses or allowances for cross-product or cross-channel offers. The OCS may also generate CDRs. To this end, a CDF can be integrated into the OCF. The CDRs that contain information generated by the online charging process are forwarded to a CGF through the Ga reference point. For the transfer to the billing domain, the Bo reference point is used, which is functionally equivalent to the above-described Bx reference point. This allows for a convergent charging but requires adaptations in the rating components as we have already discussed in the previous section. For 3GPP this is however out of scope as they only aim at functions and their interactions and not at how they are eventually realised by the system manufacturers. Consequently, no standardisation is carried out regarding these internal aspects.

3) Further Aspects: Flow-based Charging and Policy & Charging Control: One type of Charging Trigger Function that deserves special notice is the Traffic Plane Function (TPF). It is part of a bigger concept called *Flow-based*

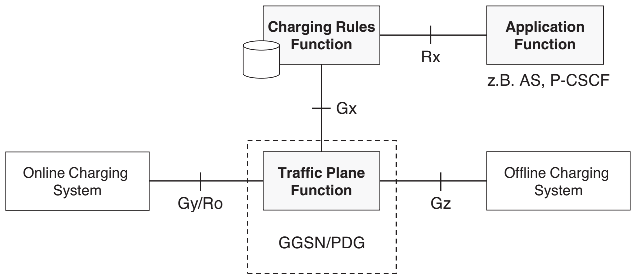

Fig. 3. Flow-based Charging Architecture of 3GPP Release 6

Charging (FBC), which realises a service-specific charging solution on the bearer level and is configurable based on policies called *charging rules*. With further work done in 3GPP Release 7 and especially in Release 8, it is about to become the pre-dominant packet bearer-level charging approach replacing the formerly separated online and offline charging solutions in the Packet-Switched (PS) Domain (cf. [16]) and also being used for the charging of alternative packet access networks, such as Interworking WLAN (cf. [17]). This work also indicates a concentration of charging functionality on the gateways and therefore a more centralised approach to charging on the packet bearer level. As such, it has to be seen in the context of this investigation as a policy-based configurability solution on the accounting level only. Additionally, it serves the synchronisation of charging processes by means of centralised design (especially, if no subsystem or service level charging is conducted anymore). Up to and including 3GPP Release 5 the smallest unit of traffic differentiation has been a single packet bearer, i.e. a single PDP (Packet Data Protocol) context. If different types of traffic were using the same bearer, this fact was transparent to the network and therefore also to the charging processes in the PS Domain. A service-specific charging was realised by means of using different *Application Point Names* (APN) for the different packet data networks or packet-based services, like Internet access, Multimedia Messaging Service (MMS), etc. In this sense, the APN identified the exit points on the GGSN to the different networks and services. To allow for a service-oriented charging below the APN/PDP context level, Flow-based Charging has been introduced with 3GPP Release 6 (cf. [18]). It adds new functionalities to the PS Domain and especially extends the capabilities of the GGSN in such a way, that it is now able to identify different data flows within a single PDP context. As depicted in Fig.3 the Flow-based Charging concept is realised by three functional elements: the *Charging Rules Function* (CRF), the *Application Function* (AF), and the already named *Traffic Plane Function* (TPF), which can be seen as a configurable Charging Trigger Function. The CRF provides charging rules and decides which rules have to be applied based on data received from the TPF, such as user- or bearer-related information, and from the AF, which provides session- and media-related information, for instance. Such rules contain information how packets belonging to a flow can be identified and how they shall be treated. For each service to be handled by Flow-based Charging, charging rules need to be specified. Two types of charging rules can be distinguished. Firstly, *predefined rules*, which are already fully configured, i.e. they contain all necessary information to be applied. They may already be installed on the TPF or provided by the CRF when needed. Already installed rules can either be active and used for flow identification or inactive, waiting for activation by the CRF. Secondly, *dynamic rules* are generated or completed dynamically by using application-specific criteria like filters to identify the flow. This information is provided by the respective AF to the CRF upon request. Dynamic rules are always installed by the CRF on the TPF. The AF provides application services to the subscriber/user for which a Flow-based Charging of packet bearer resources is required. Examples are the Proxy-CSCF and application servers. The AF may provide additional information to the CRF for charging rule selection or generation, like an application identifier, user information and packet filters (e.g. IP 5-tuple) that allow for the identification of packets belonging to the respective service data flow. The TPF is responsible for identifying and registering the user data traffic based on the charging rules provided or activated by the CRF. It is used for both online and offline charging. Flow-based Charging is therefore a convergent solution on the accounting level. In case of online charging, the TPF has also to take care of credit control issues, i.e. maintaining and supervising the assigned quota as well as of the communication with the Online Charging System (OCS). For the PS Domain, the TPF is situated in the GGSN, for 3GPP Interworking WLAN in the Packet Data Gateway (PDG). In 3GPP Release 7 the Flow-based Charging concept is extended to the so-called *Policy and Charging Control* (PCC) [19] by adding policy control features to the gateway's functionality. These mainly comprise gating and QoS control and have previously been provided by the Service-based Local Policy (SBLP) concept [20]. SBLP was originally introduced with 3GPP Release 5 to allow the IP Multimedia Subsystem (IMS) to control a packet bearer employed for a service provision regarding the QoS needed for this service. Besides, it also enables the correlation of chargeable events recorded on the bearer and the IMS level by the exchange of charging identifiers. Because of this resemblance (both are rule-based and are executed on the gateway) and the functional overlap of these two up to now separated concepts, 3GPP has carried out a full harmonisation and merging to optimise the real-time interactions in the PS Domain in 3GPP Release 7. In this process the CRF was replaced by the Policy Control and Charging Rules Function (PCRF) and the TPF by the Policy and Charging Enforcement Function (PCEF) as part of the PCC concept. Additionally, the FBC specification was withdrawn and replaced by TS 23.203 [19]. In comparison to FBC, their charging-related functionality nevertheless mainly remains the same. In 3GPP Release 8, the PCC specification is further developed to be also applicable to 3GPP and non-3GPP (e.g. WiMAX) access networks that can be connected to the new Evolved Packet Core (EPC) as devised by the SAE/LTE (System Architecture Evolution/Long-Term Evolution) initiatives of 3GPP (cf. [21]).

B. The Charging Work of the IETF/IRTF

In contrast to 3GPP, charging-relevant work by the Internet community, i.e. the IETF (Internet Engineering Task Force)

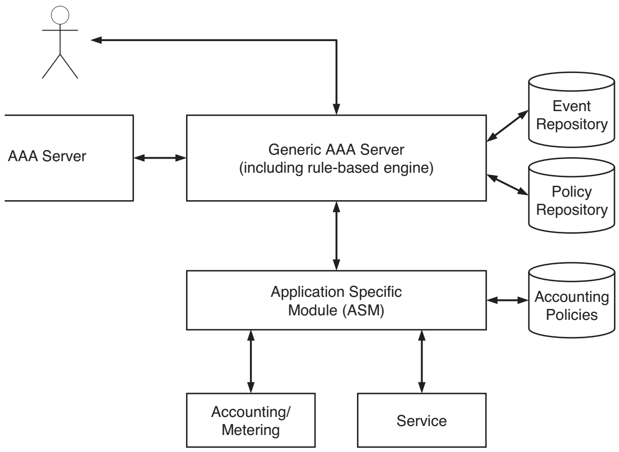

Fig. 4. IRTF Generic AAA Architecture with its components and interfaces

and IRTF (Internet Research Task Force), mainly focuses on AAA (Authentication, Authorisation, Accounting) systems and the specification of Internet protocols for accounting and charging purposes (see e.g. [22], [23], [24]). Except for [24], charging is not a topic of special interest in the Internet community, which has up until now mainly restricted its considerations to the accounting level with charging being just one application area for its accounting solutions (e.g. [2]). For this state of the art survey, the IRTF's generic AAA architecture [25] is nevertheless of importance, as two of the subsequently presented research projects explicitly built upon this solution for their vision of a future charging system.

1) Generic AAA Architecture: AAA systems have originally been directed at Internet Service Providers (ISPs) offering dial-up access via e.g. modem or ISDN (Integrated Services Digital Network). The purpose is both to authenticate and authorise dial-in users and then collect accounting data about their consumption of the provided access service resources, which explains the integration of actually distinct problem areas (authentication and authorisation on the one hand and accounting as basis for charging and billing on the other hand) into a single solution. In principle, a AAA system comprises a AAA server with user and configuration data and AAA clients located on network components, such as a Network Access Server (NAS), which request AAA services from the server. The interaction between AAA server and AAA client is governed by AAA protocols, like RADIUS (Remote Access Dial-in User Service) [22] or the more sophisticated Diameter [23]. With new requirements coming up concerning especially mobility, flexibility and scalability, the IRTF work finally has lead to the generic AAA architecture ([25], [26]) as shown in Fig.4. This architecture defines a policy-based AAA infrastructure comprising a distributed and potentially inter-domain network of interacting generic AAA servers. The special aspects of policy-based accounting – i.e. a configurable accounting based on policies – in this generic architecture are provided in detail by [27] and, with respect to metering, in [28].

2) Architectural Components and their Interactions: The generic AAA servers form the focal point in the architecture. They authenticate users, process authorisation requests and collect accounting data about the service consumption by users. The communication between them takes place by using a standardised, but not further detailed AAA protocol. The processing of incoming AAA requests is based on policy rules that are evaluated by the rule-based engine of the AAA server in order to come to an authorisation decision. These rules are retrieved from a policy repository that contains all available services requiring AAA support as well as the respective rules to be applied. Another database, the authorisation event log, allows storing time-stamped events that took place in the AAA server. Such events may also be referenced in policy rules for AAA decision making in a stateful manner. The interaction between a AAA server and the services for which the AAA functionality is required, takes place via so called *Application Specific Modules* (ASMs) that manage the corresponding service and accounting resources. This means, that besides the common interactions that can be realised by a general infrastructure, there are also requirements that are specific to the different applications and services requiring AAA functionality, like a configuration or initialisation based on service-specific protocols that are not supported by the generic AAA server. The ASM therefore can be seen as a capsule hiding these intricacies from the generic infrastructure by providing a general interface to the AAA server and carrying out the particular functionality necessary for the respective service/application or accounting equipment/meter configuration. This may also include the AAA infrastructure's transporting application-specific information or special meter configuration data that is opaque to its generic components, i.e. information that cannot be interpreted by the rule-based engines of the AAA servers but by the ASMs that are its intended recipients. If a AAA request arrives at a AAA server, it retrieves the content of the request, determines the necessary actions and queries the policy repository for applicable rules. Afterwards an interaction with an ASM may take place including the forwarding of information contained in the request. The request or parts of it may also be forwarded to another AAA server for further processing and evaluation. The accounting service in this architecture can be an integral part of the service (so-called *integrated accounting*) but also a service of its own (*discrete accounting*) [27]. In the first case, the accounting functionality is service specific and its configuration is carried out by the ASM as part of the service equipment configuration process for the user-requested service provision. For this purpose, the ASM can retrieve accounting policies from the policy repository. As a result of the tight integration of the accounting function into the service, also accounting information from service specific entities can be collected and forwarded to the AAA server as accounting records. For discrete accounting the accounting service is provided by a separate and potentially general accounting system or measurement infrastructure. Therefore, service specific accounting information cannot be collected on the service equipment itself as is the case with integrated accounting but has to be recorded by one of the meters of the measurement infrastructure. The AAA server has to request this service before the user-related configuration of the service. Accounting configuration now takes place by the accounting service's ASM which, again, has access to the accounting policies in the policy repository.

3) Further Aspects: Accounting Indications and Policies: Besides providing accounting data for charging and billing purposes, this data can also be employed for a so-called user accounting indication. This indication represents an advice of charge regarding resource consumption which has taken place within a configurable reporting interval. The user is then informed by the AAA server how many resource units he or she has consumed in the last interval. A rating function is not described. Nevertheless, a basic cost transparency is given. If accounting is a combined effort of several providers or delegated to another provider, accounting policies are exchanged between different AAA servers. The receiving AAA server acknowledges an accounting policy if enforcement of this policy is possible, i.e. the accounting task can be carried out by the available measurement infrastructure. Service denial can be a result of not being able to meet the requirements of the provided accounting policy or because the policy of the provider prohibits provision of data with the requested content or detail.

IV. State of the Art of Charging and Billing in Research

After this overview of standardised charging and billing solutions, we will now present a number of recent research projects. As we will see, some address at their time or still lacking features of the standardised solutions and therefore directly build on them. Others take a new approach. The major question will still if and to what extent they can provide the key features of a future-ready charging and billing system that we have introduced at the beginning of this article.

A. The MOBIVAS Charging Solution

The overall objective of the IST project MOBIVAS (Downloadable Mobile Value-Added-Services through Software Radio and Switching Integrated Platforms, 2000 - 2002) was to develop integrated and adaptable software platforms and systems that open up new business opportunities for value-added service providers in future heterogeneous network and service environments (cf. [29]).

1) CAB Architecture: To enable advanced service providers to realise new business opportunities, an appropriate charging and billing system is needed, that satisfies the major requirements of the envisaged service usage scenarios. One activity of the project was therefore dedicated to developing a corresponding framework and to finally proposing a solution for charging, accounting and billing which provides an advanced and configurable support for the coordination of the charging, accounting and billing processes in the operator and provider networks. Building on the 3GPP Release 4 and Release 5 solutions and also integrating AAA servers, the proposed platform ([30], [31]) consists of the so-called *CAB (Charging, Accounting, Billing) service* as the central component, two interfaces in form of Application Programming Interfaces (APIs) and the *CAB Gateway* as depicted in Fig.5. The charging architecture is separated into three layers, the transport, service and content layer on which differentiated charging information may be generated. It is primarily an offline charging architecture, as charging data collection in the operator networks draws only on offline functionality and no explicit possibilities to carry out and enforce credit control are stated. This solution allows for a one-stop billing of the end user, an automatic apportioning of the collected charges between the parties participating in the respective service provision (e.g. operator, value-added and internet service provider) and the flexible support of a variety of charging models (e.g. time-, volume-, and QoS-based charging, flat rate and one-off charges).

2) Architectural components and their Interactions: **CAB Service.** This service can either be provided by a trusted third party or forms an integrated part of the operator network itself. It comprises three components, the Charging, the Billing, and the Accounting component [32] and is responsible for collecting all relevant information to charge the user for his consumption of bearer and value-added service resources according to the operator and provider selected charging and tariff models in a single bill. The Charging component receives all CDRs generated by network and service elements and if necessary correlates them, before they are forwarded to the Billing component along with respective tariff information. The actual rating of the resource usage is then carried out by the Billing component that finally generates the bill of the user. The "Accounting" component takes care of the correct apportioning of the collected charges between the participating parties, i.e. the mobile network operators, the value-added and internet service providers involved in the charged service provision. The term "accounting" here follows the special understanding used within the 3GPP [33].

CAB Gateway. As the envisaged solution not only serves as basis for the easy introduction of new charging functionality that providers and operators want to use for their newly deployed services, but is also intended to integrate and reuse already existing charging-relevant network and service elements, the CAB service is amended by the CAB gateway as well as an interface that builds on the at that time current OSA (Open Service Access)/Parlay specification and extends its functionality where required. OSA/Parlay itself is a standardised suite of open Application Programming Interfaces (APIs) that allow for an easy and secure integration of the telecommunications network's capabilities with IT applications ([34], [35]). The MOBIVAS extension mainly aimed at deploying also advanced charging-related services like advice of charge, provision of account information or electronic bill presentment. Charging information can be directly reported to the CAB service which might be the case for certain value-added services. It can be communicated through the extended OSA functionality or via the CAB gateway. The latter is the case for non-OSA compliant parts of existing charging solutions such as the offline charging functionality of the packet-switched domain/GPRS part of a mobile operator network (Charging Gateway Function, CGF), the IP Multimedia Subsystem (Charging Collection Function, CCF) or AAA servers and other metering devices in Internet Service Provider networks. The CAB gateway is used for the collection of bearer (transport) and service usage related charging information which the different existing charging functionalities generate by providing the respective interfaces

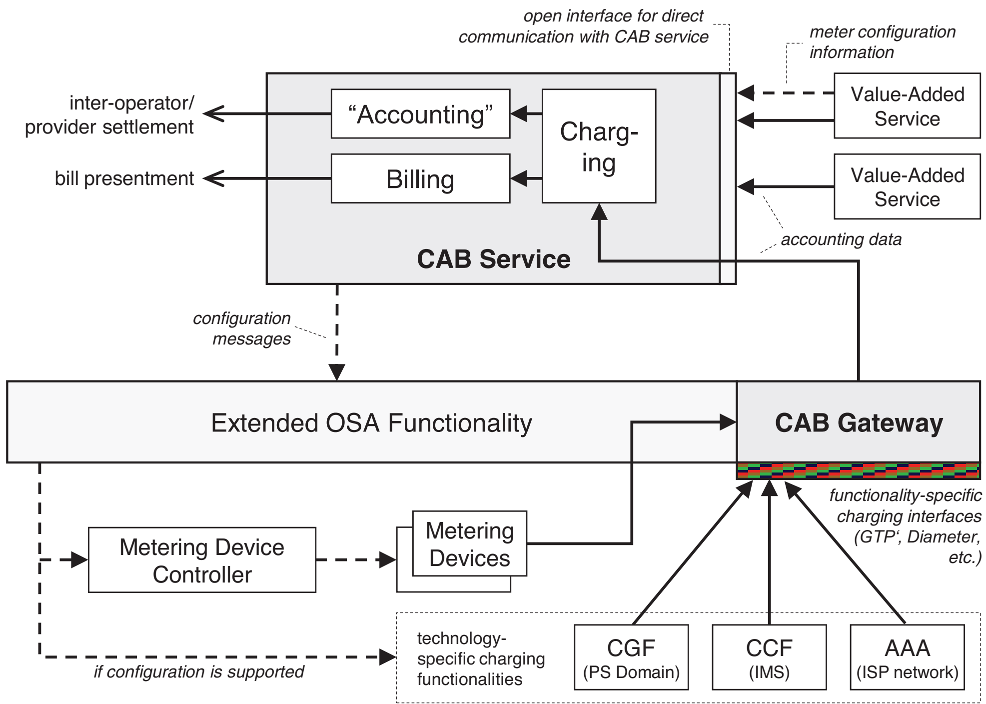

Fig. 5. MOBIVAS CAB Architecture with its components and interfaces

that they normally require. Mentioned examples are GTP' (GPRS Tunnelling Protocol derivative used for CDR transport – in this case the CGF would have to be integrated into the CAB-Gateway), Diameter (for AAA servers) but also implementations of specific APIs. The collected information is then forwarded by the CAB gateway to the CAB service.

Metering Devices. The extended OSA functionality also makes it possible to configure charging relevant elements involved in the respective charging task. This possibility especially refers to so called Metering Devices (MDs), i.e. dynamically configurable meters that monitor IP network traffic on the flow level and allow for a content-based charging. A Metering Device Controller (MDC) controls these MDs and serves as mediator between the CAB service which request the collection of charging-related information and the MDs by mapping incoming request through the extended OSA functionality on corresponding COPS (Common Open Policy Service) messages that are sent to the MDs. The MDs only records charging-relevant information of those flows for which this is requested in the respective configuration policy. The necessity for such Metering Devices was given, as MOBIVAS addressed at that time current 3GPP Release 4 [32] and Release 5 [30] specifications, which had to be extended by new functionality necessary for the advanced requirements regarding the support of different charging models that came along with the ideas of MOBIVAS. This also included the traffic differentiation below PDP context/APN level, which was not yet included as functionality in the 3GPP specifications but was needed for a service- and content-based charging.

3) Conclusion: The MOBIVAS charging solution already partly implements configurability of both the accounting layer and of the billing component regarding tariffs. New service or network elements may be introduced either by having them use the open interfaces of the CAB service or by providing the service or network-element specific interfaces on the CAB Gateway. Nevertheless, the MOBIVAS solution is restricted to offline charging and while one-stop-billing, electronic bill presentment and advice of charge after service usage are present, it lacks capabilities regarding a predictive and online cost transparency. The need to synchronise distributed accounting processes is indirectly solved by correlation, which, depending on the involved charging elements, may lead to a significant overhead and therefore to potential resource inefficiencies. With its configurable metering devices that allow for a differentiated charging based on data flows, MOBIVAS that built on 3GPP Release 4 and Release 5 anticipated the FBC concept introduced with Release 6.

B. The Moby Dick Charging Solution

Based on at that time ongoing IETF and IRTF work on Quality of Service, Mobile IPv6, and the AAA (Authentication, Authorisation, Accounting) Framework, the IST project Moby Dick (Mobility and Differentiated Services in a Future IP Network, 2001 - 2003) set itself as objective the definition, implementation and evaluation of an IPv6-based architecture both providing end-to-end QoS and mobility support to drive forward the convergence of 3G mobile networks and the Internet (cf. [29]). In order to propose a possible solution for a 4G network, functionalities have to be present that offer authentication and authorisation, mobility, quality of service as well as charging in an integrated fashion as it is already the case for 3G mobile telecommunications networks. For the Internet, in contrast, these tasks have traditionally been carried out by three independent IETF architectures without charging being an issue addressed separately, but subsumed as possible application area into the work done on accounting (e.g. [2], [36]). Thus, the charging work in the project [37] was mainly directed at extending IETF and IRTF AAA architectures with charging aspects which then support both QoS and mobility mechanisms in such an integrated fashion taking into account the heterogeneity coming along with the 4G vision.

1) AAAC Architecture: The result is an enhanced generic AAAC architecture ([26], [37], [38]) as depicted in Fig.6 with the work of the IETF and IRTF forming the basis (cf. Section III-B). This work has then been extended where it seemed necessary, i.e. primarily regarding metering, charging and auditing capabilities. It is still restricted to offline charging, as no credit control is exercised and the charging and rating processes are decoupled from the actual metering and accounting by a database that is only queried for new data in certain time intervals.

As shown in Fig.6 the core of the AAAC system is a Diameter server (cf. [23]) extended by a series of databases and a Charging Module. For the communication between different AAAC systems and with the AAAC clients, which are responsible for metering and accounting, the Diameter Base Protocol is used. AAAC clients are located on external entities in the form of Application Specific Modules (ASMs) which enable the AAAC system that takes the server role in this interaction to communicate with these entities. Besides being located on the external entities, other ASMs may also be integrated into the AAAC system itself, as is the case for the QoS ASM that provides an interface for the QoS solution to interact with the AAAC system. AAAC clients are responsible for collecting accounting information. They offer an interface to the metering functionality (AAAC/Metering Interface), e.g. on an Access Router or other measurement points chosen by the operator. This interface allows the AAAC client to configure the meter for which flows what information is to be recorded as well as to send a trigger at session start. The meter then also uses this interface to report the recorded network activity based on the passed configuration settings. The Moby Dick metering solution is based on the IETF Real-time Traffic Flow Measurement (RTFM) metering framework [39] which is integrated in a modified form into the overall AAAC solution.

2) Architectural Components and their Interactions: **The Setup Phase.** Authentication and Authorisation. If a user switches on his or her user equipment (UE) in a Moby Dick environment, only a care-of address is assigned to this UE, which only allows for its sending registration messages but no usage of services is possible. In order to do so, the UE has first to send an authentication request to the access router that also implements an AAAC client. The AAAC client then forwards this request to the AAAC system that is responsible for the respective access router. It is always the home AAAC system of the user that has to carry out authentication of the user and that for roaming cases has to check whether roaming is allowed for the respective foreign network. Thus, if the user is roaming, the AAAC system that initially receives the registration message from the AAAC client is a foreign AAAC system which consequently cannot authenticate the user. It has therefore to forward the message to the user's home AAAC system, which then queries its User Profiles database that contains authentication information regarding all its subscribed (home) users as well as DiffServ Code Points of those services, the users are each allowed to use, i.e. their respective SLA (Service Level Agreement) established with the operator/provider. If this authentication procedure was successful, the home AAAC system contacts the Home Agent informing it about the user's current care-of address. After the answer of the Home Agent, the home AAAC system in the roaming case sends this positive response that also includes the user profile with all necessary information for service provision to the foreign AAAC. Parts of this profile are then sent to the AAAC client in the access router, which subsequently allows the requested resource usage. Besides, the access router also exercises policy control together with a QoS broker based on the QoS specific parts of the user profile which are used to control the access router's QoS configuration. After authentication and authorisation of the requested service have successfully taken place, the UE may request services within the authorised limits, i.e. according to his SLA contained in the user profile. Particular services are requested by the UE's sending messages to its current access router with the DiffServ Code Points of the requested service set. This leads to the respective reservation of resources by Moby Dick's QoS solution. Additionally, the AAAC client also starts an accounting session at the home AAAC system by sending a Diameter accounting start request. In its answer to the accounting start request the AAAC system includes an accounting policy, i.e. accounting configuration settings to be applied to the service and resource consumption that takes place within the new accounting session. These settings, i.e. flow filter specifications and other parameters are then employed to configure and start the metering infrastructure belonging to the AAAC client.

The Service Usage Phase. Accounting and Charging. During service usage, the AAAC client receives data about the consumed resources from the respectively configured meters and forwards it to the home AAAC system with Diameter accounting interim requests. This dialog may occur via a foreign AAAC system in a roaming scenario and may also include re-authorisation requests. At the end of the service session or because of a handover, the accounting session is terminated by the AAAC client's sending a Diameter accounting stop request. For a handover case the new access router's AAAC client consequently starts a new accounting session by sending the appropriate Diameter request. All accounting data belonging to such a session contains the same session identifier and when received by the home AAAC system is correlated and aggregated there and stored in the Accounting Database for further processing. The Charging Module is responsible for calculating the price based on the tariff applicable to the respective service usage and resource consumption. In more detail, the collected accounting data saved in the Accounting Database together with tariff and SLA (Service Level Agreement) information stored in the User Profiles Database are drawn on for the rating process. As in Moby Dick only offline charging is investigated for a post-paid use case, the Charging Module only periodically queries the Accounting Database for newly stored and unrated accounting data which is then retrieved as input for the rating process.

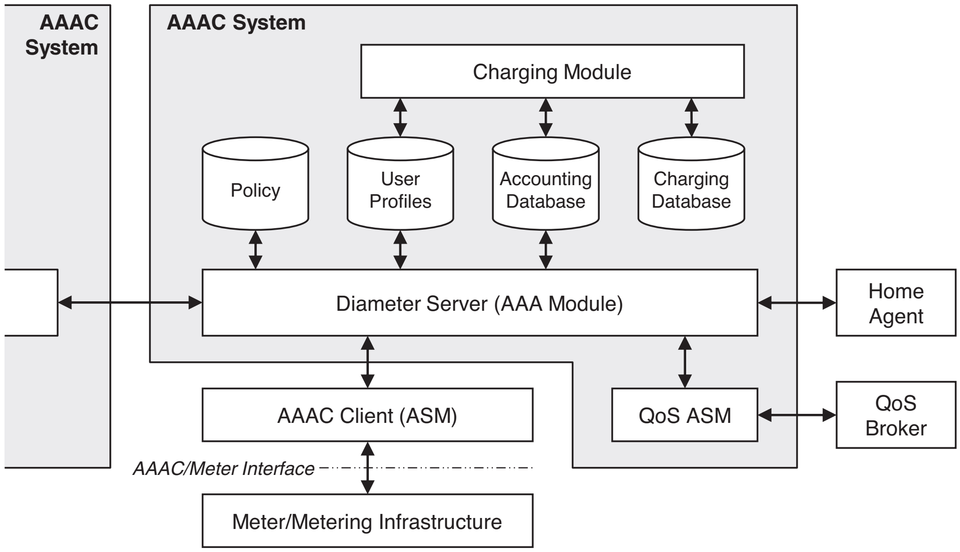

Fig. 6. Moby Dick AAAC Architecture with its components and interfaces

Based on a session identifier that has been initially assigned by an AAAC client to a user requesting some service, the appropriate tariff and SLA information can also be retrieved from the User Profile Database as necessary further input for the rating of the accounting data. The resulting charges are then written to the Charging Database that is located on the charged user's home AAAC system. With the help of a web-based user front-end, users are provided with the possibility of an electronic bill presentment in the Moby Dick prototype.

3) Conclusion: One may say that the Moby Dick charging solution is a configurable architecture focussing on offline charging. By concentrating on flow metering with RTFM, the project has laid emphasis on the transport level and the configurability of the corresponding accounting infrastructure. As a consequence, the introduction of new services generating chargeable events on the service-level is not further addressed, but by implementing service-level ASMs it should be possible to integrate such elements to report service-level data directly to the AAAC server. Flow meters as propagated by RTFM will not be of use for this purpose, except if it is possible and from an efficiency and scalability point of view acceptable to identify service-level events by deep packet inspection on the transport level. Synchronisation is achieved by correlation which may again lead to inefficiencies, but as accounting is concentrating on the transport layer and here apparently on the access routers, this may not be of significance. However, transport level accounting follows a differentiated approach based on flows and already includes the possibility for QoS-based charging.

C. The Daidalos Charging Solution

The IST project Daidalos (Designing Advanced network Interfaces for the Delivery and Administration of Location independent, Optimised personal Services, 2004 – 2008) aimed at devising an open and scalable beyond 3G communication infrastructure hiding all the intricacies of the underlying heterogeneous access network technologies from mobile users thereby allowing the seamless provision of a wide and diverse range of personalised services (cf. [29]).

1) The A4C System: Regarding the charging system, the project draws on the results from the above-described Moby Dick project, i.e. it is based on the IRTF AAA Architecture integrated with mobility and QoS and extended by charging and now also auditing capabilities. The A4C (Authentication, Authorization, Accounting, Auditing, and Charging) system, as it is called in Daidalos, especially enhances the offline system designed and implemented in Moby Dick to also support online charging.[40] The main objective is to support both payment methods prevalent today, i.e. post-paid and, in contrast to Moby Dick, pre-paid as well. For the online mechanism, the system employs the Diameter Credit Control Application [24] with the A4C system taking the server role and the Accounting Gateway (AG) functioning as client. This Accounting Gateway represents the accounting system towards the A4C system and is responsible for correlating accounting data belonging to the same or related chargeable events. All accounting data metered on the transport layer, like QoS parameters and transferred volume information, are collected by another component called the Central Monitoring System which is responsible for OSI-layer 3 metering conducted on access or edge routers, for instance. Service usage information is recorded on the service level (application layer, OSI-layer 7), such as the number of messages sent, which necessitates a correlation of related information.

2) Architectural Components and their Interactions: Online Charging. The online charging concept ([40], [41]) is reservation-based and credit control is always carried out for a number of services that form a so-called service bundle. The operator or provider can decide which services to bundle, for instance based on service usage behaviour of their customers, i.e. which services are most often used together. If a customer uses one or more services from a service bundle, a time interval is calculated that represents the reserved credit for the customer's using the bundle's services, assuming that all these services are maximally used, e.g. a data transfer takes place all the time during this interval and also employs the highest possible bandwidth. The service bundle concept is in essence quite similar to the credit pool proposed in [24] but while the units assigned to a credit pool are abstract such that the service units of a specific type, such as volume or time, need to be translated by means of a translation factor derived from the rating parameter [4] to this abstract unit, the Daidalos service bundle always has units of time assigned to it in form of a time interval. By aggregating the credit control procedure for several services, both approaches reduce the risk of credit fragmentation as well as the number of credit checks and subscriber account balance updates and therefore the signalling and server load in contrast to credit control that is independently done on a per-service basis. The advantage of this interval-based approach is that also services with non-linear tariff functions can be part of a service bundle, which is not the case for credit pools as the translation factor for converting service-specific units to the abstract pool unit is a scalar multiplier and therefore only allows for linear tariff functions. On the other hand, this also means, that for the Daidalos approach all tariff functions need to be transformed beforehand into corresponding time-based functions which imposes certain restrictions on possible tariff functions, that is, they have to be at least piecewise smooth and also monotonically increasing. Additionally, if the charging of the service is based on events, like sending a message or accessing certain content, and the frequency of these events cannot be foreseen and sensibly restricted within a given time interval or the events have different prices, the time interval-based approach cannot be applied. For instance, in the example application of the interval-based online charging approach provided in [40], the number of messages a customer is allowed to send per minute was restricted to a maximum of four. Thus, the maximum service consumption within a given time interval and with it the corresponding maximum charge can be derived, which is then used to calculate the respective time interval until the next credit check has to take place to avoid credit risk. For this calculation the so-called *Time Interval Calculation Algorithm* (TICA) is used. If a pre-paid customer requests a service, the Daidalos solution leads to the Accounting Gateway's contacting the A4C system in the respective customer's home network to carry out a credit check and request an initial quota, i.e. a first time interval. This request is forwarded to the credit control component in the A4C system which is responsible for calculating this time interval by executing TICA. As a subsequent request for a new quota will take a certain amount of time during which the customer may consume resources, the time interval calculated by TICA must take this fact into account to avoid credit risk. The time interval that represents the reserved credit therefore has to be longer than the time interval that is actually returned to the AG as quota by exactly the time span that is needed to process a new quota request and, if need be, i.e. the customer runs out of money, to terminate service provision. The latter may also take place if the computed time interval falls short of a certain lower bound that describes the shortest possible time interval the accounting system is able to support. How this case, i.e. that there is still a certain amount of money left on the customer's account, is handled, is at the operator's discretion. One possibility is that the A4C system sends an updated list of services to the AG that contains the services the customer is still allowed to use. Thus, a direct interaction between the charging process and the service rendered exists. If the calculation of an initial time interval was successful, it is returned to the AG and service provision may commence. During this interval the AG collects and correlates all recorded accounting information forwarded to it, which are sent to the A4C system when the assigned interval expires and a new time interval needs to be requested. Until this is assigned, the customer may continue to use services from the service bundle and further accounting data may therefore arrive at the AG, which will be reported to the A4C system with the next request. The accounting data forwarded to the A4C is rated and the corresponding amount of money is debited to the customer's account. A certain portion of the remaining balance is then used to calculate the new time interval to be assigned as quota to the AG responsible for the respective service bundle. When the subsequent time interval expires, also the accounting data collected during the preceding credit check phase is sent to the A4C system with the new request. Again, rating and debiting takes place and a new time interval is calculated. If the resulting time interval is finally shorter than the minimally supported interval, the AG's request for a new time interval is denied by sending a stop message. The AG communicates this decision to the components involved in the respective service provision and in the end sends all collected but not yet reported accounting data to the A4C, which takes care of the respective balance account transactions.

3) Further Aspects: Charging for Ad-hoc Access Extensions: Besides the online charging topic, Daidalos also investigates charging in so-called ad-hoc stub networks that extend the radio coverage of infrastructure networks like that of a mobile operator or an hotspot Internet access provider and how such a solution can be integrated into a AAAC/A4C architecture.[42] Besides charging service consuming users, here the rewarding of users who are willing to cooperate in the mobile ad-hoc network by forwarding traffic also becomes an issue. Since ad-hoc networks are by definition dynamic in nature, i.e. new nodes may join and member nodes may leave the network at any time, charging and compensation are no trivial tasks. In the Daidalos project, the Secure Charging Protocol (SCP) [43] is used to address this issue. It is based on the idea, that all parties that participate in an ad-hoc forwarding task should benefit when a packet reaches its destination and therefore have an incentive to cooperate. As the forwarding parties cannot derive a direct advantage from a successful communication between two corresponding parties, they need to be rewarded or compensated for their effort. SCP provides means to allow for such a compensation, which is achieved by each so-called Forwarding Node's (FN) in the ad-hoc access extension adding information to the forwarded packets proofing that it participated in the respective forwarding task by means of a hash chain and a shared secret. In this scheme, it is the responsibility of the node that finally delivers the forwarded packet to the actual recipient or to an Access Router (AR) and which is therefore called the Last Forwarding Node (LFN) to collect this information, i.e. the proofs, and take care that the AR receives it. The AR which is part of

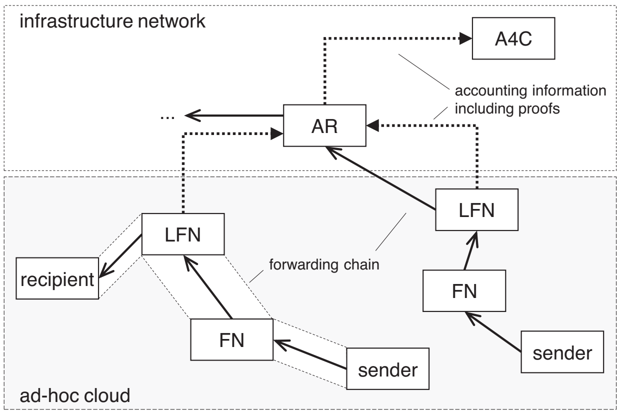

Fig. 7. Secure Charging Protocol Architecture and Daidalos Integration

the infrastructure network as depicted in Fig.7 and which connects the infrastructure network with the ad-hoc extension then forwards it to the A4C system which will respectively charge and compensate all participating parties based on the provided information. This takes place in an offline fashion. An integration with the online charging extension presented above, is however not comprised by the procedure described in [43].

4) Conclusion: We can summarise that the Daidalos charging solution builds on the Moby Dick solution, by also addressing service-level charging and by integrating the online charging mechanism. This solution for the charging of prepaid customers can be seen as a special form of online charging extension that is based on the Diameter Credit Control Application and employs a special type of credit pool that only uses time units as quota and as a result only supports a session-based charging. Event-based services are nevertheless supported by defining certain restrictions on their usage that allow for their tariff functions' being included in these special kinds of quota. More detailed information about TICA and a performance evaluation can be found in [44].

D. The 3GET Charging Solution

The main objective of the project 3GET (3G Evolving Technologies, 2003 - 2007), is the conception and development of next generation technologies for the exhaustive provision of mobile speech, data and multi-user services with special focus on the evolutionary further development of 3G/UMTS systems. 3GET was initiated and partially funded by the German Ministry of Education and Research (BMBF; cf. [45]).

1) Service-oriented and Convergent Charging Architecture: As part of 3GET, a service-oriented and convergent charging architecture has been devised ([46], [47]) from scratch to reduce the complexity of the charging task in current mobile telecommunication systems that is marked by independent and uncoordinated charging processes distributedly running on several architectural levels. This makes necessary either a coordination by manual configuration each time a new service or access technology is introduced or a correlation of collected data based on charging identifiers that for this purpose need to be communicated between the distributed processes by appropriate means that also have to be provided for each new service or access technology.

The 3GET architecture is characterised by a functional separation into three building blocks, called domains that interact with each other. Each of these domains, shown in Fig.8, is composed of several functions that realise the respective functionality.

2) Architectural Components and their Interactions: **Charging Domain.** The Charging Domain comprises the entire knowledge about services. For each service to be offered by some operator or provider a tariff has to be defined and appropriate rules for configuring the charging and accounting processes have to be available. These tariff and configuration rules form the charging policy for the respective service and are stored and maintained by the Charging Policy Management. Besides, the Charging Domain hosts the user accounts and the accompanying management functionality (User Account Management). The processing of charging records - i.e. rating, preparation for transfer to the billing domain etc. - is located here as well (Rating and Billing functions). The billing domain itself is not in the scope of the 3GET project. Authentication and Authorization Domain. While the Charging Domain "knows" about the services and how they have to be charged, the Authentication and Authorization (AA) Domain focuses on the subscriber and his eligibility to utilise certain service categories. Basic contract and subscription data for all subscribers are stored here and are provided to the Charging Domain when subscribers request services. On the basis of this information, it is possible for the Charging Domain to retrieve the appropriate charging policy and to configure its components, like Rating, for the upcoming charging task.

Accounting Domain. The Accounting Domain is the third building block and is responsible for collecting usage data being relevant to a particular service's charging. In contrast to the domains introduced above, the Accounting Domain is agnostic with respect to services and subscribers. The network elements within this domain are configured on the basis of so-called accounting policies, which contain information retrieved and refined from the corresponding charging policy by the Charging Domain. These accounting policies are then passed to the Accounting Domain's Accounting Control function, which further transforms them into directions that are used to configure those network and service elements that will do the actual resource usage metering. These so-called Accounting Entities eventually report the metered resource usage after a possible first correlation by the Collector function to the Charging Domain.

The received accounting data is subsequently rated and depending on the fact whether the corresponding user is charged online (mainly but not necessarily pre-paid customers) or offline (post-paid customers) the rated data records are either prepared for the transfer to the billing domain for user bill generation or taken as input for an update of the user account and the credit control process that checks whether the user's online account balance allows for a continuation of service usage and resource consumption, respectively. If this is not the case, the respective services' termination or a redirection to a re-charging server has to be initiated by the charging system.

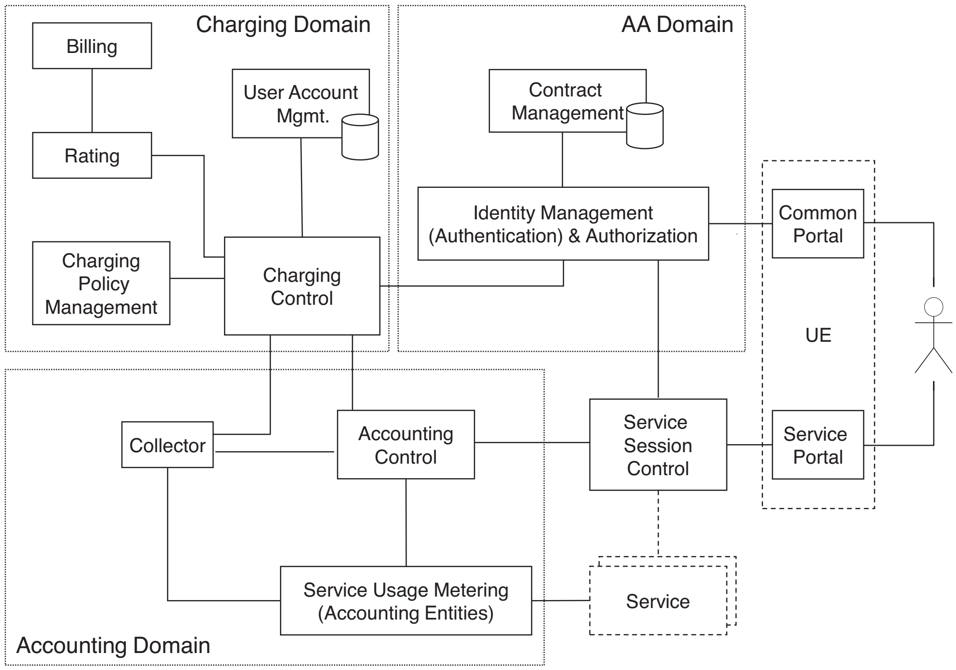

Fig. 8. 3GET Charging Architecture with its components and interfaces [47]

3) Conclusion: The 3GET charging architecture realises a general concept for charging in mobile telecommunication networks in particular and in IP networks in general through the application of policy-based networking methods and through independence from the respective service to be charged as well as from the user's payment method (pre-paid, post-paid). The question how to use policy-based networking for charging is answered in 3GET by the principle of service orientation and the hierarchical transformation of charging rules starting at the level of service descriptions and ranging down to the level where single meters at certain measurement points are configured for collecting information about data flows or other chargeable events such as accessing specific content or using particular value-added services. This also allows for a synchronisation of charging processes. Besides this extended configurability, another key feature of the 3GET charging architecture is convergence regarding online and offline charging, i.e. that irrespective of the charging mechanism selected for some user, the same charging models and as a result of this also the same tariff models may be applied to all charged services.

E. The ScaleNet Charging Solution

The project ScaleNet (Scalable, efficient and flexible next generation converged mobile, wireless and fixed access networks, 2005 – 2008) aims at providing solutions for next-generation networks enabling an efficient, integrated, flexible and cost effective working of a system comprising different wire line and wireless access and network technologies. Moreover ScaleNet addresses the problem arising from an increasing heterogeneity and complexity of future communication systems. Like 3GET the project ScaleNet was initiated and partially funded by the German Ministry of Education and Research (BMBF; cf. [48], [49]).

1) Refinement of the 3GET Charging Architecture: The charging related parts of the 3GET project mainly produced conceptual results in form of a general policy-based charging architecture. In the ScaleNet project, these concepts were turned into solutions easily applicable to mobile networks according to 3GPP Release 6, Release 7 as well as envisaged by the System Architecture Evolution (SAE) work by means of a conceptual refinement, further development and subsequent implementation. As can be seen in Fig.9 that depicts results of this refinement work the abstract 3GET domains have been replaced by a Charging System that mainly takes the place of the Charging Domain and an Accounting Infrastructure that represents the former Accounting Domain and remains mainly unchanged when compared to the 3GET approach. In contrast, the AA Domain has more or less vanished from the proposed charging architecture. This is due to the fact, that in a 3GPP network user authentication and service authorisation are already present, for instance, in the Home Location Register (HLR), the Home Subscriber Server (HSS) and the AAA Server for the 3GPP Interworking WLAN, respectively. The functionality of the AA Domain as described in the 3GET concept is therefore already implemented outside the charging system except credit control which forms an integral part of online charging and adds another level of authorisation by checking the subscriber's eligibility to use some service based on his or her current account balance and the price of the service usage. This credit authorisation aspect is clearly to be provided by the charging system and necessitates ways to interact with service provision and AA components to enforce decisions resulting from the online charging process. In the ScaleNet charging approach one central idea is that charging

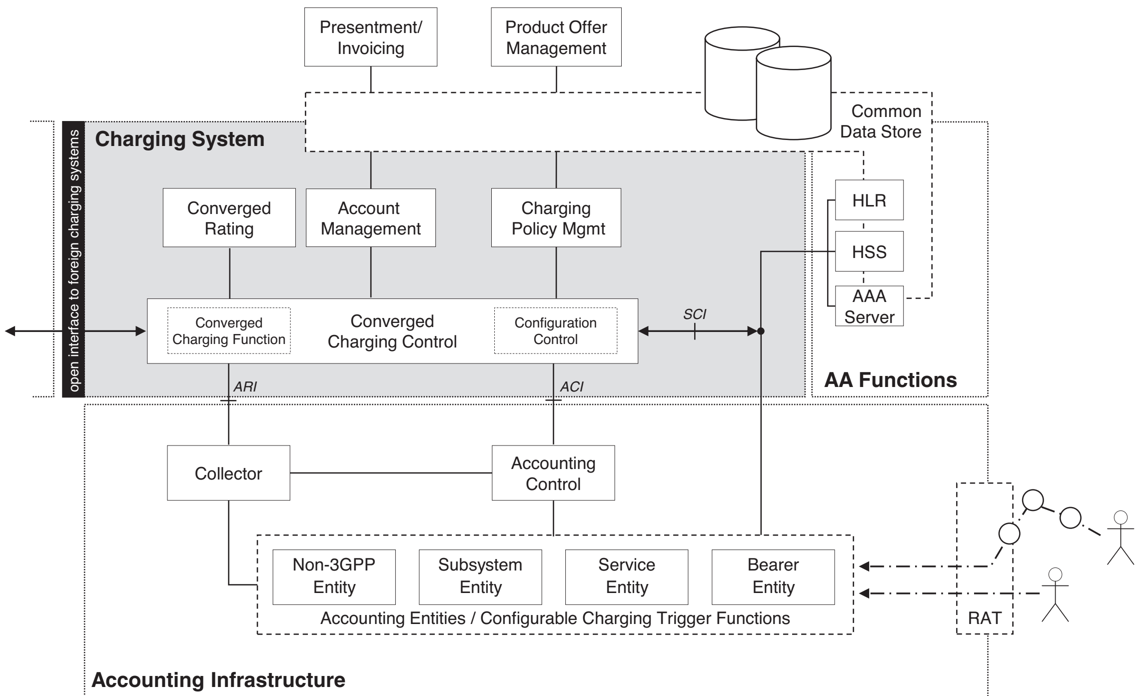

Fig. 9. ScaleNet Charging Architecture with its components and interfaces

represents a system service that has to be requested by end-user services if charging support is needed. Network or service entities (e.g. the IMS Application Server, the Proxy-CSCF, the MMS Relay/Server, etc.) that control the provision of these end-user services therefore have to request the charging service for the respective service provision which triggers the corresponding configuration process in the charging system and in the accounting infrastructure. The exact charging characteristics, i.e. whether online or offline charging shall be applied or which tariffs have to be used, are not part of these requests and the requesting entities will not even know about these with one exception. When online charging is to be applied to the service provision, then there is the possibility that the charging system will contact the requesting entity to inform it that the user has run out of credit or some budget limit has been reached. Any such entity has to support these interactions through the respective interface which is called Services and Control Interface (SCI). How the entity reacts to this kind of message is at the operator's or provider's discretion. Normally, this would lead to a service termination including a cause indication or to a re-direction of the user to a re-charging server.

2) Architectural Components and their Interactions: Based on the information provided in a charging request, i.e. primarily unique user and service identifiers together with service-specific information needed for the subsequent configuration process, the Configuration Controller, that is one of the two main sub-functions of the Converged Charging Control, calls the Charging Policy Management to retrieve the applicable charging policy for the user/subscriber-service combination. The Charging Policy Management is responsible for providing the appropriate charging policy. To this end it queries the Common Data Store (CDS) which is a potentially distributed or federated operator-wide customer and product database containing all information relevant for the generation of the charging policy, like user contract and tariff information, possible cross-channel discounts or loyalty programs the user participates in and which therefore needs to be taken into consideration for the charging of the respective service usage. Besides, also operator-defined parameters governing the charging and accounting process may be included which the Charging Policy Management needs to consider when generating the charging policy. The charging policy that the Charging Policy Management returns is made up of two parts. The first refers to the configuration of the charging system itself. These so-called Charging Needs mainly determine the charging mechanism (online or offline) to be applied as well as the configuration of the rating process with the respective service unit prices and possible discounts or bonuses. For the latter cases, the reaching of a certain number of consumed service units (also combined with the ones consumed in earlier service sessions or for other services) might lead to a credit entry or even a re-rating of already rated accounting data. For these cases, special counters need to be maintained and considered by the rating engine. The second part, the so-called Accounting Needs, define the accounting task, i.e. what charging-relevant service usage or consumption information needs to be reported to the charging system, in which interval and format reporting shall take place as well as information about how the respective service usage can be identified. The Configuration Controller then sends these accounting needs as XML-formatted Accounting Policy to the accounting infrastructure through the Accounting Configuration Interface (ACI), whereupon the accounting infrastructure tries to configure itself for the accounting task at hand. After this has successfully taken place, the Charging Controller is informed about the outcome and success is reported to the entity that requested the charging service. This response through the ACI also includes a unique session identifier to be used for all further communication between the requesting entity and the charging system. At this point, service provision may commence and while the service is subsequently being rendered, all accounting data recorded by the Accounting Entities and collected by the accounting infrastructure's Collector is reported to the Converged Charging Function, which is the second major sub-function of the Converged Charging Control. This is done through the so-called Accounting Reporting Interface (ARI)

Online Charging Concept. Based on the unique identifier passed along with the Accounting Policy, the Converged Charging Control has to decide whether the received accounting data belongs to an online or an offline charging session. This differentiation was selected for the ScaleNet charging approach, because with the integration of the online and the offline charging mechanisms into a single convergent solution, the question arises, whether this solution will be cost-effective from an economic point of view. Online charging-related accounting data has hard real-time requirements, since service provision cannot be continued without a response within a well-defined upper time bound and without accepting an increased credit risk. Otherwise, accounting data generated for offline charging is relatively uncritical regarding its processing, but in a converged solution like the one devised in 3GET/ScaleNet, both would use the same processing components. These have necessarily to be designed for the more stringent and demanding online case and combined with the additional load caused by offline-charged customers a non-optimised solution would come hand in hand with higher capital and operation expenses, so that the advantages of a convergent solution would probably be levelled out. To prevent this, the approach chosen for the ScaleNet architecture is to prioritise accounting information that is used for online charging. On the other hand, consumption information of users that are charged offline are buffered by the Converged Charging Function until the load of the rating engine drops below a certain threshold. Another option is to wait until off-peak hours for their being processed. However, a more differentiated approach is needed when hybrid subscribers or multi-user schemes need to be supported by the charging system, as missing data from the services, that are charged offline, would cause credit risk or impede budget supervision. In such cases, the fact whether a service or user under credit control is used or active needs to be taken into account as well. Regarding online charging and the credit control process, the 3GET and ScaleNet charging approaches do without an explicit online charging dialogue as specified, for example, by the Diameter Credit Control Application [24]. Rather, the reporting intervals of the accounting infrastructure are configured in a way, that no or only a negligible credit risk arises from this fact. Small account balances lead to short reporting intervals and therefore to a more frequent updating and checking of the subscriber account, i.e. an overdrawing is prevented or at least discovered relatively early after it took place, so that appropriate actions can be initiated. In this sense, the approach pursued here lies between the dialog-based mechanism used in current online charging systems based on Diameter and hot billing solutions. Reporting intervals hereby always refer to the respective service units, such as seconds, kilobytes, etc. that are employed in the chosen tariff and charging model, respectively.

3) Further Aspects: Another difference to the 3GET charging architecture is the absence of an explicit interface component that links the charging system with the billing domain. This is a result of online-offline charging convergence and the presence of a converged rating engine which removes the need for a separate rating in the billing domain and therefore reduces its functional scope to invoicing and presentment. As the Account Management in the ScaleNet charging system not only means simple balance management and updating but also storing all charging-relevant information about the service usage, all information necessary for the billing of the user or subscriber is also present in the CDS. For invoicing or electronic bill presentment purposes, the necessary information can therefore be retrieved from this CDS and no extra and consequently redundant data storing is necessary. For charging tasks that span multiple domains, i.e. different operator networks, an interaction between charging systems is possible through an open interface that allows for the exchange of accounting configuration and service usage information that would in a single domain case normally be communicated through the ACI and ARI interfaces to and from the accounting infrastructures. Future work will also need to address scenarios with higher levels of trust (e.g. direct configuration of accounting resources in the foreign network is allowed) or even delegation of parts or of the whole accounting and charging process to the foreign network as this has not been done within ScaleNet.

4) Conclusion: The ScaleNet charging architecture represents a configurable and online-offline converged solution based on charging policies that are part of service descriptions from which accounting policies are dynamically derived. New services can be easily integrated by providing a corresponding service description including charging policies to be applied. As a prerequisite, if the charging of the new service comprises new service elements, these must be configurable to allow for service-level charging. The synchronisation of charging processes is achieved by the demand-driven configuration of the accounting infrastructure which limits the correlation need and thereby increases efficiency regarding system and network load. Cost transparency is also addressed by one-stop-billing as well as an advice of charge feature providing cost indications before (predictive), during and after service usage. Its provision controlled by operator policy and user-specific settings.

F. Charging Solutions in Peer-to-Peer Networks (MMAPPS)

In contrast to the other charging architectures previously presented in this article, the one developed in the IST project MMAPPS (Market-Managed Peer-to-Peer Services, 2002 – 2004) project follows a peer-to-peer-based approach. MMAPPS itself was directed at developing and testing a software system providing management support for peer-to-peer services. In order to allow for economically viable business models for all kinds of service providers in a decentralised environment, three problems had to be solved, namely the missing possibility to provide incentives for the participants to offer resources to the peer-to-peer services and to prevent free-riding, the lack of a resource allocation that is economically efficient as well as the need for scalable communication mechanisms that support an appropriate delivery of the services to their users, see [29].

G. MMAPPS Accounting and Charging Architecture

The MMAPPS accounting and charging system, that mainly addresses the issue of accountability in peer-to-peer environments and the problems associated with it, is integrated into the service-oriented peer-to-peer middleware devised in the same project [50]. This middleware serves as distributed run-time environment for services and is functionally realised by all peers participating in the peer-to-peer network. On a single peer the local middleware implementation abstracts from the respective system platform and offers a uniform interface for the applications and services it executes through common APIs. By this way, also the distributed nature of the system is hidden from them. Common standard protocols allow a communication between the run-time environments on the different peers. The Accounting and Charging (A&C) system [51] provides its services generically and independent of the chosen so-called accounting schemes. The generic services of the A&C system are the calculation of charges by applying tariffs, a rule-based configuration of accounting schemes and the maintenance of accounting sessions. The accounting schemes are in essence account-keeping solutions and implementations and should not be confused with the term "accounting" as it is otherwise used throughout this text. These schemes can be centralised on a peer, distributed over a number of peers, follow other decentralised approaches or even be token-based compensation schemes that are realised by micro- or pico-payments ([52], [53], see also [54]). The separation between the account-keeping solution and the A&C system by means of well-defined interfaces (the A&C Core Interface and the Accounting Scheme Interface) allows for the implementation of a wide variety of such accounting schemes independent of the charging and accounting functionality, which means that accounting schemes can be used interchangeably depending on the requirements of the respective situation or environment at hand. In a sense, these accounting schemes represent a possibly distributed, decentralised account management function while the A&C system contains all the necessary functionality for session management as well as charging and rating of peer resource consumption and provision (Fig.10). It is therefore the task of the A&C system to receive and collect resource consumption information, to rate it, to derive the corresponding charge (or reward) and then to pass it as balance update to the employed accounting scheme. The accounting scheme then has the responsibility to transfer these balance updates to the appropriate accounts to be debited or credited. Regarding the accounts two types are distinguished in the MMAPPS solution. The first is the so-called session account that keeps track of the accounting information that is generated in the course of a single session, like the amount of data downloaded or the number of service units provided by the respective peer. The second is the so-called peer account that maintains charging-related information that comprises all sessions that the peer has participated in, regarding both the provision and the consumption of resources and services. Accordingly, the session account represents in a way the charging session or rewarding process of a service provision while the peer account holds the account balance of the peer that can be interpreted as its monetary credit or debit, priority level or reputation depending on the application scenario and the applied accounting scheme.

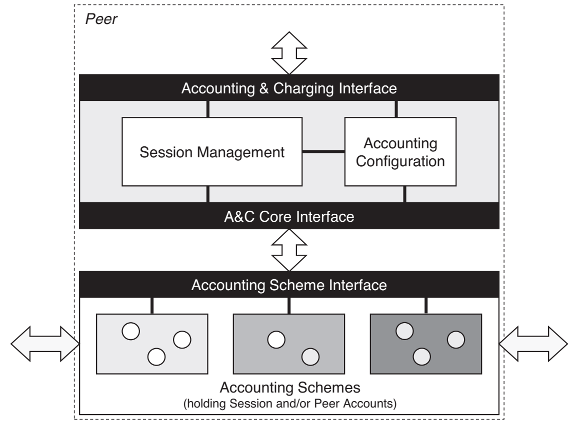

Fig. 10. MMAPPS Accounting and Charging (A&C) Architecture with its components and interfaces

1) Architectural Components and their Interactions: The A&C system provides its service to other services and applications of the MMAPPS peer-to-peer middleware through the Accounting & Charging Interface. This interface allows for a rule-based configuration of the A&C system, for the creation of so-called accounting sessions which more or less represent charging sessions, for the forwarding of charging-related information to the A&C system as well as for querying accounts. Thus, if a peer starts using a chargeable service, the service requests the creation of a new accounting session through the Accounting & Charging Interface. The request is processed by the Session Management module which creates and maintains this accounting session that also includes the SLA and tariff to be applied to charging-related information collected in the course of the respective session. Additionally, rules can be defined that govern the behaviour of the A&C system during its operation. This is for instance done, when a new application is installed to configure the A&C system as well as the accounting scheme needed for or requested by the application. Between the A&C system and the accounting schemes two interfaces are defined as already mentioned above. The A&C Core Interface is a call-back interface that allows the accounting schemes to report account balances and to notify other peers about new sessions. The Accounting Scheme Interface is implemented by every accounting scheme and serves for scheme instantiation, account maintenance and querying purposes. It is used by the A&C system. These two latter interfaces are internal to the overall charging and accounting solution, i.e. they cannot be used by other peer-to-peer middleware components. During the provision of the service, the service sends charging-related information to the

A&C system indicating the resource consumption which is accompanied by its provision. The Session Management module receives this information, selects the applicable tariff and rates the service-specific consumption information. Based on the resulting charge, so-called balance updates are generated and send to the Accounting Configuration module that notifies the accounting scheme belonging to this service about the balance updates. The selection of the correct accounting scheme is governed by the configuration policy that was provided during application or service installation. The instances of the A&C system on the participating peers do not directly communicate with each other, instead all communication is carried out by and is therefore also encapsulated in the respective accounting scheme that is used for the service. The balance update notifications are consequently handed over to the appropriate accounting scheme through the Accounting Scheme Interface and the accounting scheme then takes care about how these updates reach the correct account, be it a session or peer account. To do so, an accounting scheme employs its own protocol for the communication between all peers that implement that scheme and hold the accounts belonging to it. The same applies for account balance queries that need to be forwarded to the respective accounts. Where the accounts are held, depends on the particular implementation of the used accounting scheme. As already mentioned a wide variety of schemes is possible reaching from local and centralised approaches to remote, distributed and redundant or replicated ones [52]. Also trusted third-party peers are possible. In all these cases balance updates will be propagated to the respective peer accounts.

2) Further Aspects: A Token-based Accounting Scheme: A different accounting scheme also examined in MMAPPS is a token-based approach ([52], [53]) where the compensation between service-providing and service-consuming peer is realised by a micro-payment scheme. For that, each peer holds a certain amount of tokens in its account that it may spend to use the services that other peers provide. Double-spending is avoided by introducing an account holder peer, i.e. a remote peer that holds an account replica of another peer. If this other peer wants to spend tokens, a service-providing peer may query the account holder peer whether the offered tokens have not already been consumed before. The spending of tokens is therefore always logged by the respective account holder. Additionally, the spending of tokens is split into two phases. First, the token is sent to the service-providing peer as is, which allows the recipient already to check its validity. Only after the service has been provided, the service-consuming peer sends the token now signed by itself to the service-providing peer. This signature validates the token. The service-providing peer cannot directly spend such received tokens, though, but has to exchange them against new tokens. As no central authority exists in the peer-to-peer system that could issue tokens, this role is taken over by so-called trusted or super peers employing threshold cryptography. This means that each of these peers holds a part of a special private key that allows for a shared signing of new tokens and for which all peers in the system know the public key. A service-providing peer therefore sends in received tokens to a trusted peer which determines their validity and value and then creates new unsigned tokens of equal value. If a quorum of other trusted peers have been found by the token creating peer to also sign the new tokens with their private key part, it returns the signed and by this signing validated tokens to the service-providing peer that requested the token exchange. The latter can then spend the newly created tokens for other services offered by the peer-to-peer network.

3) Conclusion: The A&C system proposed by the MMAPPS project differs significantly from the other presented architectures and solutions as it is fully decentralised and peer-to-peer-based. It serves the charging of service-consuming peers and the remuneration of service-providing peers. These peers employ the services of the A&C system when interacting with each other. In this sense, no separate accounting system is present, as the service providing peer directly requests its compensation based on service requests from the A&C system by providing the respective accounting data. Configurability, online charging capabilities and cost transparency are also rather characteristics of the respectively employed accounting scheme than of the overall solution, as we have, for instance, seen with the token-based accounting scheme described in the previous section. How new services (i.e. peers) are introduced and how their configuration then takes place (i.e. what do they report to/request from the A&C system regarding the provided services) is however not explicitly mentioned. In the overall picture of future network and service environments, we would rather expect to find such an approach in areas of the network that are decentrally managed or are of a stub or ad-hoc nature than in the parts controlled and operated by large organisations or providers.

V. Comparison of State of the Art Charging and Billing Architectures

In the previous chapters we presented a detailed survey of recent developments in charging and billing architectures comprising 3GPP and IETF/IRTF standardisation work as well as several research projects that were or are still active in this area. The objective has been to investigate to what extent a proposed set of key features of future charging and billing systems, namely cost transparency, online charging capabilities and convergence, the capability to easily introduce new services, the synchronisation of charging processes and (self-)configurability, are present in the considered state of the art systems.

Wrap-up. Table I wraps up the findings of our survey. Besides the key features in the subsequent rows, the first two rows concisely describe the system design and the concept scope of the respective architectures or solutions. We note that with respect to the feature of online charging and convergence the summary is split up into two rows for the sake of clarity.

When we look at Table I and take into account the time when the listed architectures and solutions were conceived, the trend becomes visible that indicates an increasing convergence of the included features and with them of the developed charging and accounting solutions regarding their functionality and scope.

Convergence of development strands. In Fig.11, we depict the convergence of charging and accounting solutions in

TABLE I
COMPARISON OF DESCRIBED CHARGING ARCHITECTURES AND SOLUTIONS

	3GPP Charging Architecture	**3GPP FBC/PCC**	**IRTF AAA Architecture**	**MOBIVAS**	**Moby Dick**	**Daidalos (online extensions)**	**3GET / ScaleNet**	**MMAPPS**
System design	centralised OCS, interface to billing domain for offline charging	accounting solution, centralised on gateways	accounting solution with centralised AAA server	centralised charging system (CAB service)	centralised AAAC server	centralised A4C server, Accounting Gateway and Central Monitoring System	centralised convergent charging system	decentralised peer-to-peer solution, also depends on chosen accounting scheme
Scope of Concept	all charging levels (bearer, subsystem, services), defined in separate specs	only packet bearer, on gateway (GGSN, PDG)	all charging levels (based on specialised ASMs), but no charging in AAA server	all charging levels (transport, service/ content, subsystem)	focus on transport/bearer level (based on specialised ASMs)	all charging-levels (transport and service level)	all charging levels (bearer, subsystem/ control, services)	compensation for peer-to-peer services
Cost transparency	simple AoC defined, billing out of scope	n/a (only accounting)	user accounting indication	one-stop-billing/ electronic bill presentment, AoC after service usage	electronic bill presentment	electronic bill presentment, AoC not described	AoC before, during and after service usage, based on operator policy and user settings, one-stop-billing	not directly addressed, balance reports possible
Online charging capabilities	present	present (interface to OCS)	not present	not present	not present	present	present	present (depends on accounting scheme, incl. micro-payments)
Online-offline convergence	separate charging mechanisms	present (TPF offers both online and offline interfaces)	n/a (only accounting for offline)	n/a (only offline)	n/a (only offline)	separate charging mechanisms	present	n/a (inherently distributed, different accounting schemes)
Introduction of new services	mainly handled by new specification, cf. MMS, PoC	handled by adding new charging rules	on accounting level by employing Diameter Accounting protocol (possible new ASM necessary)	based on open interfaces/extension of CAB gateway	not further addressed as focus on transport/bearer level, probably by service-level ASMs	ASM-based on accounting level, definition of time-based tariff function and service bundle	new service element must be configurable to allow for service-level charging	A&C must be capable of interpreting reported accounting data from service
Synchronisation of charging processes	manually configured or correlation based on charging identifiers	n/a (only accounting)	not directly addressed	correlation of CDRs by CAB service	correlation in AAAC server, focus on transport layer	correlation on Accounting Gateway, correlation between charging levels not described	demand-driven configuration limits correlation need to a minimum	only separated charging processes described
Configurability	no overall solution, accounting based on policies (e.g. charging characteristics), also manual configuration during deployment	static and dynamic charging rules (interaction with AF), equivalent to accounting policies	accounting policies	by extended OSA functionality, dynamic rule generation (additional information from VAS), tariff definitions to configure billing component	accounting configuration based on policies, tariff selection based on user profiles	accounting configuration based on policies, tariff selection based on user profiles	based on charging policies as part of service descriptions from which accounting policies are dynamically derived	rule-based configuration of accounting schemes, no configuration of services described regarding accounting data to be reported

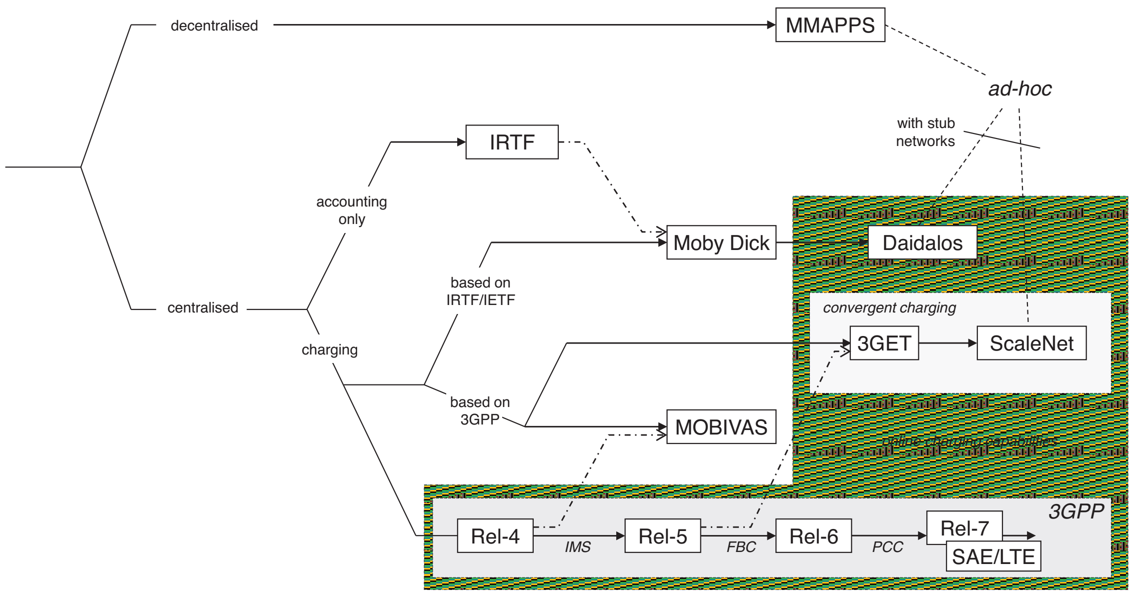

Fig. 11. Charging Taxonomy showing converging development strands

form of a "taxonomy" . Here, three principal strands can be observed:

1) in the middle, a strand of development directed at *centralised* charging solutions that originate from the Internet community building on the *IRTF AAA Architecture* and gradually extending this accounting-only solution with offline and online charging capabilities;
2) below, a *centralised* strand originating from the *mobile telecommunications sector*. There, online and offline charging mechanisms have already been available for circuit-switched technologies for quite a while, but with packet-based technologies' gaining importance to finally replace them, also the charging solutions employed there had and have to be extended in a still on-going process to address the changing situation with new services and technologies being added over the time;
3) above, a development strand, which is *decentralised* and peer-to-peer-based, and represented by the MMAPPS solution in this chapter. It is separated from the other two strands both due to its requirements and due to the resulting decentralised nature of the charging solutions to be employed in such environments.

As can be seen, both centralised strands are converging to a point regarding their provided features where they will fulfil the same needs and requirements of charging in future fixed-mobile converged networks. Although still quite separated, there are also first contact points between decentralised networks and future fixed-mobile converged networks, when mobile ad-hoc stubs are intended to serve as access extensions for operator networks by means of a multi-hop forwarding. Charging and billing systems will also have to support possibilities to reward or compensate this forwarding service's provision if operators or providers want to offer incentives for their customers to participate in this forwarding.

Fulfilling future requirements. Regarding configurability the same convergence can be found and configurability based on policies is an already widespread feature present in all architectures considered. Differences only exist in the level of this configurability, i.e. whether it is restricted to the accounting layer, whether it treats accounting and rating separately or whether an overall concept is chosen that combines both levels and derives specialised charging and accounting policies from more general charging policies or service descriptions. However, the question how these configurations can be automatically and dynamically adapted to changes in the network, in the tariff or in the accounting system itself is addressed by none of the considered architectures or solutions. This means, that the current state of the art of accounting, charging and billing does not provide mechanisms to realise self-configurability. Especially, the increasing complexity in future network and service environments with the need to decrease costs, will require more than just simple configurability, e.g. by providing Application Programming Interfaces (API) or policy-based approaches limited to parts of the overall system that are configured and administered by hand. Rather the entire system should configure itself and adapt its configuration when developments in its environment require it to do so. On the other hand, self-configurability can also serve as an important enabler for an efficient and continuous operation by adapting the system to changing situations (e.g. either transient or longer-lasting changes in usage behaviour of the customers that may lead to partial or complete failures because of overload) as well as by a synchronisation of charging and accounting processes. This argumentation gains special emphasis when current trends are carried further and the up until now quite static operator-centric business models experiences significant changes; especially so, if information and communication technology finally becomes ubiquitous and pervasive and an interworking becomes possible between all kinds of equipment and devices – as envisioned for example by the Ambient Networks [55] project – which means that everyone can eventually become a service provider. This is possible, for instance, by allowing other users to access public networks via own access points, by forwarding messages as part of an ad-hoc network, by providing services other than communication and by many more. Dynamic pricing including negotiation between the participating parties about both monetary and non-monetary remuneration therefore also becomes an important functionality in such future charging and billing solutions. Because of these changes, the classical relationship between a "small" private user and a "big" operator or provider ceases to be the only possible case, which also has consequences for charging the service user and remunerating the service provider. Here, alternative compensation schemes come into play. Compensation (cf. [56]) is thereby an umbrella term used for the mutual exchange of all kinds of non-monetary and monetary values that might take place in such future mobile and fixed network interaction scenarios. Compensation is broader in scope than charging, as charging as a term is restricted to compensate on a monetary basis between the classical business roles (users/subscribers, operators, providers). In addition, the term Compensation also comprises alternative schemes, where the involved parties exchange only non-monetary values (e.g. barter transactions) and where these parties cannot necessarily be mapped on the classical roles. Additionally, new business models and roles come into existence, like service providers that offer certain technical capabilities necessary for a correct compensation that the exchanging parties lack or specialised clearing houses that settle claims or possible disputes. Thus, charging and billing systems in such long-term future converged hybrid networks could much more resemble a situation in-between, i.e. centralised, operator-driven approaches when large providers or operators are involved, peer-to-peer interactions for the small provider scenarios and well-defined interfaces or interfacing solutions to seamlessly connect both worlds.

ACKNOWLEDGMENT

The first author wishes to thank his former colleague Mr. Gerald Görmer (Nokia Siemens Networks, Chair of 3GPP Charging Management) for his valuable comments and the many insights into 3GPP standardisation.

REFERENCES

[1] 3rd Generation Partnership Project, "Charging architecture and principles (Release 9), 3GPP TS 32.240," June 2010.
[2] B. Aboba, J. Arkko, and D. Harrington, "Introduction to Accounting Management, RFC 2975," 2000.
[3] M. Koutsopoulou, A. Kaloxylos, A. Alonistioti, L. Merakos, and K. Kawamura, "Charging, Accounting and Billing Management Schemes in Mobile Telecommunication Networks and the Internet," in *IEEE Commun. Surveys & Tutorials*, vol. 6, no. 1, pp. 50 - 58, 2004.
[4] 3rd Generation Partnership Project, "Charging architecture and principles (Release 8), 3GPP TS 32.240," December 2008.
[5] 3rd Generation Partnership Project, "Service Aspects; Charging and Billing (Release 11), 3GPP TS 22.115," November 2010.
[6] Z. Ezziane, "Charging and pricing challenges for 3G systems," in *IEEE Commun. Surveys & Tutorials*, vol. 7, no. 4, pp. 58 - 68, 2005.
[7] TM Forum, "Introduction to TM Forum," http://www.tmforum.org/IntroductiontoTMForum/5749/home.html (November 2010).
[8] TM Forum, "Enhanced Telecom Operations Map (eTOM) The Business Process Framework," http://www.tmforum.org/BusinessProcessFramework (November 2010).
[9] 3rd Generation Partnership Project (3GPP), www.3gpp.org.
[10] 3rd Generation Partnership Project, "3GPP Scope and Objectives," 2007, http://www.3gpp.org/ftp/Inbox/2008_web_files/3GPP_Scopeando310807.pdf (November 2010).
[11] 3rd Generation Partnership Project 2, "About 3GPP2," 2007, http://www.3gpp2.org/Public_html/Misc/AboutHome.cfm (November 2010).
[12] 3rd Generation Partnership Project, "Charging Data Record (CDR) file format and transfer (Release 9), 3GPP TS 32.297," December 2009.
[13] 3rd Generation Partnership Project, "Diameter charging applications (Release 9), 3GPP TS 32.299," October 2010.
[14] 3rd Generation Partnership Project, "IP Multimedia Subsystem (IMS) charging (Release 10), 3GPP TS 32.260," October 2010.
[15] 3rd Generation Partnership Project, "Online Charging System (OCS): Applications and interfaces (Release 10), 3GPP TS 32.296," October 2010.
[16] 3rd Generation Partnership Project, "Packet Switched (PS) domain charging (Release 10), 3GPP TS 32.251," October 2010.
[17] 3rd Generation Partnership Project, "Wireless Local Area Network (WLAN) charging (Release 9), 3GPP TS 32.252," October 2009.
[18] 3rd Generation Partnership Project, "Overall high level functionality and architecture impacts of flow based charging; Stage 2 (Release 6), 3GPP TS 23.125," March 2006.
[19] 3rd Generation Partnership Project, "Policy and Charging Control architecture, 3GPP TS 23.203 (Release 10)," September 2010.
[20] 3rd Generation Partnership Project, "End-to-end Quality of Service (QoS) concept and architecture (Release 5), 3GPP TS 23.207," October 2005.
[21] J.-J. P. Balbás, S. Rommer, and J. Stenfelt, "Policy and Charging Control in the Evolved Packet System," in *IEEE Commun. Magazine*, vol. 47 no. 2, pp. 68 - 74, 2009.
[22] C. Rigney, S. Willens, A. Rubens, and W. Simpson, "Remote Authentication Dial In User Service (RADIUS), RFC 2865," 2000.
[23] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, and J. Arkko, "Diameter Base Protocol, RFC 3588," 2003.
[24] H. Hakala, L. Mattila, J.-P. Koskinen, M. Stura, and J. Loughney, "Diameter Credit-Control Application, RFC 4006," 2005.
[25] C. de Laat, G. Gross, L. Gommans, J. Vollbrecht, and D. Spence, "Generic AAA Architecture, RFC 2903," 2000.
[26] C. Rensing, Hasan, M. Karsten, and B. Stiller, "AAA: A Survey and a Policy-Based Architecture and Framework,"in: *IEEE Network*, vol. 16, no. 6, pp. 22-27, November/December 2002.
[27] T. Zseby, S. Zander, and G. Carle, "Policy-based Accounting, RFC 3334," 2002.
[28] Georg Carle, Sebastian Zander, and Tanja Zseby, "Policy-basiertes Metering für IP-Netze," in *Proc. GI/ITG-Fachtagung Kommunikation in Verteilten Systemen (KIVS 2001)*, Germany, 2001, pp. 21 - 34.
[29] Information Society Technologies (IST), "IST Project Fact Sheets," http://cordis.europa.eu/ist/projects/projects.htm (November 2010).
[30] M. Koutsopoulou, A. Kaloxylos, and A. Alonistioti, "Charging, accounting and billing as a sophisticated and reconfigurable discrete service for next generation mobile networks," in *Proc. 56th IEEE Vehicular Technology Conference (VTC 2002-Fall)*, pp. 2342 - 2345, 2002.
[31] M. Koutsopoulou, A. Kaloxylos, A. Alonistioti, and L. Merakos, "A platform for charging, billing, and accounting in future mobile networks," in *Computer Communications*, vol. 30, no. 3, pp. 516-526, 2007.
[32] M. Koutsopoulou, C. Farmakis, and E. Gazis, "Subscription management and charging for value added services in UMTS networks," in *Proc. 53rd IEEE Vehicular Technology Conference (VTC 2001-Spring)*, pp. 2162 - 2166, 2001.
[33] 3rd Generation Partnership Project, "Vocabulary for 3GPP Specifications (Release 10), 3GPP TR 21.905," November 2010.
[34] 3rd Generation Partnership Project, "Open Service Access (OSA); Stage 2 (Release 8), TS 23.198," December 2008.
[35] The Parlay Group, "The No-Nonsense Guide to Parlay/OSA," 2005.
[36] J. Loughney and G. Camarillo, "Authentication, Authorization, and Accounting Requirements for the Session Initiation Protocol (SIP), RFC 3702," 2004.
[37] J. Jähnert, J. Zhou, R. L. Aguiar, V. Marques, M. Wetterwald, E. Melin, J. I. Moreno, A. Cuevas, M. Liebsch, R. Schmitz, P. Pacyna, T. Melia, P. Kurtansky, Hasan, D. Singh, S. Zander, H. J. Einsiedler, and B. Stiller, "The 'Pure-IP' Moby Dick 4G architecture," in *Computer Communications*, vol. 28, no. 9, pp. 1014-1027, 2005.

[38] P. Kurtansky, Hasan, B. Stiller, D. Singh, S. Zander, A. Cuevas, J. Jähnert, and J. Zhou, "Extensions of AAA for Future IP Networks," in *Proc. IEEE Wireless and Communications Networking Conference 2004 (WCNC 2004)*, Atlanta, USA, pp. 1516 - 1521, 2004.

[39] N. Brownlee, C. Mills, and G. Ruth, "Traffic Flow Measurement: Architecture, RFC 2063," 1997.

[40] P. Kurtansky and B. Stiller, "Prepaid Charging for QoS-enabled IP Services based on Time Intervals," Computer Engineering and Networks Laboratory TIK, ETH Zürich, 2005.

[41] P. Kurtansky and B. Stiller, "Interval-Based Prepaid Charging of QoS-Enabled IP Services," in *Internet and Network Economics, Lecture Notes in Computer Science (LNCS), Vol. 3828*: Springer-Verlag, pp. 325 - 335, 2005.

[42] S. Sargento, T. Calcada, J. P. Barraca, S. Crisóstomo, J. Girão, M. Natkaniec, N. Vicari, F. Cuesta, and M. Ricardo, "Mobile Ad-Hoc Networks Integration in the DAIDALOS Architecture," in *Proc. IST Mobile and Wireless Communications Summit 2005*, 2005.

[43] J. Girao, B. Lamparter, D. Westhoff, R. L. Aguiar, and J. P. Barraca, "Linking Ad Hoc Charging Schemes to AAAC Architectures," in *Security in Ad-hoc and Sensor Networks, Lecture Notes in Computer Scienece (LNCS), Vol. 3313*: Springer-Verlag, pp. 66 - 81, 2005.

[44] P. Kurtansky, P. Reichl, and B. Stiller, "The Evaluation of the Efficient Prepaid Scheme TICA for All-IP Networks and Internet Services," in *Proc. 10th IFIP/IEEE International Symposium on Integrated Network Management (IM '07)*, Munich, Germany, pp. 284 - 293, 2007.

[45] 3GET: 3G Evolving Technologies, http://www.3get.de (March 2009).

[46] U. Föll, C. Fan, G. Carle, F. Dressler, and M. Roshandel, "Service-Oriented Accounting and Charging for 3G and B3G Mobile Environments," in *Proc. 9th IFIP/IEEE International Symposium on Integrated Network Management*, Nice, France, 2005.

[47] R. Kühne, U. Reimer, M. Schläger, F. Dressler, C. Fan, A. Fessi, A. Klenk, and G. Carle, "Architecture for a Service-Oriented and Convergent Charging in 3G Mobile Networks and Beyond," in *Proc. 6th IEE International Conference on 3G and Beyond*, London, UK, 2005.

[48] ScaleNet: Scalable, efficient and flexible next generation converged mobile, wireless and fixed access networks, http://www.pt-it.pt-dlr.de/de/1038.php (March 2009).

[49] M. Siebert, B. Xu, M. Grigat, E. Weis, N. Bayer, D. Sivchenko, T. R. Banniza, K. Wünstel, S. Wahl, R. Sigle, R. Keller, A. Dekorsy, M. Bauer, M. Söllner, J. Eichinger, C. Fan, F. Pittmann, R. Kühne, M. Schläger, I. Baumgart, R. Bless, and S. Stefanov, "ScaleNet - Converged Networks of the Future," in *it - Information Technology*, vol. 48, no.5, pp. 253-263, October 2006.

[50] J. Gerke and B. Stiller, "A Service-Oriented Peer-to-Peer Middleware," in *Proc. 14. ITG/GI-Fachtagung Kommunikation in Verteilten Systemen (KiVS 2005)*, 2005.

[51] D. Hausheer, J. Gerke, and B. Stiller, "A Generic and Modular Accounting and Charging System for Peer-to-Peer Applications," in *Proc. 14. ITG/GI-Fachtagung Kommunikation in Verteilten Systemen (KiVS 2005)*, 2005.

[52] D. Hausheer (Ed.), N. Liebau, T. Papaioannou, S. Simpson, and B. Strulo, "Accounting Patterns, Deliverable D10.1 of the MMAPPS Project," 2004.

[53] D. Hausheer, N. Liebau, A. Mauthe, R. Steinmetz, and B. Stiller, "Token-based Accounting and Distributed Pricing to Introduce Market Mechanisms in a Peer-to-Peer File Sharing Scenario," in *Proc. 3rd International Conference on Peer-to-Peer Computing (P2P'03)*, 2003.

[54] N. Liebau, V. Darlagiannis, O. Heckmann, and A. Mauthe, "32. Accounting in Peer-to-Peer-Systems," in *Peer-to-Peer Systems and Applications , Lecture Notes in Computer Scienece (LNCS), Vol. 3485*: Springer-Verlag, pp. 547-566, 2005.

[55] Norbert Niebert, Andreas Schieder, Henrik Abramowicz, Göran Malmgren, Joachim Sachs, Uwe Horn, Christian Prehofer, and Holger Karl, "Ambient networks: An architecture for communication networks beyond 3G," in *IEEE Wireless Communications*, vol. 11, no. 2, pp. 14 - 22, 2004.

[56] G. Huitema, R. Kühne, U. Meyer, H. Ensing, A. Zugenmaier, A. Bibas, O. Karasti, F. J. Rumph, and J. Siljee, " Compensation: Architecture for Supporting Dynamicity and Negotiation in Accounting, Charging and Billing," in *Computer Communications*, vol. 33, no. 15, pp. 1823 - 1833, 2010