University of Zurich
Department of Informatics

Martin Waldburger
Cristian Morariu
Peter Racz
Jürgen Jähnert
Stefan Wesner
Burkhard Stiller

Grids in a Mobile World: Akogrimo's Network and Business Views

TECHNICAL REPORT — No. 2006.05

April 2006

ifi

Grids in a Mobile World:
Akogrimo's Network and Business Views

M. Waldburger[1], C. Morariu[1], P. Racz[1], J. Jähnert[2], S. Wesner[2], B. Stiller[1]

[1] Communication Systems Research Group, Department of Informatics (IFI), University of Zurich,
CH-8057 Zürich, Switzerland, [morariu|racz|waldburger|stiller]@ifi.unizh.ch
[2] HLRS and RUS, University of Stuttgart, Germany
D-70569 Stuttgart, Germany, wesner@hlrs.de, jaehnert@rus.uni-stuttgart.de

Abstract — **The use of wireless networking technologies has emerged over recent years in many application domains. The area of grids determines a potentially huge application domain, since the typical centralized computing centers require access from anywhere, *e.g.*, from field engineers who are situated in a wireless network domain. Thus, the integration of suitable business views on mobile grids, of grid views on available technologies, and network views in a fully IP-based network domain determines the key challenge. The Akogrimo project's architecture developed, is outlined and discussed in this paper and provides the major details required to offer a fully integrated and interoperable solution for those three views of concern.**

Index Terms — **Mobile Grid, Accounting, Charging, IP-based Grid Services, Multi-domain, Dynamic Virtual Organization**

I. INTRODUCTION

Mobility has become a central aspect of life for people in business, education, and leisure. Related mobile 3G network infrastructures and user communities have surpassed corresponding Internet figures. Independent of this development, grid technology is evolving from a niche market – solely addressing the High Performance Computing (HPC) domain – toward a framework useable within a broad business context. However, while affecting largely identical complex applications, determined as user and provider domains, until now the grid community has been basically mobility-unaware. Therefore, the Akogrimo project (Access to Knowledge through the Grid in a Mobile World) [2] aims to radically advance the pervasiveness of grid computing across regions, basically by leveraging the large base of mobile users. In order to achieve this goal and in addition to embracing layers and technologies, which are supposed to make up the so-called Next Generation Grids (NGG) such as knowledge-related and semantics-driven web services, Akogrimo architects and prototypes a blueprint of an NGG. Akogrimo exploits and closely co-operates with evolving mobile Internet infrastructures based on IPv6 and integrates security as well as dynamicity aspects.

From a technical point of view, Akogrimo addresses in an integrated manner mobility, Quality-of-Service (QoS), Authentication, Authorization, Accounting, Auditing, and Charging (A4C), and security functionalities, which are provided by corresponding network-related Akogrimo middleware systems. To pursue this goal, Akogrimo addresses at the same time the Grid and the Internet with an integrative architectural view. From a user's point of view, Akogrimo will provide the technologies and concepts to establish a 'virtual home', with nomadic and mobile environments for solving complex problems across network technology and provider domains.

In generalizing the core grid concept – namely the resource-sharing concept – Akogrimo patterns these environments as Mobile Dynamic Virtual Organizations (MDVO) [34]. The MDVO concept incorporates network-identity-based concepts of personalization, profiling, privacy, security, and trust. Finally, from the provider's and, thus, business point of view, an Akogrimo world provides new business models and opportunities, making NGGs a reality and preparing them commercially viable for future use.

Thus, this overview paper will address the essential business model classifications for mobile grids as well as the key grid and networking technologies. Based on this description the development of the Akogrimo architecture handles the actor and role interrelations within mobile grids, addressing business, grid, and network views. This approach covers key characteristics as well as major functional components required and integrates a middleware design process with a business-driven view point. To enable a complete spectrum of grid services, the respective network support integrates mobility, QoS support, a SIP-based signaling infrastructure, authentication, authorization, accounting, auditing, charging, context management, and service discovery. Thus, the key concern of a mobile grid, the offering of integrated services, has been mitigated.

The remainder of the paper is structured as follows. Section II discusses the key backgrounds. While Section III outlines the basic Akogrimo Architecture, sections IV, V, and VI determine its main business, grid, and network middleware-related aspects. Section VII finally summarizes the work and draws preliminary conclusions.

II. BACKGROUND AND RELATED WORK

The areas of interest for an integrated view of grids in the wireless domain require the careful study of next-generation grids as such and their networking technology applied. This has to be completed by a study of the usefulness of grids in the wireless networking domain under business perspectives.

A. Next Generation Grids for Business

Grid principles are no longer only applied to the domain of High Performance Computing (HPC) but they are developed towards a general purpose utility technology with the focus on service virtualization [14], [15]. Beside this technological development to change the underlying infrastructure from proprietary protocols towards standardized communication based on web services [1] technology there is a trend to virtualization, not only of low level resources such as computing power and storage, but also higher level resources such as knowledge [7].

Another important dimension is the development of business models for grid-based infrastructures. The European research project Akogrimo is leveraging its attempts to implement a commercially sustainable grid solution for telecommunications operators and service providers supporting mobile users and workers on these developments. Akogrimo, thus, envisages mobile knowledge grids for business – subsequently referenced to as mobile grids for reasons of simplicity. Knowledge in this context refers to pervasive access to knowledge, reflected by adaptive workflows, rather than artificial intelligence, such as reasoning components.

Following the definition of the key properties for Grids given by Foster in [13] Akogrimo defines a *Mobile Grid* as:

"A Mobile Grid consists out of resources that are not subject to centralized control, while it supports all kind of mobility (such as device, user, and session) and communicates using standard, open, general-purpose protocols and interfaces in order to deliver non-trivial and optimized Quality-of-Service (QoS) depending on the current context of the resource or the user. Furthermore, a Mobile Grid is communicating with the underlying network in both directions enabling cross-layer cooperation between Grid middleware and the network."

In other words, both mobile and fixed grid systems are reflected from an economic viewpoint by a virtual organization (VO). The term virtual organization was originally defined as an organizational model [31] in the domain of economics. Combining this with the idea of controlled resource sharing, as envisaged by the grid community [12] and the requirements from Business-to-Business (B2B) interactions [31] a more generic view is needed if mobility aspects should be considered. Mobile grids are characterized by two key functional extensions of fixed grid systems: First, the support of mobile grid nodes provides for pervasive, location-independent access to knowledge. Secondly, mobile grid nodes may dynamically join and leave the VO. Both aspects, mobility and dynamicity, show impacts on the one hand on respective business opportunities, e.g. by implementing context-sensitive adaptive workflows, on the other

hand they demand for enhancements of the VO concept, which does not reflect mobile and dynamic aspects as yet. This includes for example session models allowing for temporary interruption of services, when a mobile node is uncovered by a communications network. Accordingly, mobile grids are represented from a business point of view by Mobile Dynamic Virtual Organizations (MDVO), which are perceived as VOs that reflect mobile and dynamic aspects and that are defined in dependence on [22], [23], [33] as follows.

"An MDVO is a temporary or permanent coalition of geographically dispersed potentially mobile individuals, groups, organizational units, or entire organizations. They pool resources, capabilities, and information in order to contribute to the operational goals of the VO according to the dynamically established contractual agreements that are accountable and chargeable across administrative domains, incorporating mechanisms for parameterized and secure authentication as well as authorization." However, note that the fully dynamic contractual negotiations required to operate an MDVO in an open environment may be replaced by out-of-band, pre-negotiated contracts, which are pre-established between major organizations. Concerning mobility support, the extension of mechanisms and algorithms required to maintain a VO covers mobility management, identity management, and mobility support protocols. Thus, a new instance of a VO, the MDVO forms the integrated concept for Akogrimo.

B. Networking Technology

The Akogrimo architecture sketched in section III is based on state of the art network technologies, covering (a) AAA (Authentication, Authorization, and Accounting) for access control and accounting, (b) Mobile Internet Protocol version 6 (MIPv6) [19], (c) SIP (Session Initiation Protocol) [30] for session handover and signaling, and (d) Differentiated Services (DiffServ) [4] as well as IntServ for QoS reservations. Each of those is briefly addressed below on the context of mobile grids.

(a) AAA systems provide standard means for user authentication, authorization of service access and accounting of service usage. These tasks are essential in commercial networks and service provisioning environments [27]. The generic AAA architecture [9] defines network components and their interactions on an abstract level, together with a proposal of a layered structure for AAA protocols. The RADIUS (Remote Authentication Dial-In User Service) protocol [29] was designed initially to support AAA functions in dial-up and terminal server access services, and it is the most widely deployed AAA protocol in current network infrastructures. Accounting support in RADIUS is specified in [28], which is restricted to network access services. Diameter [5], which is considered as the next generation AAA protocol, is based on RADIUS, however being more flexible and providing reliable data transfer, failover mechanisms, better error handling, and security mechanisms [5]. The Diameter protocol consists of the generic base protocol and various Diameter applications. The Diameter base protocol

defines Diameter entities and specifies common functionalities, including Diameter message delivery, capability negotiation between Diameter nodes, and error handling. Diameter applications, such as NASREQ [6], enable the flexible extension of the protocol, defining service-specific commands and data units. Similar to RADIUS, Diameter message parameters and data units are coded in AVPs (Attribute-Value-Pair) allowing the protocol to be extended in a flexible manner. Besides Diameter clients and servers, the protocol defines agents, which provide more flexible means in the message forwarding, supporting relay, proxy, redirect, or translation services. Additionally, Diameter enables, in contrast to RADIUS, server-initiated messages. Finally, Diameter supports secure communication based on IPSec (Internet Protocol Security) [20] and TLS (Transport Layer Security) [10] in intra- and inter-domain scenarios.

Security Assertion Markup Language (SAML) [25] is an XML-based framework for web services that enables the exchange of authentication and authorization information. A method for transporting SAML messages over Diameter protocol is proposed in [24].

(b) MIPv6 [19] defines a protocol that allows mobile nodes to be reachable while moving around in the IPv6 Internet. The new protocol simplifies the Mobile IPv4 architecture by eliminating the need of a foreign agent (FA) and enabling the binding of a mobile's node care-of address to its home address.

(c) SIP [30] defines an application layer control protocol for creating, managing, terminating sessions between two or more participants. The possible sessions include IP telephony, multimedia content distribution, multimedia conferences, and instant messaging. SIP messages used to create sessions contain session descriptors that are used for negotiation of content media types. The SIP architecture is based on SIP proxies, which are elements concerned with routing messages to a user's current location, authenticating and authorizing users for services, implementing session-routing policies and providing service-specific features to end-users. SIP implements a registration function which allows users to upload their current location for use by SIP proxies.

(d) Finally, DiffServ and IntServ define mechanisms for delivering QoS in access networks (IntServ) as well as in core networks (DiffServ). In [3], a framework is proposed for delivering end-to-end IntServ services over heterogeneous DiffServ core networks.

C. Business Models for Current Grid Systems

Grid business models vary in their actual embodiment from instantiation to instantiation. However, grid business models consist – as business models do in general – of the following building components [32]:

- A description of involved actors and assigned business roles.
- An overview of various business flows, in particular product, information, and service flows.
- The value proposition, showing benefits for actors to be implemented.

- Sources of revenue, that indicate potentials for earnings, based on financial flows within a VO.

Besides these generic constituents of business models, specific characteristics for grid business models have to be considered. Due to the fact that grid systems offer nontangible products, based on electronic services, grid economics are highly interrelated with electronic commerce business models. For e-business models, various taxonomies exist in parallel which is exemplarily documented by an embracing overview in [8]. These taxonomies are not directly comparable in the first place. However, the integrative approach followed in [11], identifies 12 common criteria for classifying electronic business models. This allows depicting typical characteristics of today's grid systems, in particular traditional high performance computing- and data-oriented grid computing as well as service grids that allow service virtualization in VOs. Accordingly, Table 1 determines a grid business model classification for grid computing and service grids, as representatives of current grid systems. As Table 1 outlines those mentioned classification criteria applied on the existing HPC-oriented grid computing and VO-driven service grid approaches, the respective business model classification for next-generation mobile grids, thus reflecting the MDVO case drawn in section II.A, has been determined as follows:

Mobile grids will target *B2C* markets in a *1-to-Many* way, in fact a *Many-to-Many* way, which however is hidden from service users, since *virtualization* is implemented by means of service aggregation. Services, thus, are composed over a distributed value chain, that is characterized by business process execution and knowledge-integration in *hierarchies*. Targeting mass markets requires a variety of charging models and tariff schemes to be supported, since various different market segments with the need for customized *pricing* mechanisms are envisaged. *Customization* is however not limited to pricing, but it includes user and device context information so that workflows adapt to context changes, which in turn has consequences on the required *security* level of business flows: In an MDVO, where sensitive context information is exchanged across administrative domains, where high-level resources such as knowledge are *integrated* and made accessible to mobile users, security aspects in the single domains as well as with respect to information flows between domains are of utmost importance – in particular for trust-building, privacy issues and commercial service provision supported by A4C mechanisms. Aiming at offering pervasive access to knowledge leads to challenging issues to be addressed as knowledge by definition is embedded in a certain field of application where it is needed to solve a specific task. Whoever is able to provide knowledge according to context information, thus, qualifies for a *differentiation* strategy. Providers however offering basic (grid) services with a low level of integration only, are assumed to follow a *cost-leadership* strategy by focusing on core competencies. High-level resource coordination in an MDVO will involve various content elements, so that high *data volumes* are expected to be transmitted. If users then

are able to solve complex problems right on scene, clearly determines a high level of *innovation* in the mobile grid, which puts successfully operating grid service aggregators and telecommunication providers into a *strong position*.

Driven by those given indications on next-generation mobile grid business models, an overview of the Akogrimo overall architecture is given in section III, and detailed views subsequently are drawn for business, grid, and network middleware aspects in sections IV, V, and VI respectively.

Table 1: Grid Business Model Classification for Existing Grid Systems (Criteria Based on [11])

CRITERIA	CRITERIA DESCRIPTION / TYPICAL VALUES	BUSINESS MODEL CHARACTERISTICS FOR EXISTING GRID SYSTEMS
User Role	Determines whether the client is related to an end user or a provider of a service which in turn is sold to other clients. / End-user-related or reseller. B2B or Business-to-Consumer (B2C)	Clients usually are related to *end users*, however not individuals but rather organizations, i.e. *B2B* customer relations are maintained. A grid service provider for example might offer to solve data- or computational-intensive simulation tasks in automotive.
Interaction Pattern	Denotes the number of involved service providers and service consumers / 1-to-1, 1-to-Many (one service provider, many service consumers), Many-to-1 (many service providers, one service consumer) or Many-to-Many [32].	1-to-1 or 1-to-Many in case of grid computing. By forming a virtual organization, as in service grids, Many-to-1 and Many-to-Many become possible, which however is not visible for service consumers. Due to today's grids' focus on niche solutions or non-commercial research projects, *1-to-1 and Many-to-1*, respectively, are prevailing.
Nature of Offerings	Describes the basic offering characteristics. / Content/information, service or complete product (including e.g. after-sales services)	Offerings usually embrace electronic *services*, such as computational power in grid computing. In a VO, various service and *content* providers are able to cooperate, whereas a service aggregator bundles basic services. Complete products typically are out of scope, since the offered electronic services represent only one step in the user's value chain.
Pricing System	Specifies the financial terms, under which a user can access the service. / Dynamic or fixed pricing with sub-categories such as free access, usage rate, subscription fee or auctions.	The pricing scheme highly depends on the actual organizational configuration of a grid-based solution as well on the type of offered services. Both, VO-externally and VO-internally, *fixed pricing* schemes, based on *service usage*, are regarded as significant choices under the assumption that relevant usage parameters are available to be metered, accounted, and charged.
Level of Customization	Determines how much a service is customized to a user's individual requirements. / Full range from mass content to completely customized content.	Since current grid systems are characterized by 1-to-1 or Many-to-1 business relations, whereas B2B relations prevail and users are end-users (cf. respective classification criteria), mass markets are not targeted. Consequently, offered services are *customized* to the user's individual requirements.
Economic Control	Depicts the organizational alignment with respect to service provision / Full range from self-organization to hierarchical design.	Independent from the existence of a VO, the organizational structure is assumed to be *hierarchical* within business grids. A central component manages the distribution of tasks to be completed. From a user's viewpoint, in a commercial environment, only one organization is assumed to maintain direct customer relations, thus hiding organizational details of a potential VO from the user.
Level of Required Security	Identifies the needed security-level to monitor and verify purchases. / Low or high	With an orientation towards direct B2B customer relations in a 1-to-1 manner, the required level of security for monitoring and verifying purchases is assessed to be *low*. In the case of service grids (VO-enabled, Many-to-1) however, monitoring has to be increased VO-internally among the various VO members.
Level of Value Integration	Qualifies the degree of integrating and transforming components over different value chain steps. / Low or high	In a VO, several steps of the VO-internal value chain are integrated from the services provided by various VO members. The VO itself offers these electronic services in a bundled form to a user which in turn integrates them into its own value chain. Thus, assuming a service grid with service provisioning in a VO, the level of value integration is considered to be *high*. The opposite (i.e. a low level of value integration) however is the case in traditional, purely high performance computing-centric grid computing.
Value/Cost Offerings	Reflects the envisaged competitive strategy. / Differentiation (adding value) or cost leadership (low-cost or low-price strategy)	Selling purely computational power implements a *cost leadership* strategy where services are imitable and offer only little value added, they however are offered to customers at comparably inexpensive conditions. In service grids, a higher level of added value is reached by offering aggregated, application domain-specific services, which allows the grid service provider to implement a *differentiation* strategy.
Scale of Traffic	Considers the expected volume of data traffic that is generated. / Low or high	Grid systems are by design useful for solving computationally and data-intensive tasks. Accordingly, most grid services are expected to generate *high* data traffic rates.
Degree of Innovation	Assesses whether existing business is directly transferred to its electronic version or the electronic business offers new functionality [32]. / Low or high	Both forms of current grid systems, grid computing and service grids, offer a *high* level of innovation, since new functionality is gained compared to non-electronic business. This includes high performance computing facilities and virtualization of electronic services in a VO.
Power Distribution	Answers whether buyer or seller is in a stronger position (e.g. with regard to negotiations). / Buyer or seller	At this early stage of commercialization, distribution of power is difficult to be assessed. However, commercial grid service providers, representing *sellers*, are not yet numerous which puts them into a better bargaining position, assuming a constant or growing demand.

III. AKOGRIMO ARCHITECTURE

The key components of the Akogrimo Architecture are shown in Figure 1. The overall architecture is composed from four major building blocks, where three of them are in a layered structure and the fourth one has a vertical coordination role. On the lower layer, the mobile IP network is depicted. The key components consist of the IPv6-based mobile network components, the network QoS provisioning components, as well as network management components. The key role of this layer it to provide logical interfaces to the higher layer in order to communicate context information to the Akogrimo architecture's higher layers and to execute the service needs from the higher layer.
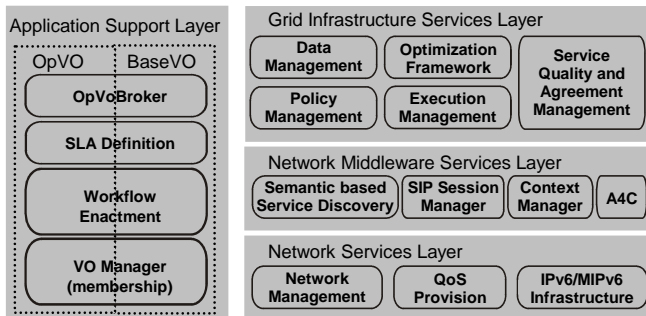
Figure 1: The Akogrimo Architecture

On the middleware layer service discovery, real-time session management, context management and authentication, authorization, accounting, and charging form key building blocks. This layer glues basically the two worlds, that of grid and that of a commercial mobile network in order to deploy grid services in a commercialized environment.

The grid infrastructure service layer consists of components that are generally seen as the fundamental elements of an OGSA-based grid infrastructure [12]. However, for Akogrimo these components have to be adapted to communicate with components on the lower layers in an integrated manner and to provide that information from higher layers down to the network infrastructure to ensure an efficient interoperation. Furthermore, additional components needed for business grids such as the management of Service Level Agreements and policy management is placed in this layer.

The application support layer provides, as mentioned above, besides its function to provide generic application domain-independent supporting services for grid-enabled applications, also the necessary components enabling the management of the cross-layer cooperation. This functionality is covered by the VO management block that enables set-up, configuration, operation, and dissolution of operational VOs, all of which is supported by the workflow enactment block. Support of this VO lifecycle and the linked lifecycle of the dynamically established or pre-negotiated Service Level Agreements and contracts define the necessary actions on the lower layers. For example, the integration into an operational VO requests the necessary settings to enable access with appropriate rights to the resources of the partners in the VO. Moreover, the defined Service Level Agreements need to be enacted and managed on the lower layers, which requires possibly the reservation of appropriate bandwidth on the (access) network layer and the configuration of AAA servers to react on validations of them.

IV. BUSINESS VIEW

From a business viewpoint, Akogrimo's interest is set on economic aspects of commercial mobile grids, which are characterized by two key functional extensions of fixed grid systems: First, the support of mobile grid nodes provides for pervasive, location-independent access to knowledge. Secondly, mobile grid nodes may dynamically join and leave the VO. Both aspects, mobility and dynamicity, show impacts on one hand on the respective business opportunities, *e.g.* by implementing context-sensitive adaptive workflows, on the other hand they demand for enhancements of the VO concept, which does not reflect mobile and dynamic aspects as yet, resulting in the respective definition of MDVOs (cf. Section II.A.).

Mobile grids are reflected by MDVOs from an organizational viewpoint, but other perspectives also have to be considered in order to gain an embracing view of relevant business aspects, particularly with regard to mobile grid business models. Following a top-down approach, starting at the organizational alignment of mobile grids and going down to the underlying technical infrastructure, two structural parts besides the MDVO are identified, namely grid infrastructure and network infrastructure. These layers are drawn from an economic viewpoint and they are not to be perceived as direct equivalents to the respective architectural layers shown in Figure 1.

With respect to the first element of business models (cf. section II.C), i.e. involved actors and assigned roles, only MDVO and grid infrastructure are regarded to cope purely with grid business models. However, since these depend on communications infrastructure to be available, respective actors and roles on network infrastructure's layer have to be included as well. Even though actors and role models show specific characteristics depending on the chosen scenario, generic constituents relevant for any mobile grid's role model are assumed to be present in all mobile grid scenarios. This assumption bases on mobile grids' distinctive characteristics from Akogrimo's commercial perspective. This means that Akogrimo scenarios adopt a mobile grid vision where, first, presumably only end-users are mobile, which does not imply that user nodes are service consumers only, but computation- or data-intensive tasks are performed on fixed grid nodes. Second, in a commercial environment, single grid services are bundled by a service aggregator that maintains end-user relations, thus taking exclusively the role of a grid solution provider towards end-users. And thirdly, grid services are aggregated from a distributed supply chain that embraces n-tier grid service providers, whereas the role of a grid service provider can show more specific instantiations, such as the one of a content provider. Figure 2 depicts the organizational alignment being typical for Akogrimo mobile grid scenarios, incorporating actors and roles from an MDVO and grid services perspective.
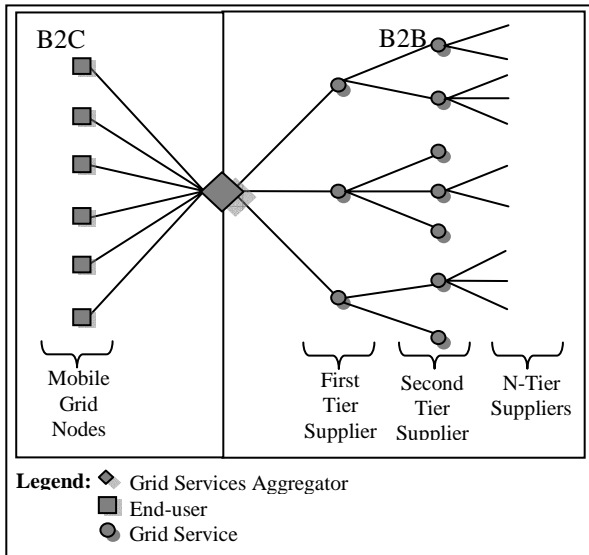
Figure 2: Generic Mobile Grid Actors and Roles in Akogrimo Scenarios

Actors and roles from the network infrastructure layer however are not explicitly shown in Figure 2 for reasons of simplicity, though assumed to be present in every scenario, since network service providers, represented for example by mobile network operators, play the indispensable role of a communications enabler. Accordingly, the following generic business roles are identified as the basic set being relevant for mobile grids that include Business-to-Consumer relations:

- End-user (potentially a separate organizational entity than the end-customer)
- Grid Service Aggregator
- Grid Service Provider
- Network Service Provider

Having denoted those four basic business roles, business models for mobile grids are analyzed subsequently in further detail using the three remaining business model building blocks, which embrace business flows, the respective value propositions, as well as potential sources of revenue (as presented in section II.C). Regarding value proposition, Table 2 lists the identified benefits for every basic business role, thus allowing to document potential values to be envisaged and achieved.

Table 2: Value Proposition for the Basic Set of Mobile Grid Business Roles

BUSINESS ROLE	VALUE PROPOSITION
End-User	End-users profit from being able to perform knowledge-intensive tasks while not being bound to a specific location anymore. Thus, services remain available even when changing locations and can be resumed after a temporary loss of network coverage. Services are reflected by context-aware, adaptive workflows, implementing various types of mobility, such as user, terminal, and service mobility [22]. Furthermore, end-users are not only able to consume service, but also to provide them, *e.g.*, by first retrieving content relevant to perform a current task, working on the content, and delivering respective results back.

BUSINESS ROLE	VALUE PROPOSITION
Grid Service Aggregator	The grid service aggregator is the producer and the vendor of a mobile grid solution. This solution makes value-added services available for end-users in form of a product. As a producer, atomic services are aggregated to service bundles. This implies that the producer runs and maintains the Akogrimo service delivery platform. As a vendor contractual relationships towards end-users, here in terms of customers, are maintained. Thus, the vendor qualifies for billing. Furthermore, the vendor copes with marketing activities and customer service.
Grid Service Provider	Grid service providers focus on their individual core competencies. Depending on the grid service characteristics, either both or only one, economies of scale and economies of scope, allow a grid service provider to implement a differentiation or cost leadership strategy while keeping transaction costs as low as possible. By the use of open, standards-based interfaces, grid service providers are able to provide their grid services in many VOs.
Network Service Provider	Network service providers offer technology-transparent access to communications infrastructure, which forms the basis for grid services on top of communications services. Mobile network operators excel as both, network service providers and grid service aggregators, due to their experience in customer relations and marketing, allowing them to move from pure access providers to providers of complex services.

Business flows and possible resources of revenue are on one hand highly interrelated, whereas on the other hand, specific indications highly depend on the chosen scenario. In general, business flows are based on technical requirements to be met. This includes guarantees with respect to QoS or non-repudiation of messages exchanged among VO members. Single business processes, regardless whether executed within an administrative domain of a grid service provider or seen from an aggregated VO perspective, all result in business flows, which in turn cause financial claims that have to be settled.

Therefore, from a process angle, A4C (see Section VI.C is regarded as key element: Authentication, authorization, accounting, charging, and billing document for every provider the complete process from service instantiation to finally billing a provider's customer. Hence, a grid accounting model, composed of meaningful accountable units that reflect process costs and resource usage is needed. Accountable units are useful on the one hand with regard to service optimization and auditing, on the other hand for price-building and billing purposes, influencing finally direct sources of revenue. In a VO however, where organization members remain legally independent, business-critical accounting information is not assumed to leave administrative domains, so that only aggregated charging records are exchanged.

Having presented various business-related aspects of commercial service provision in an MDVO leads to an in-depth investigation of the Akogrimo architecture, both from a grid and network middleware viewpoint as the two key elements by which business potentials are implemented. This includes for instance SLA management and workflow enactment on grid level (cf. section V) and A4C as well as context management on network middleware layer (cf. section VI).

V. GRID VIEW

The deployment of grids has already achieved a significant number of deployments within organizations driven by the motivation to move from large mainframes to clusters aiming at a reduction of the Total Cost of Ownership (TCO). Also in purely academic environments, where the sharing of high performance rare resources is the driver for the grid solution, cross-organizational solutions are available. However, grid solutions that enable the sharing of resources across organizational boundaries are not yet very common. The major obstacles preventing the wide take up are related to the complexity of user management, cross-organizational accounting, auditing and charging of used resources and security considerations.

A. Web Services Base Technology

All major grid toolkits nowadays are based on web service technology. Akogrimo is, despite its new challenges, coming from the business *and* network layer aiming to align with the major trend in grids and Business-to-Business relations to rely on cross organizational businesses based on web services. Besides those base technologies as defined in the WS-I Profile Akogrimo can rely on the wide range of composable specifications including WS-Addressing [35], Web Services Base Notifications [26] and from the Web Service Resource Framework [17]. Besides other specifications these are supported by all WSRF-based grid toolkits such as the Globus Toolkit 4 [16] or WSRF.NET [36]. Those have been selected as base technologies for the Akogrimo integration platform. One of the major design decisions that had to be taken within Akogrimo was on which of the layers (cf. Figure 1) the virtualization interface had to be designed. This interface provides the underlying resources and services through web service interfaces.

The originally anticipated, strictly layered approach to have this boundary between the Network Middleware Layer and the Grid Infrastructure Services Layer has been proofed to be not beneficial. Instead, and in line with the paradigm of Service Oriented Architectures (SOA) resources in the Network and Network Middleware Layer are exposed partially as Web Services and can be used by services from the Grid Infrastructure, the Grid Application Support "Layer", and application-specific services.

B. State and Asynchronous Communication

A major differentiator between stateless, distributed applications and Grid Applications rely on the WS-RF specifications for exposing state information and WS Base Notification for a more peer-to-peer like communication pattern between grid nodes. This approach is based on topic-based subscription and asynchronous message delivery.

For example, the context of a current user device might be exposed using WS-RF allowing to query and manipulate the state and also to allow these states to be transferred in case the user changes his device and wants to transfer its state to the new device. Furthermore, if the context includes value such as "Current Bandwidth" or "Screen Resolution" other grid Nodes might want to subscribe to these state changes.

Within the example of context change propagation the general concept of Akogrimo of cross-layer co-operation can be demonstrated. The collected state information from the network layer (see section VI.D) using a wide range of different mechanisms is virtualized by a web service interface and is provided to context consumers using a standardized web service interface, independent on how this information had been collected. As a result standard workflow engines or any application service can subscribe to these notifications and adapt their behavior accordingly.

C. Service Level Agreements (SLA) and Execution Management Systems (EMS)

Another basic building block for moving from e-science grids to business grids is the support of service level guarantees to the consumer of services. Akogrimo is relying in this case on the results of the GRAAP working group of the Global Grid Forum namely WS-Agreement and WS-Agreement Negotiation [18]. As part of these specifications the possibility to negotiate the content of a Service Level Agreement (SLA) is described. In the volatile context of Akogrimo, *e.g.*, where service providers might move from a high- to low bandwidth environment, the acceptance of SLAs can only be done based under certain conditions (*e.g.*, the minimum bandwidth) and through active management of information technology *and* network resources. Within Akogrimo, besides the management of computational and data resources the execution management system (EMS), which received the SLA to be enacted, is also able to benefit from virtualization interfaces provided by the network layer, since it is enabled to select from different sets of QoS bundles. Thus, if a consumer demands a certain SLA this might cause a change in its access network quality class or potentially even in a change of the access network provider.

Akogrimo is approaching SLA and QoS clearly with a concept of bilateral agreements between consumers and providers. While these agreements might be brokered they determine clearly electronic contracts between consumers and providers. This point is emphasized, since many approaches within the Grid domain aim for a global or VO-wide resource optimization strategies, ignoring potentially customer-specific behavior of provider.

D. Workflows, Operational VOs, and Application Services

A common workflow language for web service-based orchestration is the Business Process Execution Language (BPEL). As described above all grid nodes within Akogrimo are potentially subject to context changes and may consequently require a change in the execution of the workflow. For example, if the user changes from a powerful workstation that is capable of rendering complex medical data to a PDA the workflow needs to be adapted, *e.g.*, by adding a rendering service that delivers a small picture of the rendered data together with a set of meta data optimized for the low resolution client. This might involve the discovery, configuration, and integration of a new member of the operational virtual organization (see [23] for details) and a switch

of the current workflow. Furthermore, application services might be forced to change their behavior, *e.g.*, through reducing the frequency of data updates or the provision of additional textual metadata. In general, the dynamic and timely adaptation of services on the application layer is a key issue for the success of a mobile grid, since they show a major impact on the user's experience.

VI. NETWORK MIDDLEWARE VIEW

In the Akogrimo network, mobile users will be roaming across different access networks belonging to different administrative domains. Further, there is no permanent network coverage expected which leads to disconnected operations scenarios in order to maintain membership of an MDVO.

The underlying network and its components are required to provide the necessary infrastructure to deliver services to users as well as to provide supporting services for network and service providers. Figure 3 gives an overview of the Akogrimo network architecture. Users connect to the network via different end-devices, including portable devices and stationary terminals. Access networks provide wired or wireless connections via different physical access technologies. The core network connects the access networks, provides connectivity towards other networks and comprises network components that provide network and grid services. The grid middleware and the grid application infrastructure provide grid services. Figure 3 does not show network domain boundaries but the depicted components support inter-domain communication and can be deployed in several domains. If multiple network providers are present, handovers among the different networks are supported, assuming that it is underlined by a commercial agreement.
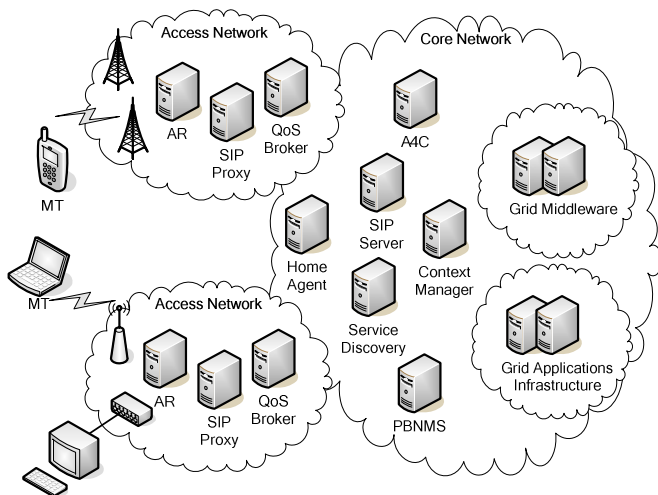

Figure 3: Akogrimo Network Architecture

Selected parts of the Akogrimo network architecture are described in the following sections in detail.

A. Mobility and QoS support

One of the main challenges of Akogrimo is providing seamless access to grid services and applications in a mobile environment. The following three different types of mobility are identified and supported [22]:

- *Device mobility*: In a mobile communication environment users are allowed to move throughout different networks having access to the same set of services.
- *User mobility*: Users of a mobile grid should have access to their services regardless of the device they use or the place and technology they use to access these services.
- *Session mobility*: Session mobility refers to seamless transfer of an ongoing service session from a device to another.

Device mobility is provided by mobile IPv6 extended with fast handover support [21]. The home agent (HA) in the core network serves as a location register for mobile terminals. User and session mobility are supported by the overall authentication process integrated in the network and the SIP infrastructure. Besides mobility the Akogrimo network infrastructure provides also QoS support. According to Figure 3 access routers (AR) in the access network provide the interface for end-terminals to connect to the network. The AR performs access control to network resources based on user authentication and meters network service usage. It receives QoS requests from the mobile terminal, controls QoS signaling and communicates with the QoS Broker. The QoS Broker is the central entity to manage QoS within the network. It decides on QoS requests, manages available network resources and exchanges these information with other QoS Brokers. Mobility and resource management are integrated in order to support QoS management during handovers and to ensure that handovers between different access networks are possible.

B. SIP Infrastructure

The integrated SIP-based communication is provided by the SIP infrastructure, supporting a common session management and session mobility. The central part of the infrastructure is the SIP server in the core network, which contains the SIP proxy and the SIP registrar for user registration (cf. Figure 3). Separate SIP proxies can be deployed in several access networks to distribute the load of the SIP server. The SIP server also provides presence information of the user to the context manager which collects and distributes these information to other network components.

C. A4C

In Akogrimo, the A4C system provides the necessary functionality for authenticating users, authorizing access to services, accounting and charging for service usage, based on agreed SLAs, as well as auditing. By its heterogeneous nature, in terms of services and service providers (SP) involved, a VO needs strong support from the A4C infrastructure for managing inter-domain security and inter-domain accounting and charging. Emerging from a network technology, the A4C will deliver AAA services not only to network services, but also to grid-based services.

A mobile grid infrastructure aggregates services from multiple SPs in order to build new applications to be delivered to end-users. The new applications will require QoS support across the different domains and will need to be

charged according to the quality experienced by the service customers. Such aggregations require trust agreements between SPs (for simplicity network providers are also seen as SPs) at the A4C infrastructure level. Trust agreements are similar to the roaming agreements between mobile operators and allow a user to consume services from multiple SPs while keeping a business relationship with a single SP - his home domain. A user usually has an initial contact with his home domain represented by signing a contract between the two parties, and then a permanent security relationship between them. The trust relationship between the user and his home domain and the trust relationships of that domain and other SPs enable the user to establish temporary, transitive trust relationships with new SPs thus allowing the user to access services provided by other SPs. The newly created trust relationship however does not imply any business relationship between the user and the new SP.

Challenges in designing and implementing the A4C system arise from the need to extend mechanisms developed initially for network services in order to deliver their functionalities to mobile grid services. Akogrimo's A4C infrastructure is based on the Diameter protocol and it follows the IETF AAA generic architecture. The main components of A4C are the A4C server, the A4C client and the SAML authority. Their physical deployment in an existing network and grid infrastructure is depicted in Figure 4.

The A4C server performs authentication of users, authorization of network services, accounting of service usage and charging of service consumption. Besides these tasks, it also handles the inter-domain communication with A4C servers of other SPs based on existing cross-domain trust agreements.
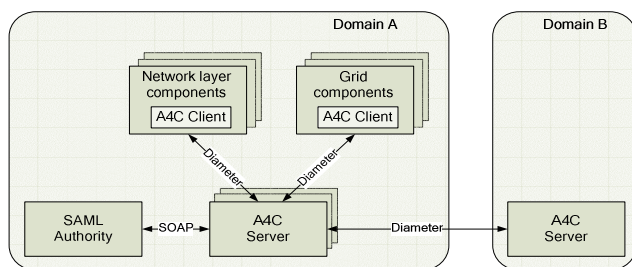

Figure 4: A4C View of the Network Architecture

The SAML authority is a component built based on [25] and is used by the A4C server for providing single-sign-on support across multiple domains. Its task is to generate and check SAML assertions during user authentication sessions.

The A4C client is a module used by any network or grid component that requires A4C functionalities. Its main task is to establish and manage communication with the A4C server in order to provide access to the A4C infrastructure to any component that needs it.

The following paragraphs shortly describe the tasks performed by the A4C system.

1) Authentication:

Two different authentication mechanisms are used in Akogrimo. Network authentication is the first one that is used when a user attaches to an access network (AN). During the network authentication a user is identified based on a username and a password. If the AN belongs to the home SP the authentication is processed by the local A4C server, otherwise the local A4C server will redirect the request to the home domain of the user for further processing. As result of a successful authentication the user receives a SAML artifact (IDToken) which will be used for further service request as a proof of being an authenticated user. An IDToken is generated by the SAML authority from the user's home domain during the authentication process, and it can only be validated by the SAML authority that generated it. The validation of IDTokens is the second authentication mechanism used in Akogrimo.

2) Authorization

In Akogrimo, service authorization has two distinct aspects: on the one hand there is network services authorization which allows the user to have network connectivity and on the other hand there is a more complex authorization mechanism that is used inside a VO for enabling the interoperation of different services belonging to different SPs. The A4C server will perform network services authorization, while the authorizations needed for the VO internal operations will be performed by VO management components. The Akogrimo components that will make use of the authorization mechanisms of the A4C server are the QoS broker and the VO manager. Whenever a user attaches to an access network, the QoS broker responsible for that access network will request from the A4C details on the QoS that needs to be de1livered to the user. The VO manager will contact the A4C server whenever a user will request to join a VO. The VO manager requests the generic user profile (GUP) of a user. Based on a user profile, further authorization decisions will be made at the VO level.

3) Accounting

The A4C server controls and manages the accounting process. It collects and stores accounting records, which describe the usage data for a certain service session. An accounting session is related to the service usage of a user. Service usage data is gathered on different components by metering network traffic, grid resources utilization or other user activities. The components which perform metering and generate accounting data are access router (AR) and the grid monitoring component. A commercial mobile grid platform for being successful is required to offer different levels of QoS and charge its users based on the QoS they receive during service provisioning. The A4C system performs this task by specifying for every service metrics that define the quality of that service as well as the degree of the service usage. The acceptable values for these metrics are part of an SLA agreement and are continuously monitored and accounted for during the service session.

4) Auditing

Services in Akogrimo are monitored for the parameters negotiated in the SLA. The auditing component gathers information on the SLA violations occurring during service consumption. The data recorded during auditing processes will be used later on during the charging process.

5) Charging

Based on the accounting records collected during accounting process and the charging specification stored in the A4C server, the charging subsystem calculates the charge for each service session and creates associated charging records. A charging record contains the result of the charge calculation and it includes attributes like the username, accessed service, domain where the service was provided, currency used, and the charge for the service. To support service provisioning in a multi-domain environment where users receive a single bill from their home provider, the exchange of charging records between providers is required. Charging records are used both for the transfer of end user charges and for the exchange of settlement data between providers.

D. Context Management

The Akogrimo context manager gathers basic context data, processes and refines it according to common Akogrimo context requirements and allows context consumer to add extension modules to handle domain-specific context inference. The overall Akogrimo context management approach is shown in Figure 5.

The context manager gathers raw context data from different context data sources across well-defined interfaces. Given the variety of possible context data sources, the context manager needs to handle multiple protocols for collecting data from different sources. The context manager has a set of rules (policies) describing the logic for context gathering (*e.g.*, which sources provide context, at what schedule, action to take when something fails). These rules are configured through an administrative interface. The context manager refines raw context data to provide standard formats (*e.g.*, geographical coordinates, temperatures), performs processing according to common Akogrimo requirements and stores the context data in a context database.
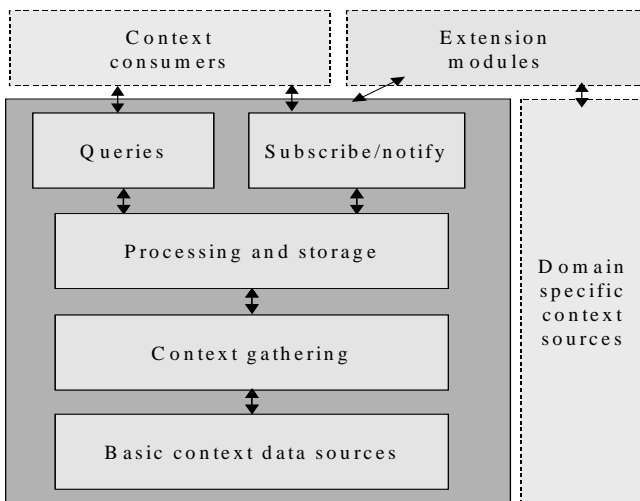

Figure 5: Context Management Functional Layers

Context consumers can either execute queries towards the context manager to get specific context information or subscribe for notifications about changes in certain context data. The context manager is responsible for distributing context data to context consumers accordingly. To handle domain-specific context inference, the context manager offers interested parties interfacing with extension modules. One example is the translation of location coordinates to a map of a building, such that the context will also specify which room a person is in and what physical facilities he is able to use. Context extension modules also allow context consumers to combine basic context information with additional domain-specific context sources, e.g. data from specific medical sensors.

E. Service Discovery

The functional role of the service discovery service (SDS) is to allow a service requestor to find an appropriate service provided by a service provider. A service provider must describe the service capabilities (service description, SD) and publish the service through the service discovery service. The SDS maintains a list of available services and their capabilities. Since within the Akogrimo architecture very diverse services must be handled, a unified service description is a key issue. The different types of service discovery protocols target different types of services, such as people, places, or web services. In order to embrace the greatest number of SD types several protocols will be used, which may encounter service directory protocols and their access protocols.

Within the Akogrimo environment, the general service discovery system (GSDS) is in charge of supplying a particular service from a particular client request. GSDSes have to deal with services of very different nature. On the one hand, with the so-called local services/resources such as printers, beamers, virtual hard disks, and on the other hand with the grid services that can be defined as stateful web services, which are suited for the semantic grid.

The system in charge of discovering local resources (services) is called local service discovery system (LSDS), whereas grid service discovery system (GrSDS) is the name for the module dealing with proper Grid Services as presented in Figure 6.
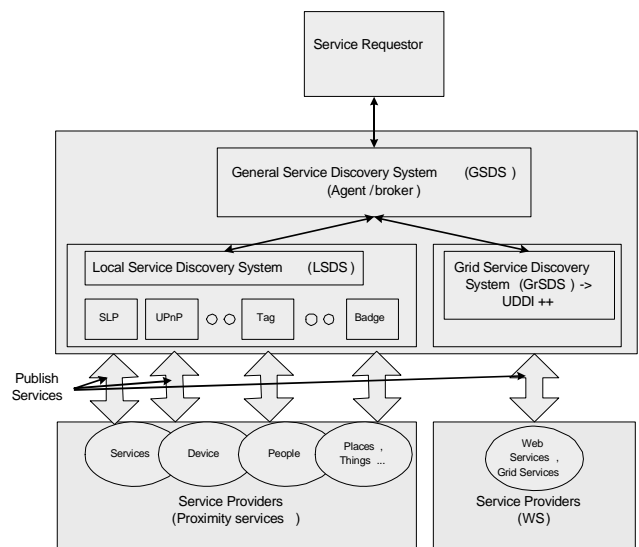

Figure 6: Service Discovery

VII. Summary and Conclusion

In this overview paper, a new approach on grid business models – mainly driven by the Akogrimo context, however, taking general requirements into account – has been developed in which the detailed list of classification criteria as well as key characteristics have been exploited. Based on an extensive related work discussion, the Akogrimo Architecture has been outlined by addressing business, grid, and network middleware views. Driven by a three layer model, the network services layer, the network middleware services layer, and the grid infrastructure services layer, the grid application support in terms of virtual organization (VO) memberships, workflow enhancements, and Service Level Agreements (SLA) can be granted. This approach results in a highly flexible system, which addresses modern mobility requirements at the same time.

To show the integrated understanding of networking and grid aspects, three different views have been introduced. While the business view of Akogrimo focuses on economic aspects of commercial mobile grids as such, the grid view determines the services, infrastructure, and technology available for now. Finally, the network view provides insights into the supportive distributed systems, which provide mobility and Quality-of-Service (QoS) support, signaling support, and authentication, authorization, accounting, auditing, and charging (A4C) support, context management support, and service discovery support. Thus, these elements form the base for a single technology solution, in which networking and grid approaches are harmonized in the sense of a common A4C system and joint signaling system.

As the project Akogrimo is being implemented in its current state, the architectural design steps undertaken have resulted in an integrated view of grids and networks. This integration is believed to leverage off the effort seen so far in pure grid approaches, which do not address explicitly the potential knowledge of underlying networking technologies. Neither the support of roaming users nor the integrated provisioning of full-fledged and automated accounting schemes have been visible so far in existing systems. Thus, Akogrimo provides a unique solution for accessing knowledge through the mobile link.

In the near future there will be a prototype available, which addresses key areas of the grid and networking integration. Driven by implementation restrictions and at the same time standardization requirements, a SOAP and diameter-based system will form the basis for experimental work and evaluations.

Acknowledgements

References

[1] D. X. Adamopoulos, C. A. Papandreou: *Enabling Web Services: Towards Service Grids*. Ninth International Symposium on Computers and Communications 2004 (ISCC 2004), Vol. 1, June & July 2004, pp 38-43.
[2] Akogrimo: *Access to Knowledge through the Grid in a Mobile World*. European Union project FP6-2004-IST-2-004293, http://www.mobilegrids.org, October 2005.
[3] Y. Bernet, P. Ford, R. Yavatkar, F. Baker, L. Zhang, M. Speer, R. Braden, B. Davie, J. Wroclawski, E. Felstaine: *A Framework for Integrated Services Operation over Diffserv Networks*. IETF RFC 2998, November 2000.
[4] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, W. Weiss: *An Architecture for Differentiated Services*. IETF RFC 2475, December 1998.
[5] P. Calhoun, J. Loughney, E. Guttman, G. Zorn, J. Arkko: *Diameter Base Protocol*. IETF RFC 3588, September 2003.
[6] P. Calhoun, G. Zorn, D. Spence, D. Mitton: *Diameter Network Access Server Application*. IETF RFC 4005, August 2005.
[7] M. Cannataro, D. Talia: *Semantics and Knowledge Grid: Building the Next-generation Grid*. IEEE Intelligent Systems, Vol. 19, No. 1, January-February 2004, pp 56-63.
[8] R. Cao, Y. Chai: *The Study of the Electronic Commerce Business Models*. 2004 International Conference on Systems, Man and Cybernetics (SMC 2004), Vol. 5, October 2004, pp 4089-4094.
[9] C. de Laat, G. Gross, L. Gommans: *Generic AAA Architecture*. IETF RFC 2903, August 2000.
[10] T. Dierks, C. Allen: *The TLS Protocol Version 1.0*. IETF RFC 2246, January 1999.
[11] M. Dubosson-Torbay, A. Osterwalder, Y. Pigneur: *E-Business Model Design, Classification, and Measurements*. Thunderbird International Business Review, Vol. 44, No. 1, January-February 2002, pp 5-23.
[12] I. Foster: *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*. First IEEE/ACM International Symposium on Cluster Computing and the Grid, May 2001, pp 6-7.
[13] I. Foster: *What is the Grid? A Three Point Checklist*. GridToday, Vol. 1, No. 6, June 2002, http://wwwfp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf.
[14] I. Foster, C. Kesselman: *The GRID: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers Inc., 1999.
[15] I. Foster, C. Kesselman, J. M. Nick, S. Tuecke: *The Physiology of the Grid. Grid Computing*. Making the Global Infrastructure a Reality, John Wiley & Sons, 2003, pp 217-249.
[16] The Globus Alliance: *The Globus Toolkit*. http://www.globus.org/toolkit/, March 2006.
[17] The Globus Alliance: *The WS-Resource Framework*. http://www.globus.org/wsrf/specs/ws-wsrf.pdf, May 2004.
[18] Global Grid Forum (GGF): *GGF-GRAAP Working Group Documents*. http://forge.gridforum.org/docman2/ViewCategory.php?group_id=71&category_id=659, March 2006.
[19] D. Johnson, C. Perkins, J. Arkko: *Mobility Support in IPv6*. IETF RFC 3775, June 2004.
[20] S. Kent, R. Atkinson: *Security Architecture for the Internet Protocol*. IETF RFC 2401, November 1998.
[21] R. Koodli: *Fast Handovers for Mobile IPv6*, IETF RFC 4068, http://www.ietf.org/rfc/rfc4068.txt, July 2005.
[22] A. Litke, D. Skoutas, T. Varvrigou: *Mobile Grid Computing: Changes and Challenges of Resource Management in a Mobile Grid Environment*. 5th International Conference on Practical Aspects of Knowledge Management (PAKM 2004), December 2004.
[23] C. Loos, S. Wesner, J. Jähnert: *Specific Challenges of Mobile Dynamic Virtual Organizations*. eAdoption and The Knowledge Economy: Issues, Applications, Case Studies, 2005.
[24] A. Olivereau, A.F. Gomez Skarmeta, R. Marin Lopez, B. Weyl, P. Brandao, P. Mishra, H. Ziemek, C. Hauser: *An Advanced Authorization Framework for IP-based B3G Systems*. 14th IST Mobile & Wireless Communications Summit, Dresden, June 2005.
[25] Organization for the Advancement of Structured Information Standards (OASIS): *Security Assertion Markup Language, SAML*, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security, October 2005.

[26] Organization for the Advancement of Structured Information Standards (OASIS): *Web Services Base Notification 1.2 (WS-BaseNotification)*. Working Draft 03, http://docs.oasis-open.org/wsn/2004/06/wsn-WS-BaseNotification-1.2-draft-03.pdf, June 2004.

[27] C. Rensing, H. Hasan, M. Karsten, B. Stiller: *AAA: A Survey and a Policy-based Architecture and Framework*. IEEE Network Magazine, Vol. 16, No. 6, November 2002, pp 22-27.

[28] C. Rigney: *RADIUS Accounting*. IETF RFC 2866, June 2000.

[29] C. Rigney, S. Willens, A. Rubens, W. Simpson: *Remote Authentication Dial In User Service (RADIUS)*. IETF RFC 2865, June 2000.

[30] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler: *SIP: Session Initiation Protocol*. IETF RFC 3261, June 2002.

[31] L. Schubert, S. Wesner, T. Dimitrakos: *Secure and Dynamic Virtual Organizations for Business*. eAdoption and The Knowledge Economy: Issues, Applications, Case Studies, 2005.

[32] P. Timmers: *Business Models for Electronic Markets*. Journal on Electronic Markets, Vol. 8, No. 2, April 1998, pp 3-8.

[33] M. Waldburger, B. Stiller: *Toward the Mobile Grid: Service Provisioning in a Mobile Dynamic Virtual Organization*. IFI Technical Report 2005.07, ftp://ftp.ifi.unizh.ch/pub/techreports/TR-2005/ifi-2005.07.pdf, August 2005.

[34] S. Wesner: *Towards and Architecture for the Mobile Grid*. it-information Technology, Vol. 47, No. 6, June 2005, pp 336-342.

[35] World Wide Web Consortium (W3C): *Web Services Addressing (WS-Addressing)*. http://www.w3.org/Submission/ws-addressing/, August 2004.

[36] WSRF.NET Forum: *WSRF.NET*. http://www.cs.virginia.edu/~gsw2c/wsrf.net.html, March 2006.