Project no. FP6-033636

NESSI-Grid

Networked European Software and Services Initiative - Grid

Instrument: Specific Support Action (SSA)
FP6 – Priority IST-2.5.4 Advanced grid technologies, systems and services

Deliverables D.1.2 & D.1.4
Grid Vision and Strategic Research Agenda

Due date of deliverable: 31 October 2007
Actual submission date: 30.11.2007

Start date of project: 1st May 2006 Duration: 30 months

Lead contractor for this deliverable: SAP AG
Revision: V.2.0 (30th November 2007)

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)	
Dissemination Level	
PU Public	Yes

Keywords:
Business Grids, SOKU, Agility, Adaptive IT, Virtualisation, SLA enforcement, NESSI, SRA

Document History

Version	Date	Features
0.9	September 2006	Early draft for SRA workshop in Brussels
1.0	October 2006	First official SRA version • general vision • three short term scenarios • high-level state-of-the art • challenges based on short-term scenarios
1.2	May 2007	Intermediate version for the Community for the SOI-NWG workshop
1.5	July 2007	Draft for second official SRA version • revised introduction (sharpened context; added stakeholder description) • extended/revised scenario (added mid- to long-term scenarios; refined derivation from vision; refined impact description) • refined state-of-the art (refined structure; extended scope) • refined challenges (extended according to full set of scenarios; added technical drivers; added correlation with requirements) • added description of community involvement process • added business indicators section
1.9	November 2007	Final draft of Second Official SRA Version • Revised introduction and executive summary. • New section (1.1) with the relationship between NESSI, NESSI-Grid, SOI-NWG, TG1, etc. • New section (2.2) with requirements for business applications. • Revised SOTA section. • Revised Challenges section. • New section (6) with analysis of results from the call from contributions to NESSI-Grid SRA. • Revised business & market indicators section. • Revised conclusions. • New and revised annexes on community contribution.
2.0	November 2007	Second official SRA version

Table of content

Executive Summary

The Networked European Software and Services Initiative (NESSI) aims to create a strategic research agenda for European research in services and their foundations. NESSI-Grid supports the Service Oriented Infrastructure (SOI) NESSI Working Group (NWG) in its task of defining a vision and strategic research agenda (NESSI-Grid SRA) for grid-like infrastructures used in business environments and in particular in NESSI scenarios. A second version of both the NESSI-Grid vision and SRA, including extensive community feedback, is presented in this document. It constitutes the main contribution made by the SOI NWG to the NESSI overall vision and SRA. This document addresses all stakeholders in this field, in particular researchers, industry and policy makers, to get a precise understanding about envisioned scenarios, state-of-the-art, challenges, business impact and actual roadmaps and recommendations for realizing the vision.

The Driver

The main driver behind this SRA is the fact that businesses in future service-oriented economies need to act in a more agile fashion than ever before. This will be possible only if high-level business requirements can be translated into lower level ICT requirements with a high level of automation, so that ultimately the ICT environment adapts automatically to changing business needs.

The Vision

We envision ***Business Grids as the adaptive service-oriented utility infrastructure for business applications***. They will become the general ICT backbone in future economies, thus achieving a profound economic impact. The adoption of Business Grids is expected to happen in 3 steps: first as the ICT backbone within enterprises, second as a basis for hosting scenarios and ultimately as the general ICT-infrastructure for service-oriented economies. In this way, Business Grids will eventually support the emergence of new types of application.

Business Scenarios

The industry-driven approach of this SRA is evident in its identification of the most relevant core business scenarios and associated business requirements. The scenarios cover short-term perspectives (basic enterprise, hierarchical enterprise, hosting) but also mid- to long-term setups (extended enterprise, merger & acquisition, virtual organizations, dynamic outsourcing, value chains and mega services).

Research Challenges

Research challenges for realizing the vision of Business Grids are systematically derived from the business scenarios, technology trends and state of the art. They flow together in three key challenges:

KC 1: **New system architectures** that harmonize service architectures (SOA) and infrastructure architectures (SOI), advance the structure of multi-tier, federated and Internet scale architectures, support all kinds of business models, applications and emerging hardware environments and provide transparent and integrated

access for all relevant stakeholders (architects, engineers, operators, customers, …).

KC 2: **Advanced system lifecycle approaches** including engineering, deployment, composition, provisioning, management and decommissioning phases that support transparent knowledge tracking, feedback loops, prediction and simulation, allow for a clear separation of concerns between different stakeholders (business vs. IT, developers, providers, customers, …) and support the full variety of business scenarios (from traditional data centres to complex service value networks) while adhering to overarching sustainability requirements.

KC 3: **Advanced infrastructure technologies** in terms of hardware (energy efficient, flexible allocation, virtualization …), middleware (new multi-tier system design, flexible storage systems, harmonized virtualization on all layers) and related programming models (parallel programming, multi-core) that meet the required flexibility of the networked economy.

1 Introduction

The Networked European Software and Services Initiative (NESSI) aims to create a strategic research agenda for European research in services and their foundations. NESSI-Grid supports the Service Oriented Infrastructure (SOI) NWG in its task of defining a vision and strategic research agenda (SRA) for service-oriented infrastructures used in business environments and, in particular, in NESSI scenarios. The anticipated service-oriented infrastructures are supposed to leverage traditional Grid infrastructures, which ultimately aim at providing resources as a utility, to general purpose applications and business scenarios. The term *Business Grids* has been adopted to describe those infrastructures.

This document presents the second version of both, a business-driven vision for grid-like infrastructures as well as a first approach to a strategic research agenda and its related research challenges. Rather than describing vision and SRA in two separate documents, we decided to combine both deliverables into one document as they have a significant overlap. The vision as such is explained on a high level in the first section but also refined through various business scenarios, presented in Section 3. On the other hand, the business requirements used to characterize those scenarios in more detail as well as the order of the scenarios is already a first step towards the SRA. An account of the state of the art from a scientific point of view has been given in Section 4. The SRA is then refined by concrete research challenges in Section 5. Section 6 presents the analysis of the results of the community involvement process conducted to gather input the wider grid community and giving the timeline for the next phases of this process. Section 7 introduces an approach to characterize relevant business and market indicators. Section 8 provides a short conclusion.

This document is addressing all kinds of stakeholders, in particular industry, researchers and policy makers. *Industry stakeholders* will get a vision of future IT infrastructures, their related business scenarios, current technologies/challenges with respect to those scenarios as well as the specific business impact. *Researchers* get guidance for industry-relevant research topics. *Policy makers* get an understanding of possible areas of influence and anticipated market impact.

1.1 Context

Launched in September 2005 by 13 partners and enlarged in June 2006 to 22 partners and over 200 members, NESSI aims to provide a unified view for European research in Services Architectures and Software Infrastructures that will define technologies, strategies and deployment policies fostering new, open, industrial solutions and societal applications that enhance the safety, security and well-being of citizens.

The Strategic Research Agenda (SRA) is the tool through with a European Technology Platform such as NESSI details its strategy, its challenges and how it intends to meet them. The role of the SRA is to translate the NESSI vision into action and to guide NESSI's implementations during the entire lifetime of NESSI.

The NESSI SRA is coordinated by the NESSI SRA Committee, which was formed by NESSI's Steering Committee. The Committee's mission is to define and update the NESSI SRA. Like all NESSI Committees, it is an extension of the Steering Committee and in particular is open only to NESSI partners. The specific missions of the NESSI SRA Committee are:

- to be the guardian of the integrity of NESSI within the scope of the NESSI SRA
- to define the NESSI skeleton and principles
- to redact and maintain the NESSI SRA based on the input from the NESSI Working Groups
- to trigger the creation of horizontal and vertical NESSI Working Groups
- to implement communications that are related to the NESSI SRA

The content of the NESSI SRA is elaborated by both the partners and members of NESSI. NESSI Working Groups (NWGs) constitute the core elements which contribute to the content of the NESSI SRA. NESSI Working Groups are further classified in Horizontal NWGs, Vertical NWGs and the SME NWG. Horizontal NWGs are primarily intended to provide input on the various technological research areas in scope for the NESSI Strategic Research Agenda. Vertical NWGs are primarily intended to link the business application domains (e-Health, e-Government, …) with the NESSI Strategic Research Agenda. Last but not least, the SME Working Group is intended to foster active participation by the ICT SMEs to NESSI and to ensure that the issues related to SMEs are taken into account.

Besides contributing to the NESSI SRA, Horizontal NWGs are aimed to:

- Promote the creation of NESSI communities interested in a given research topic
- Link NESSI with the relevant research initiatives through common members
- Disseminate NESSI concepts within their research community
- Participate in related activities, such as the creation of standards, in coordination with the relevant NESSI Committees

NESSI-Soft is a Specific Support Action (SSA) launched at FP6 whose objective is to support the work of NESSI at global level. Particularly, NESSI-Soft is responsible for:

- The setting up and operation of the NESSI-Office, whose goals are to provide NESSI (and NWG in particular) with administrative, logistics, secretarial and organisational support.
- Providing tools to the NESSI Board and Steering Committee to monitor the evolution and coordination of NESSI activities.
- Supporting activities of the NESSI SRA Committee as defined previously

The Service Oriented Infrastructure (SOI) NWG is one of the Horizontal NWG and is focused on the new generation of ITC infrastructures that will support development and execution of component services and their provision as utilities.

The NESSI-Grid project is a Specific Support Action (SSA) launched at FP6 to support activities of the SOI NWG, with a special focus on providing the necessary support to structure and develop its contribution to the NESSI SRA in the specific area of ITC infrastructures and Grids. This contribution is materialized in the present NESSI-Grid SRA document.

In addition, NESSI-Grid provides support to activities of the Technical Group 1 (TG1) defined by the EC to coordinate debate and joint collaboration among Grid-related FP5, FP6 and FP7 EU Projects on architectural challenges which have to be addressed in the transition of the Now-Grid to the next-generation Grids. Providing the basis for this debate and joint collaboration, NESSI-Grid coordinates the development of a series of whitepapers. Note that the NESSI-Grid SRA is aimed to define a SRA focused on business environments in NESSI scenarios, while activity in TG1 is not devoted to generation of a SRA as such and, on the other hand, is not limited to address business environments in NESSI scenarios.

Besides the above, both NESSI-Soft and NESSI-Grid SSAs provide support for communication and interlinking activities at NESSI level. Both types of activities are aimed to achieve the widest dissemination and adoption of the NESSI Vision.

1.2 Vision

The main driver behind this SRA is the relationship between high-level characteristics of service-oriented economies [1] and supporting Information and Communication Technology (ICT) infrastructures. This is depicted in Figure 1.

Businesses in future service-oriented economies need to act in a more agile fashion. This will be possible only if high-level business requirements can be translated into lower level ICT requirements with a high level of automation, so that ultimately the ICT environment adapts automatically to changing business needs. Figure 1 shows the different stages of this translation process.

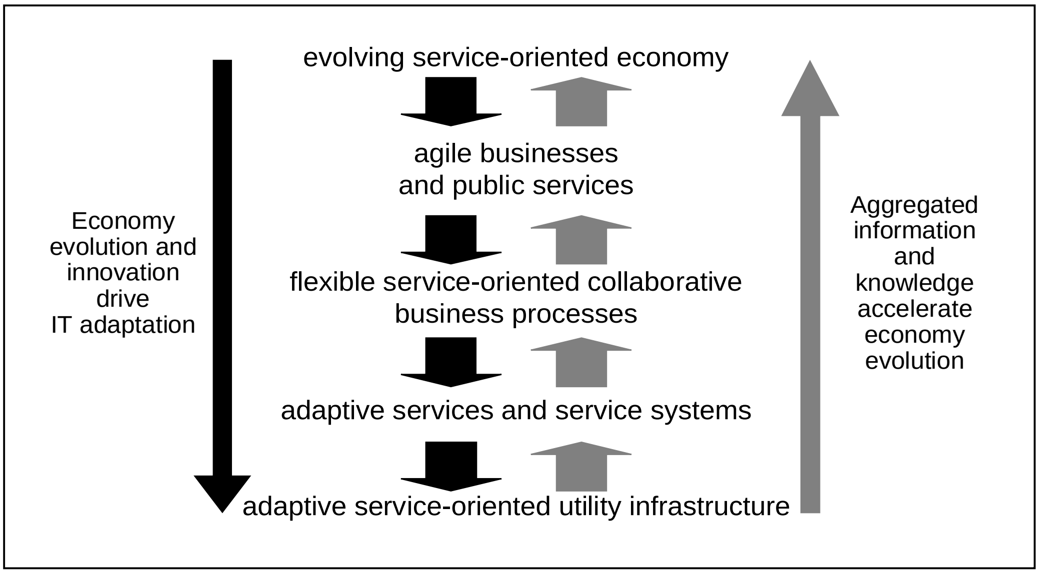


Figure 1: Agile businesses drive adaptive IT

In detail, an evolving service-oriented economy will lead to new requirements in the sense that existing businesses and services must react more quickly to changing circumstances. This means that they must adopt more flexible service-oriented business processes and be able to exploit collaborative relationships. Software technologies enabling the implementation and composition of flexible services will be important in achieving this and are now being seriously addressed by the software industry. This can be seen in the widespread trend towards service-oriented architectures and business processes built upon this architecture. At the infrastructure level, an adaptive service-oriented utility infrastructure will allow the dynamic, on-demand allocation and assembly of resources needed to support service components and assemblies. A crucial prerequisite for this layered adaptability is proper information flow translating higher-level business goals into lower-level technical goals and aggregating lower level characteristics to higher-level capabilities.

We call this envisaged adaptive service-oriented utility infrastructure "***Business Grids***". Business Grids will become the general ICT backbone in future economies, thus achieving a profound economic impact. In this respect they differ significantly from traditional e-Science grids. These focus on specific application areas often aiming to provide shared access to specialised high performance computing resources or datasets. Typically they deal with independently executing, stateless batch jobs which can be easily moved around a network and executed based on some given input files. In contrast to this, Business Grids will provide a general purpose infrastructure for arbitrary business applications, primarily aimed at providing business flexibility efficiently. Typically, business applications do not exist as independent executables but involve a complex technology stack, containing, for example, application servers, databases, other middleware components and co-existing and cooperating business applications. In addition, usage patterns are often interactive or session-based. Transactions are used across all tiers to guarantee data consistency.

We anticipate that Business Grids will become the foundation for ICT infrastructure via the following scenarios:

1. Business Grids will be the ICT backbone for enterprise solutions.
2. Business Grids will support hosting scenarios for small and medium sized enterprises.

Based on these two initial scenarios and as the interworking between administrative domains becomes commonly accepted and is supported as part of the infrastructure this will evolve to the following:

3. Business Grids will provide the ICT infrastructure to support service-oriented economies and eventually support the emergence of new types of applications.

It is, of course, important to note that these developments will require changes in business culture which will undoubtedly influence the timescales for change.

1.3 Technical scope

Figure 2 shows the main ICT layers of existing business solutions.

Applications (e.g. collaborative business processes)
Business logic (e.g. business web services)
Middleware (e.g. application server)
Infrastructure (e.g. OS, hardware)

Figure 2: ICT layers of business solutions.

Broadly speaking, the infrastructure layer is about the provisioning of any kind of hardware resources (compute power, storage, network, sensors, actuators) together with associated low-level software components (e.g. operating systems). Middleware covers generic software functionality such as application servers (e.g. J2EE), databases, messaging services, portal frameworks and also serves for hiding technical complexity (distribution, heterogeneity) from application developers. The business logic layer includes specific business functionality (e.g. processing a purchase order) while the application layer combines business logic in concrete usage scenarios (e.g. a procurement system). It is important to note that these layers do not have well-defined boundaries but are presented as an illustration of different technical areas which are relevant within business solutions.

Business Grids primarily target the infrastructure layer and partially also the middleware layer. They follow the paradigm of a Service Oriented Knowledge Utility (SOKU) [2]. The overall SOKU vision goes far beyond infrastructure problems as it “identifies a flexible, powerful and cost-efficient way of building, operating and evolving IT intensive solutions for use by businesses, science and society”. Applying the SOKU paradigm to Business Grids means that Business Grids offer infrastructure resources to higher levels according to the following main principles:

- service-oriented (i.e. dynamic allocation and assembly of resources via infrastructure services),
- knowledge-assisted (i.e. translating high-level business requirements to infrastructure requirements and infrastructure capabilities vice versa),
- utility (i.e. immediately available, dependable usage, predictable).

Business Grids need to be able to participate in the execution of the business models of the applications they support in an accountable fashion, addressing issues including auditing, billing, and linkage of resource consumption to business goals. In this context, Business Grid solutions will span several ICT layers.

1.4 Non-goals

Following the described scope it becomes clear that research questions related to higher layers of ICT solutions are out of scope for research in Business Grids. In particular this is the case for business models, human aspects, and business services.

- *Business Grids will not address isolated low-level grid business models.* Several trends in hardware, software, and the relative cost of IT operations conspire to make it nearly impossible to construct viable business models just around low-level grid services in isolation from the supported applications[1]. The added value stems from the actual services offered on top of the grid which may include provision of resources as managed services with assured performance levels.

- *Business Grids will not address human aspects* such as political, organizational or cultural issues.
  Business Grids present themselves as a utility; they are invisible to the user and even largely to the developer of applications, business logic, and middleware.
- *Business Grids will not address issues of business service engineering and service provisioning (other than provisioning of infrastructure services).* This will be addressed in other parts of the overall NESSI-SRA.
- *Business Grids* will not address aspects of public networking infrastructures which are typically provided by a network operator.
  Instead they focus on the infrastructure elements that belong to the providers of business services.

1.5 Methodology

Business Grids are about the provisioning of infrastructure resources to business services as infrastructure services. In order to derive a comprehensive and consistent research agenda for Business Grids this SRA adopts a multi-perspective approach by integrating societal trends, business scenarios, technical trends, and technical constraints. A key concept is the *administrative domain*. As resources are always bound to an *administrative domain* (there is always an entity with ultimate authority over the management of a resource), the approach of this document is based on the distinction between the following two cases: provisioning of resources within one domain and provisioning of resources shared between several domains.

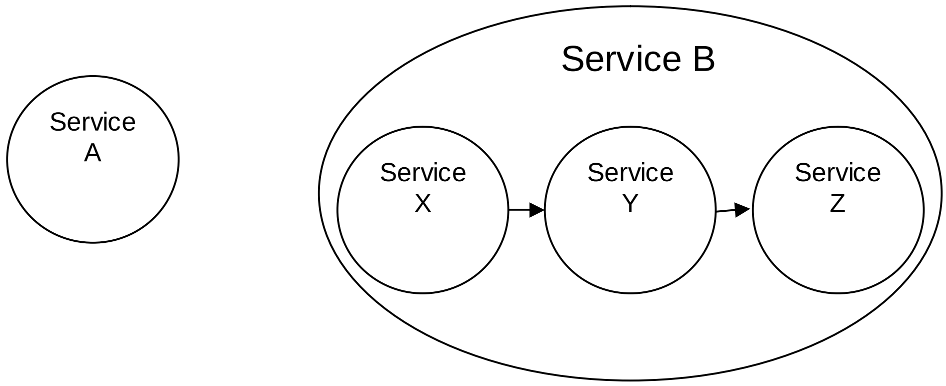
Consequently this document is structured as follows. Section 2 introduces some underlying concepts of this SRA, namely the notion of services, a technical characterization of business applications and a basic model for describing Grid scenarios with respect to administrative domains and some associated basic business characteristics. Section 3 details the most important business scenarios that can be expressed by means of the basic model and specifies their related business requirements. Section 4 reflects the state of the art, both in academia and industry. Section 5 presents the main research challenges derived from the previous analysis. Section 6 summarizes the main results of the community involvement process followed by Section 7 which details the approach for describing relevant business and market indicators. Section 8 ends with a brief conclusion and identification of overarching key challenges.

[1] Compute cycle, network and storage costs will continue to decrease for the next few years. Purchased network capacity, software and higher-level service costs, including infrastructure management will remain relatively unchanged. Thus the share of pure hardware costs in overall IT operation costs becomes more and more irrelevant, inhibiting companies from basing their business models on commodity resource provisioning (argument taken from [3]).

2 Basic Model

2.1 Services

At the core of the NESSI vision are services. Consider the diagram below:



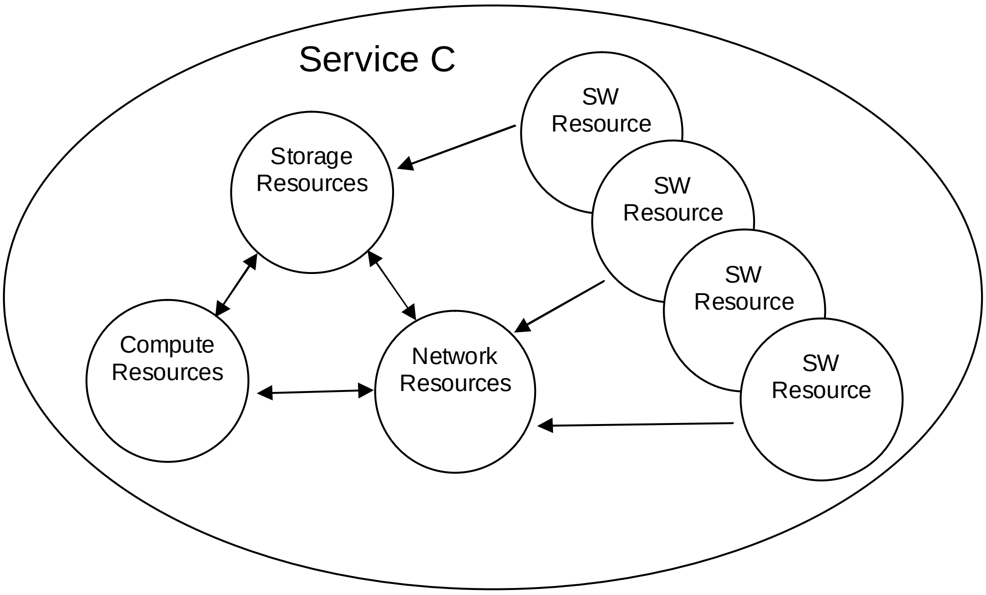
A service provides a given capability. It may be "simple"; e.g. service A, in that it provides all of the required functionality contained within itself. Alternatively, a service may be "complex"; e.g. service B, in that it may itself be made up from other services combined in some fashion. Note that the diagrams above are not intended to show the "ownership" of any given entity. It is perfectly possible that Services X, Y, Z and B are sourced from different organisations.

Services are provided and consumed. This is the case whether they are "simple" or "complex". Relationships between organisations can be expressed in terms of this provision or consumption of services.

- A service may be provided to one organisation for its own consumption.
- A service may be provided to an organisation which consumes it to in turn provide another (complex) service.
    - Services thus provided may be provided "as-is"
    - Services thus provided may be augmented or extended in some (value-added) fashion

Services may have explicit or implicit relationships with one another. These relationships may in turn be expressed or maintained by other services (e.g. by a brokering service, a lookup service, an identification service etc.). Relationships may be static or dynamic. These service relationships allow common business models to be supported.

A service (whether simple or complex) implicitly uses resources to deliver its capability. An example of this is shown in the diagram below:



Service C uses a number of underlying hardware (e.g. compute, storage, networking, etc.) and software (e.g. operating system, applications, licences, etc.) resources. Note that there may also be other types of resources used by some services. These resources are implicitly inter-related. As in the case of the services described earlier, there is no implied “ownership” of any resource in this diagram. An organisation may use resources provided by another to deliver a service. Resources, as with services, may be provided or consumed.

By treating these underlying resources as just another case of a service, then we can see that we can use the same models (implicit / explicit and dynamic / static relationships) to describe the relationships of both the underlying resources to one another as well as the relationships of these resources to the services they support. These relationships between services as well as those between services and their underlying resources need to be well integrated and optimised. Whilst this simple model implies that the relationships between different services and the relationships between services and resources may be expressed in a similar fashion, this does not imply that they can be implemented in a similar fashion. Currently, relationships between different services lack the efficiencies and optimisations that are found between resources and the executable software they support.

2.2 Business applications

In this section we describe the nature of business applications and which are their requirements. Charactering business applications and their requirements is critical since the combination of business scenarios and business applications requirements will define the context of the challenges to be overcome by the future business grids. Additionally, business applications have a number of requirements currently not addressed by grid technologies what will be key in identifying the gap between future business grids and traditional grid technology.

Online Processing

Business applications consist of a mix of online and batch applications. A large fraction of enterprise applications are inherently interactive and therefore require to be online. Online means that requests should be executed in a timely fashion and the client should be provided with short response times. A sharp evidence of this requirement of business applications are industrial benchmarks for service oriented infrastructure. For databases, the main benchmark organization is Transaction Performance Processing Council (TPC). All but one TPC benchmarks are for online applications such as online transaction processing (OLTP), like TPC-C and the new TPC-E, TPC-App for online application server and web service applications, and TPC-W for web applications. Only the TPC-H benchmark for decision support (evolution of online analytical processing, OLAP, and data warehousing) has less strict online requirements due to the large length of the queries run on these systems. Another example of the online requirements comes from the multi-tier enterprise applications based on J2EE for which the current benchmark is jAppServer from the Standard Performance Evaluation Corporation (SPEC). This benchmark focuses exclusively on online applications.

All these standards have tight requirements on the response time and throughput of the applications. The response time requirements are typically set with strict constraints on its statistical distribution (e.g. 90% of the requests should have a response time below 2 seconds and the rest of the requests cannot have an average response time higher than 10%). The fact that all industrial benchmarks focus on online applications is first, because online is considered a crucial requirement of business applications, and second, because online applications are the ones that pose most significant technological challenges and enable to distinguish among the best commercial products.

Stateful Nature and Consistency

Another important feature distinguishing a large fraction of business applications is related to the fact that business applications are stateful since on the one hand, they manage persistent data, and on the other hand, many of the interactions are conversational what requires manipulating session state. Statefulness is an important requirement since it poses important challenges for service oriented infrastructure. Session data restricts which sites can process requests for a particular session. Persistent data is many times shared what is challenging in itself, since wherever requests are processed they need to access the shared persistent data in a timely manner.

Persistent data typically contains vital business information. This means that is consistency is also a vital requirement. This consistency requirement is being satisfied by most business applications by means of transactions. Transactions provide atomicity from two different perspectives. Firstly, they provide failure atomicity. This guarantees that in the advent of failures, either a transaction is completely successful or if the failure happens during its execution the final effect is as it was not executed. This semantics prevents the corruption of persistent data when failures happen. Secondly, it also provides a simplified application programming model since application programmers do not have to deal with concurrency issues. Applications have the illusion of interacting in isolation with the system. That is, transactional systems guarantee that a concurrent execution of transactions in the system is equivalent to a serial execution what tremendously simplifies the programming of applications. This is in fact the reason why most service oriented infrastructures include transactional semantics as part of their support (databases, application servers, web services, enterprise application buses, etc.).

Multi-Tier Service Architectures

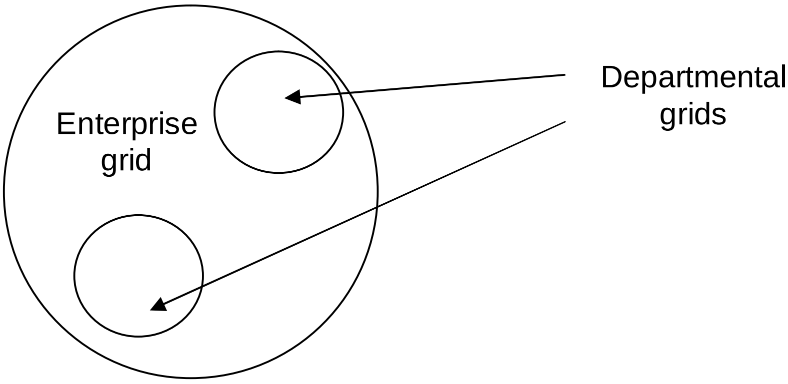
Modern business applications are based on SOA. Typically, they rely on multi-tier frameworks as, for instance, J2EE. Multi-tier architectures contain specialized containers for different purposes, web server for web the presentation tier, application server, for the business logic, and the database tier for persistent data. The fact that most business applications rely on these architectures sets a requirement for business grids to support them, since re-implementing all applications is not feasible. This requirement also means that business grids should not only understand multi-tier architectures, but also should take care of gridifying the different tiers. This implies that consistency should be preserved in the interaction between the different tiers, even in the advent of failures. Consistency involves exactly-once and transactional semantics. Transactional semantics was introduced in the previous requirement. Exactly-once semantics states that a request should be executed exactly once in the advent of failures. That is, in the advent of a failure it should be replayed to guarantee at-least once semantics. But there should be provisions to prevent duplicate request processing in order to also guarantee at-most once semantics.

2.3 Enterprise grids

An “enterprise grid” is the starting point for this discussion. This is a grid within a single enterprise. The main factor identifying an enterprise grid is that it exists within a single administrative domain. An administrative domain in this context is defined by the existence of a common set of management policies. It is likely to also include a specific set of identifiers to refer to relevant entities. The enterprise grid is made up of the lowest level resources: computing, network, storage, and basic services that make these resources available for use.
Noteworthy, the term grid in this and the following 2 subsections does not refer to any existing or envisioned technology/middleware. Instead, it just describes the assembly of resources according to organizational principles and their related management policies.

2.4 Departmental grids

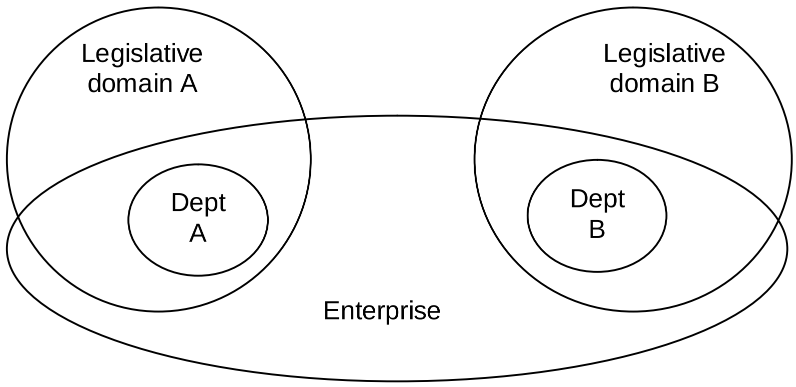
Enterprise grids naturally include those grids implemented across different departments within a wider organisation. A “departmental grid” is implicitly a subset of an enterprise grid. The diagram below shows this relationship:



The enterprise grid can be regarded as a “grid of grids” i.e. a specific instance of a “system of systems”. It could also be regarded as demonstrating some other hierarchical relationship between the departmental grids that make it up. The depth of this relationship will define the capabilities of the enterprise grid and the attributes that it exhibits. A very basic relationship may mean that the ‘enterprise grid’ is nothing more than a loose collection of departmental grids and as such adds little value or capability. At the other end of the spectrum is a more tightly integrated vision that delivers a synergistic value to the entire enterprise through the harnessing of all resources in an enterprise. The ability of a given departmental grid to harness resources within and external to itself will vary, based upon the degree of integration. Some enterprise IT organisations are already setting up internal ‘utilities’ to provide services and resources to other parts of the enterprise. In general, these enterprises have a higher degree of sharing / utilisation than others that are less integrated.

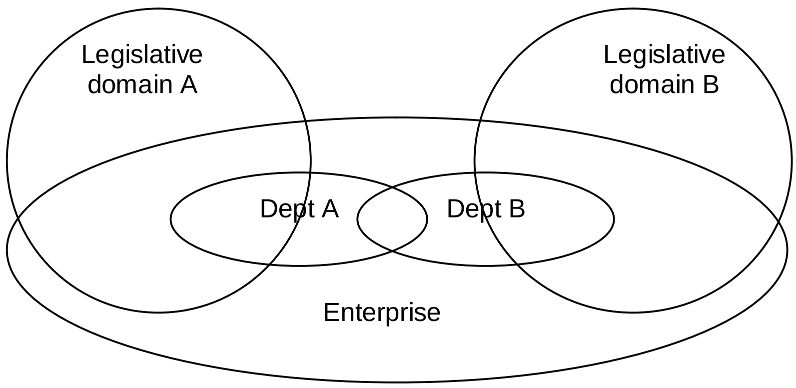
Within an enterprise, resources or services will be owned (in the sense of ultimate authority over usage) by someone. These may be part of an identified ‘grid’ within a given department or might be owned by ‘corporate IT’ in some other fashion. In any case, they ought to be able to be utilised (subject to the appropriate permissions etc.) without necessarily formally being part of a departmental grid. Using the ‘internal utility’ analogy, a resource/service within such a utility can be used by a departmental grid without being part of it, though implicitly if the internal utility is built using grid-like primitives then there is sharing between grids demonstrated here too.

In general, any departmental grids within the same enterprise will come under the same set of enterprise administrative policy rules. For instance, they are both likely to have to conform to similar security or data protection requirements set by the enterprise as a whole. This is not always going to be the case however. Any enterprise that spans multiple countries (and hence multiple legislative domains) may have different local legal requirements (say around the protection of personal information) which need to be complied with. Similarly, some enterprises do not mandate rules to be adopted (or some departments choose to ignore them!). In any case, there are some problems to be solved in terms of what resources can/cannot be used by whom. Consider the following diagram:



In the diagram above, we see two different departments (A and B) within a single enterprise. Each department is subject to a different legislative domain (A & B respectively). It can be seen that if there is NO overlap between the two legislative domains then there is a fundamental limit in terms of what can be shared between them. Note that this does not necessarily preclude the sharing of computational resources, but may well completely prohibit the sharing of data. This will have implications in terms of how isolation between different services and resources can be achieved (and proven to the satisfaction of the respective legislator).

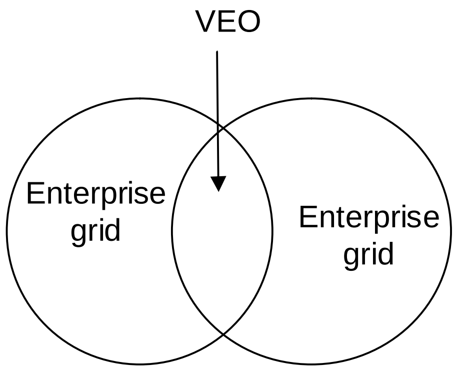
The reality of the situation is usually such that any given enterprise will have a consistent subset of administrative rules that comply with all of the legislative domains within which they operate. This will allow for the sharing of resources without major restrictions thus:



It is not the purpose of this SRA to define mechanisms to overcome shortfalls in enterprise (or departmental) administrative policy setting or in circumventing legislative restrictions. Departments that are inherently forbidden to share resources with others will be unable to do so regardless of any NESSI framework. The NESSI framework should allow such restrictions to be expressed and enforced. However, in those environments where there are consistent legal rules, then the NESSI framework should allow for the appropriate sharing of services and resources according to these rules.

2.5 Virtual Enterprise Organisations

Virtual Enterprise Organisations (VEOs) form when two or more enterprise grids intersect; more correctly when the two administrative domains (and hence their enterprise grids) overlap and share resources. The reality of a VEO is that only a subset of the overall grid within an enterprise is likely to be "contributed" to this virtual organisation. The majority of the systems within a given enterprise grid are likely to remain outside the VEO. The diagram below shows this VEO relationship:



For the most part, contracts in the business world are between two parties. We believe that the abstractions used to describe the interaction of two administrative domains are good enough in general to be a means to describe the interaction of multiple domains (e.g. companies). Consider a tri-party agreement. This may be formed between the parties A, B & C in such a way that individual bi-party relationships describe the interactions, i.e.

- A – B
- B – C
- C – A

An alternative might be the case where the three parties form a consortium D, in which case, the relationships are between the parties and the consortium itself, i.e.

- A – D
- B – D
- C – D

Two kinds of inter-organisation relationship are typical within the business world. The first is a "master / slave" relationship. This is an asymmetric relationship where one party takes on the rules and policies laid down by the other. A typical example of this might be a sub-contractor interaction with a large manufacturer. All of the sub-contractors must meet the rules of the large manufacturer as a condition of doing business with it. One administrative domain has priority over another, such that the policies of the VEO describing the relationship may be indistinguishable from the policies of the master.

The second relationship is "peer-to-peer". This is a symmetric relationship between the two parties. Most inter-organisation relationships fall into this category. This differs from the "master / slave" relationship in that neither administrative domain has priority over the other. The policies of the VEO describing the relationship are likely to be some combination of the policies of the contributing organisations.

3 Business Scenarios

This section describes a number of common / emerging business scenarios. The purpose of these scenarios is twofold: First, they serve for exemplifying the general Business Grid vision in a concrete context. Second, they form the basis for concrete business requirements which impose various technical requirements and challenges and ultimately lead to the overall strategic research agenda.

Context: The shift towards a service-oriented economy

An important context for all the scenarios discussed in this section is the overarching NESSI vision on the anticipated transformation towards a service-oriented economy. It "starts with a fundamental change in our approach to computing technology, moving away from the physical view of data processing devices to the functional view of data transformation services and that innovation in businesses and public services drive the evolution of these services down to the technology level. In this context, NESSI grand challenge is to '*transform the Internet to service your life*'.

The idea is to have a service-oriented economy based on a new generation of the Internet characterised by the following core properties:

- *Alive through services* – the service model enables to continuously improve the functionality available to users, making the system so responsive that it seems alive
- *Pervasive through trust* – the services are not only available everywhere, anytime, but are also pervasively trusted through the combination of strong end-to-end security and availability properties and a fully fledged collaboration between regulatory, technical and infrastructural environments.
- *Rich through knowledge* – the system supplies knowledge and expertise-based functions for the valorisation of business logic services and thus confers a competitive edge to the European economy
- *Invisible through ICT* – the system decouples provisioning and management of the infrastructure from use of services, which allows users and service providers to focus on the areas where they are best positioned to create value. This boils down to requirements, scenarios etc." [127]
- In terms of future IT infrastructures, the NESSI vision assumes that "IT resources are adapted automatically as a service-oriented utility according to high-level demands" [127] so that ultimately a "Dematerialisation of IT" [127] takes place.

Methodology

The presentation of business scenarios is structured according the basic model introduced in Section 2. The initial enterprise scenario is used as baseline based on which all the other scenarios are described as extended (and consequently more complex) setups. Each scenario will focus on its specifics which make it different from any predecessor scenario.

Scenarios differ in when we expect them to become relevant in the context of Business Grids. The first three business scenarios (enterprise, hierarchical enterprise, hosting) are assumed to have most short-term relevance. Others (such as Value Chains) are assumed to have more long-term relevance.

The presentation of each scenario is introduced by describing the general setup, followed by the scenario-specific vision of Business Grids and its expected business impact. Then, the specific business requirements are detailed (including their relevance for the business) out of which research challenges are later derived in Section 5.
Business requirements are discussed according to the following major categories:

- *Functional & commercial issues*, i.e. issues stemming from functional requirements typical for enterprises that want to use business grids. Especially important in this area is the commercial context under which business grids are operated. This covers issues such as service level agreements, cost and revenue agreements, etc.
- *Dependability*, i.e. the trustworthiness of a computing system in terms of availability, reliability and safety. Dependability covers issues such as redundancy, autonomy, quality of service etc.
- *Security* which we list as separate issue from dependability because of its paramount importance in business interactions. Security covers issues such as confidentiality, authentication, integrity, authorization, etc.
- *Performance*, i.e. the system performs and uses resources in an efficient, predictable and accountable way.
- *Interoperability*, i.e. the inter-working with other systems including legacy systems.
- *Manageability*, i.e. the easy, transparent and low-cost management of systems in particular including their maintenance.
- *Governance*, i.e. the specification and assurance of requirements and policies at various levels.
- *Flexibility*, i.e. the capability to react on/implement changed business requirements as fast as possible. Specific issues in here are about scalability etc.
- *Sustainability*, i.e. the ability to meet environmental criteria in both purchasing and design decisions.

These categories are neither complete nor completely independent from each other. However, they serve as our major perspective under which we present the following requirements as well as the various research challenges discussed in Section 5. Except for the first one, all categories focus on non-functional requirements and as such are often hard to express, measure and quantify in current business systems.

3.1 Enterprise

Setup

Basic enterprises consist of one homogeneous administrative domain. Following Section 2, they can be supported by basic enterprise grids which should provide general-purpose infrastructure services within one domain and which should ultimately act as the ICT backbone for the complete IT infrastructure.

Vision & Expected Impact

The vision of Business Grids for the Enterprise follows the SOKU vision (service-oriented, knowledge-assisted, and utility) as it aims for

- Providing IT resources as a service to the whole organization; this will significantly improve the flexible usage of IT resource within an organization, in particular for non-technical experts.
- Transparency on IT resources and related business figures; this will be a milestone for further industrialization and professional usage of IT within enterprises.
- IT provisioning and management as a utility; the simple and dependable usage of IT will allow for significant operational savings compared to today’s systems which still require a lot of manual maintenance.

Overall, Business Grids will change the role of IT in companies from being seen as a cost driver to be an enabler for agile businesses.

Business requirements

The basic enterprise scenario imposes the following major business requirements.

Functional & Commercial

- To support **fast provisioning** of systems (contribute to shortened time to market) and **low costs** (TCO).
- To support the **reliable and secure management of business data**. This is vital to any business as all of its internal activities eventually materialize in changed business data.
- To allow for **operating IT infrastructures as a business**. This includes complete transparency on how system parameters or setup decisions affect the total cost of ownership.

Dependability

- To allow for dependable service behaviour according to agreed non-functional service properties.
- To allow for **high availability** of (possibly virtualized) infrastructure resources in case of failures or maintenance activities. This contributes to overall high availability of business solutions facing the situation that already short downtimes in service-economies can lead to significant loss of money or even a company’s bankruptcy.
- To support the **balancing of availability levels with economic costs**. This allows enterprises to tune the actual availability to the most profitable level.
- Enterprise grids should be **autonomic** in the sense that they should be able to independently repair and recover in the event of errors, including complete failure of system components.

Security

- To reflect **security policies on infrastructure level** thus guaranteeing integrity and confidentiality of business data at the lowest possible level.
- To support application domain specific end-to-end (infrastructure to user) security demands.
- To support several levels of secure authentication to guarantee the possibility of providing several different levels of services to different classes of users.

Performance

- To support **prediction and accounting** of the non-functional behaviour and in particular the performance of applications, services and resources. Thus, providers can plan and manage their environment properly, users can know in advance what they can get and the resources used can be accounted for properly. This is especially important in service-oriented systems where services from different providers are composed to make higher-level services.
- To support **massive enterprise job scheduling** by pre-emptive and/or planned allocation/booking of the necessary infrastructure resources.

Interoperability

- Enterprise Grids can be implemented **transparently to arbitrary business applications**. Thus, the benefits of enterprise grids can be applied to complex existing IT stacks (including Grid environments) in place today.
- They should be **standards based and offer effective interoperability** with other standards based services within and outside the enterprise.

Manageability

- To allow for **homogeneous, low-cost, secure, easy and transparent planning & management** of arbitrary infrastructure resources. This leads to maximum exploitation of available infrastructure resources with the lowest possible effort.
- To allow for harmonization of different co-existing management tools.
- To remain fully-functional while undergoing **incident based or planned maintenance**.
- To enable **on-demand provisioning and termination of services**.

Governance

- To support the **transparent translation of business requirements and policies** to infrastructure capabilities and vice versa thus supporting their reflection and enforcement at infrastructure level.
- To reflect **enterprise policies on infrastructure level** via appropriate logging, tracking and auditing of resources and services. Such infrastructure support for compliance offers most robust and reliable implementation of governance policies even providing some resilience to badly behaving software, unforeseen security holes or erroneous sizing.

Flexibility

- To support **flexible changes of business processes and applications** which impact on the underlying infrastructure, such as dynamic allocation of additional resources. This supports the overarching goal of business agility, a major prerequisite for successful interaction in rapidly changing markets.
- To **dynamically adapt to meet demand** at any point in time in its operations to support sudden ‘bursty’ or unpredictable peak workloads on demand.

Sustainability

- To support **sustainable operation** of data centres with lowest possible environmental resource consumption. Apart from simple cost savings this becomes highly important for public and customer acceptance of IT solutions.

3.2 Hierarchical enterprise

Setup

Larger enterprises are typically organized via a hierarchy of departments, each of which following the general enterprise policies while potentially deviating from them e.g. due to different legislative constraints.

Vision & Expected Impact

The vision of Business Grids for the Hierarchical Enterprise is that they allow for a harmonized enterprise-wide infrastructure that is built upon a hierarchy of departmental grids. Again, Business Grids will follow the SOKU vision as they

- Allow for sharing of IT-resources in a service-oriented way.
- Transparently reflect the organisational structure of different enterprises.
- Support automated IT provisioning and management on enterprise level (where possible) and department level (where necessary).

Overall, Business Grids will be an enabler for efficient, flexible and transparent IT operation in complex enterprise setups. In particular, they will be an enabler for realizing changes in enterprise's organisational structure to meet internal and external demands.

Business requirements

In addition to basic enterprise scenarios, hierarchical enterprises impose the following major business requirements.

Functional & Commercial

- To support both **central and local service level and cost level agreements** as different departments might be operated as almost separated business entities (e.g. different cost centres).

Security

- To implement both **central and local security policies at the infrastructure level** and to support them at higher service levels.

Performance

- To support distributed infrastructures by considering their impact on resource management and performance prediction.

Manageability

- To respect **departmental boundaries in resource management** as departments might be willing to share some resources while others are used exclusively.

- To support **a consistent approach to management from central or local resource and service delivery platforms** including autonomic conflict management and resolution.

Governance

- To **reflect enterprise policy hierarchies on infrastructure level**, either following a hub-spoke model (where the degree of centralised control and policy consistency with central directives is high) or a distributed model (where the degree of centralised control and policy consistency with central directives is low because individual policy controllers retain significant freedom within a loosely defined Enterprise IT policy framework).

Flexibility

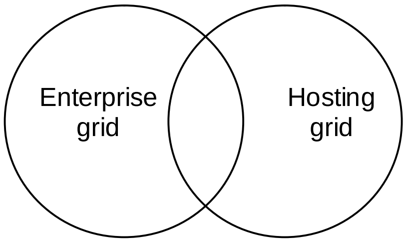
- To **propagate changes in the departmental structure** of the enterprise automatically to the IT infrastructure.

3.3 Hosting

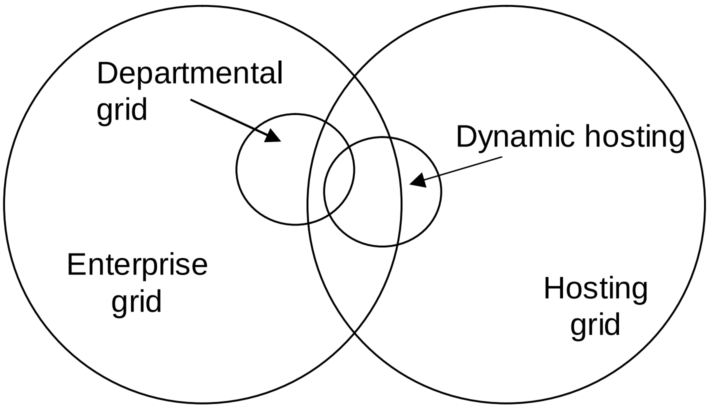
Setup

Hosting environments are about providing resources and/or services to customers. The actual offer may comprise pure resources, predefined services or environments for running customer-defined services (e.g. computing on demand). An enterprise that provides hosting facilities can also be described using the terminology of an enterprise grid. Today, most hosting can be described to be 'static'; in that the resources and services are pre-allocated by the hosting company to the using company well in advance of them being used and are typically allocated for a long period of time (weeks, months, years). The relationship between the hosting and using companies is well-defined and also typically long-term. Static hosting relationships are quite predictable and hence can be billed in a well understood fashion. In the future, these static relationships will likely become more 'dynamic'. Dynamic hosting is typically much more short-term in nature. Hosting resources and services are allocated on demand and in as near to real time as possible. The length of time that a resource or service is allocated to the using company may potentially only be for a few seconds or minutes. Relationships between companies may similarly only exist for the length of time that the resources or services are being used. Dynamic hosting is much more unpredictable and is likely to require new billing and accounting mechanisms.

A hosting relationship is a variant of the Virtual Organisation, in that it is simply a special relationship between two administrative domains. Here we see a simple environment where the hosting company (represented by the enterprise on the right) provides resources to the consuming enterprise on the left.



Inter-relationships that are more complex can also exist; for example the provision of dynamic resources by a hosting company to a departmental grid within an enterprise. Such an inter-relationship can be shown thus:



Some complexity starts to arise here. There is a two-level relationship between the two enterprises. The former is the VEO that is a high-level interaction between their two administrative domains (a company-to-company agreement). The relationship between the departmental grid and the dynamic hosting environment is then a special case relationship within, albeit still encompassed by the overlap of the two administrative domains. This special relationship is likely to be a subset of the VEO relationship. These problems are relatively easy to solve in the case of a static hosting situation. As the hosting needs of the using company become more dynamic, the challenges of building (and maintaining) relationships, of creating the necessary security and interactive linkages and of managing the interaction between the two administrative domains become significantly harder problems to solve.

Vision & Expected Impact

Apart from adding transparency and dependability Business Grids will support highly dynamic hosting scenarios. These are characterized by

- Automation support for negotiation and setup of hosting relationships and service level agreements.
- On demand (or at least near real-time) allocation of resources and services.
- Comprehensive management of the complete hosting lifecycle from enactment to decommissioning,

Business Grids will significantly impact hosting scenarios by supporting new hosting business models (e.g. software as a service, computing on demand) and by enabling short-term relationships (down to a few seconds or minutes). Overall, they allow for much more flexible interactions between different organizations.

Business requirements

Apart from the characteristics discussed above, basically all the requirements listed for (hierarchical) enterprise scenarios apply in hosting scenarios but to an even more rigorous extent. These and some further specific requirements are detailed below.

Functional and Commercial Requirements

- To support **reliable and secure accounting and billing, especially for dynamic, short-term hosting relationships**. In particular, they must support it for different business (e.g. software as a service), financial (e.g. pay as you go, monthly contract, pay in advance, pay in arrears, volume discounts) and licensing models.
- To support **SLA monitoring and penalties management** in order to deal with violated SLAs.
- To support **decisions when to break SLAs**. A provider may service different clients and it is possible that, at some times there are difficulties in complying with all SLAs. The provider may decide that it is more cost-effective to break one or more SLAs in order to be able to comply with the rest (e.g. to prioritise a large or important customer over a smaller one).
- To support dynamic software license management.

Dependability requirements

- To support **flexible, customer-specified dependability levels** for hosted solutions which are automatically mapped to the infrastructure setup and which can even be changed over time and per requested service. Thus, customers can choose the required availability level and balance that with the resulting costs.

Security requirements

- To guarantee **strongest security and isolation** of hosting resources between different customers and hosting providers. As hosting customers put highly sensitive data into a hosting grid they need complete confidence in their secure and isolated management.
- To provide strongest protection against external threats, e.g. from the Internet.
- To support **policy management services** maintaining security policies, enabling access rights and enforcing restrictions for all involved actors consistent with the terms of existing SLAs.

Performance requirements

- To support **highly efficient operations**, e.g. by intelligent sharing of resources among hosting customers. This is of special importance in mass hosting scenarios
- To support **prediction and enforcement of performance characteristics** (response time and resource usage) and to derive them from high-level requirements. This allows the provider to offer appropriate SLAs and plan resource deployment to deliver against them efficiently.

Manageability requirements

- To support **remote infrastructure management**. Though a hosting environment may belong to the administrative domain of the provider, it may be physically collocated with the client's data centre. Such setups require strong support for the remote management of the hosting environment, in particular the hosting infrastructure.
- To support **mass automation** including automatic software deployment in order to allow management effort to scale to very large numbers of hosting customers.

Governance requirements

- To support **governance and compliance for both providers and customers**.

Flexibility requirements

- To provide **hosting services in a highly dynamic and on-demand way** in order to satisfy spontaneous customer needs in extremely short time frames (seconds or minutes).

3.4 Extended enterprise

Setup

Enterprises may intend to extend the scope of their IT-managed activities beyond the traditional boundaries of data centres and desktop machines. For example, pervasive scenarios aim for the integration of pervasive devices (PDAs, mobile phones, RFID systems etc.) into the ICT backbone of a company. Real-time scenarios aim for the integration of real world entities via sensors/actuators (e.g. in shop floor integration or asset tracking) thus exhibiting more real-time or event-driven characteristics. Other remote devices such as satellites may be relevant as well.
Extended enterprises and their special-purpose devices have some characteristics which clearly separate them from traditional data centres:

- The heterogeneity of devices is much higher. In particular capabilities, interfaces and also connectivity of devices may differ significantly and sometimes have quality characteristics much lower than typical data centre resources.
- Device capabilities may quickly vary over time (i.e. connectivity, battery).
- Both, devices and services may be mobile.
- The association of devices with administrative domains may be more complex: devices may belong to the core enterprise domain, may be interlinked by concepts of a virtual organization and finally may even change their ownership (e.g. when they are sold possibly as part of another good to another party).

Vision & Expected Impact

The vision of Business Grids for the Extended Enterprise is that they allow for a harmonized enterprise-wide infrastructure that includes all types of devices. Again, Business Grids will follow the SOKU vision as they

- Allow for seamless access to a large variety of IT resources in a service-oriented way.

- Ensure transparent life cycle management that considers overlapping and changing administrative domains.
- Support automated IT provisioning and management on enterprise level (as appropriate) and device level (as appropriate).

Overall, Business Grids will be an enabler for increased real-world and real-time awareness (both globally, and locally) and will support overall enterprise flexibility by enabling local interaction which takes place close to the problem point.

Business requirements

The extended enterprise scenario reveals the following specific business requirements.

Functional and Commercial Requirements

- To support **real-time processing** on events happening to properly react to varying context. IT services will take benefit of rapid adaptation to the changing environment, including device and/or user mobility.
- To support **user mobility** (not linked to only one device) and service mobility (transferring sessions among devices).

Dependability

- To provide a **reliable access and usage** of resources in pervasive devices (PDAs, mobile phones, sensors, etc) taking into account their context (i.e. geographic location) and availability (i.e. connectivity intervals).
- To allow for **disconnected operation** in the backend in case some special devices are not online (due to device or network failures). Predictable disconnection and backend operation are essential for assuring a company’s ability to continue operation even in presence of failures (which are quite likely to happen with pervasive and real-world devices).

Security

- To enable proper **access management** for central IT operations and local device users/owners.
- To support **lightweight security methods** and complementary methods (usage logging, reputation system, assessment from actual security conditions, etc.) whenever possible. This allows pervasive devices to accomplish security requirements while saving resource usage.
- To support **integrity and confidentiality** of operations and shared data. Especially in pervasive scenarios, access to enterprise infrastructure from mobile devices must be provided via secure channels to ensure data integrity.

Performance

- To allow for **scalable event handling** for possibly thousands or millions of events.
- To realize an **intelligent resource management** (in particular, energy management). As resources of special-purpose devices can be quite limited it is of great importance to manage them intelligently in order to maximize the business benefit.

Interoperability

- To establish **standard management and access interfaces** for highly heterogeneous devices, network interfaces and protocols.

Manageability

- To allow for **balanced management procedures** that properly balances the needs and capabilities for local and central device management.
- To establish a **proper lifecycle management** for pervasive/real-time devices that might change ownership over time.

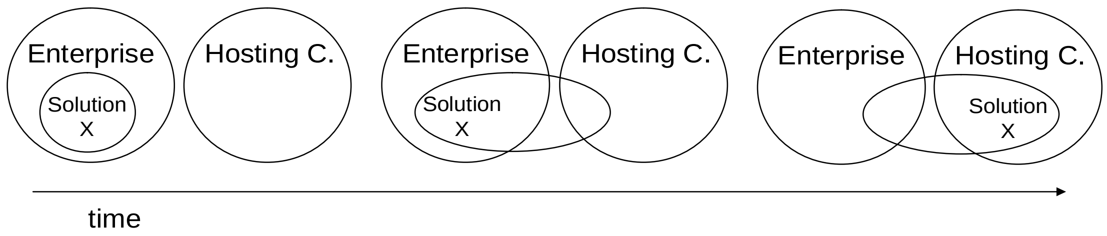
Flexibility

- To support **dynamic reconfiguration of devices**. This is essential for agile businesses to adapt the scenario/process under which pervasive and real-world devices are used.

3.5 Dynamic Outsourcing

Setup

Decisions on whether IT systems are operated in-house or via an external hosting provider may change over time as the strategies of organizations evolve. Traditionally, outsourcing decisions do not just depend on strategic requirements but have to carefully balance these with the practical effort, overhead and side effects (e.g. interruption of business) of an outsourcing process. The following picture shows such a dynamic outsourcing process where a solution X, initially provided by an enterprise infrastructure is migrated to a hosting company. Ideally, solution X is fully operational during the whole migration process.



Vision & Expected Impact

Business Grids are envisioned to support dynamic outsourcing (and likewise insourcing) processes by realizing the SOKU vision in the following aspects:

- Automated migration of IT solutions/services between different administrative domains (service-orientation).
- Comprehensive and secure management of migration process and handover across administrative and security domains (knowledge-assisted).
- Ensuring full availability and normal operation during all phases of the migration process (utility).

Business Grids will significantly impact organization's ability to change and adapt their company strategy in terms of IT provisioning decisions. In particular, Business Grids will prevent organizations from being locked by previous decisions which over time turned out to be inappropriate.

Business requirements

In addition to the previous and in particular the hosting scenario dynamic outsourcing reveals the following specific business requirements.

Functional and Commercial Requirements

- To provide **planning support** for migration processes and the impact and side-effects of outsourcing decisions.
- To support **migration** of running IT solutions/services between administrative domains and possibly across wide-area networks.

Dependability & performance requirements

- To support **flexible, customer-specified dependability/performance levels** for all phases of a migration process. Similar to the hosting scenarios, customers shall be able to choose the required dependability/performance level and balance that with the resulting costs.

Security requirements

- To guarantee **strongest security and isolation** of all services and resources which are involved in a migration process.

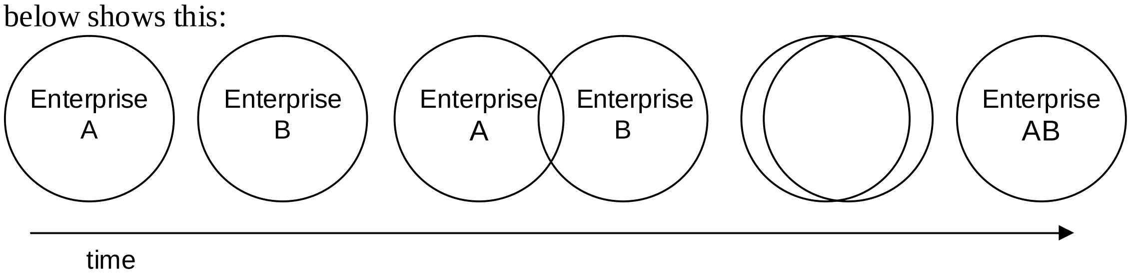
Manageability requirements

- To support a **transparent migration process** including a proper handover across administrative and security domains.

3.6 Mergers & acquisitions

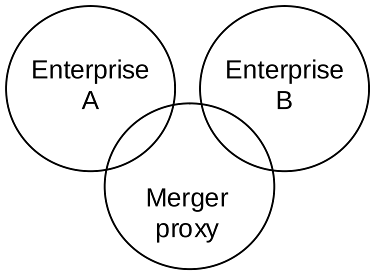
Setup

Merging with, or acquiring another company is an increasingly common method for companies to grow. The traditional means by which companies merge typically takes a long time to fully complete. This can be envisaged using the enterprise grid model in one of two ways. The first occurs when two enterprise grids merge over time. The diagram below shows this:



The two enterprise grids that are initially separate form a virtual organisation that links their two grids together in a limited fashion. Over time, as the grids coalesce, the virtual organisation encompasses more and more of each individual enterprise until a single merged enterprise remains. In the case of a merger, the relationship is likely to be more of a symmetric one. In the case of a takeover or acquisition, the relationship between the enterprises is more likely to be asymmetric.

One other form of merger / acquisition mechanism is also possible here. This introduces the concept of a merger proxy, as shown below:



The merger proxy is an external entity that provides a link between the two enterprises. Both enterprises form virtual organisations with this proxy and use the services of the proxy to merge logically. Such a proxy might be used to facilitate a new range of business services in this arena.

Vision & Expected Impact

Business Grids are envisioned to support scenarios of mergers and acquisitions by realizing the SOKU vision in the following aspects:

- Providing services supporting the actual merger/acquisition process in terms of managing the consolidation of administrative domains, infrastructure resources Automated migration of IT solutions/services between different administrative domains (service-orientation).
- Comprehensive management of merger/acquisition processes in terms of IT landscapes, components, processes and users (knowledge-assisted).
- Supporting an almost transparent shielding of the merger/acquisition process from the users of business solutions (utility).

Business Grids will significantly impact the dynamicity of economies as they lower the technical complexity for merger/acquisition processes.

Business requirements

Merger and acquisition scenarios reveal the following specific business requirements.

Functional and Commercial Requirements

- To provide **planning support** for merger & acquisition processes and the impact and side-effects of them.
- To provide **support services** for conducting the actual merger/acquisition process.

Performance requirements

- To support **resource consolidation** of previously heterogeneous resources into one overall and logically homogeneous resource pool.

Security requirements

- To support **harmonization and consolidation** of security policies and associated mechanisms.

Manageability requirements

- To support a **transparent merger/acquisition process** clearly reflecting the various stages of a merger/acquisition process.

Interoperability

- To support **standards for IT infrastructures** that go beyond service-/interface-level but allow for consolidation of resources.

Flexibility

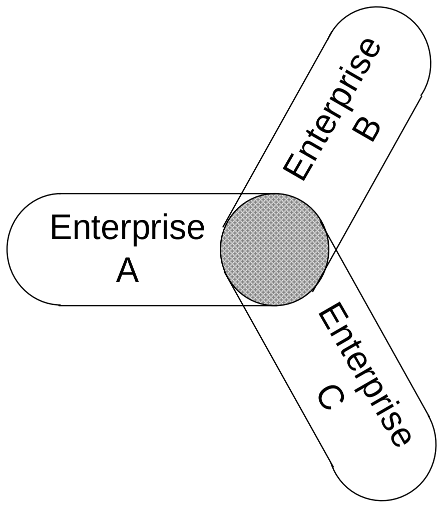
- To support different **merger/acquisition styles** including staged processes, symmetric/asymmetric processes and merger proxies.

3.7 Virtual organisations

Setup

As already sketched in Section 2.5 Virtual Enterprise Organizations (VEOs) are built whenever two or more administrative domains overlap and share resources. This may happen in a master-slave or peer-to-peer modus. Moreover, virtual organizations may be formed at different levels, e.g. via sharing more high-level services such as Web services or by sharing more low-level resources such as files, processing capacity etc.

Though many of the presented scenarios already revealed various VEO-like characteristics we believe that the full vision of VEOs has a couple of unique aspects that qualify the discussion of VEOs as a separate scenario. Furthermore, specific issues gain further practical relevance as soon as at least three administrative domains overlap.



Vision & Expected Impact

Business Grids are envisioned to become a trusted infrastructure for the formation and management of virtual organizations. They realize the SOKU vision in the following aspects:

- Providing support services supporting the formation and management of virtual organizations (service-orientation).
- Supporting transparent specification of VEO policies which may be quite complex and fine-grained (knowledge-assisted).
- Embedding compliance with VEO-aware and specific policies into the infrastructure (utility).

Business Grids will significantly impact the simplicity for initiating and maintaining dynamic virtual organizations thus contributing to the overall agility of economies.

Business requirements

Virtual organizations reveal the following specific business requirements.

Functional and Commercial Requirements

- To provide logically **central support services** for the initiation, management, monitoring of virtual organizations (e.g. central registries) and the enacted, distributed business processes. Contractual details have to be specified and (preferably) automatically negotiated between the partners in a VEO.

Dependability & performance requirements

- To support **accountable and isolated access** so that resources cannot be used or changed in an unintended way.
- To support **reputation awareness** of participants (i.e. their performance in terms of fulfilling promised SLAs) based on monitoring of usage control.

Security requirements

- To embed SLA-aware **policy enforcement** into the infrastructure (e.g. of security policies and associated mechanisms) according to the negotiated collaboration/interaction pattern.
- To allow for **secure access and usage** to the resources of the partners, e.g. by maintaining and distributing usage control at the service (global VEO policies) and the computational levels (local policies + global policies enforced at node level).

Manageability & governance requirements

- To support a **transparent specification/management** of all documents necessary for a VEO, e.g. policies reflecting specific VEO roles (initiator, contributor etc.), collaboration/interaction patterns, SLAs etc.
- To support **automated VEO management** activities including proper access interfaces (console, API, GUI).

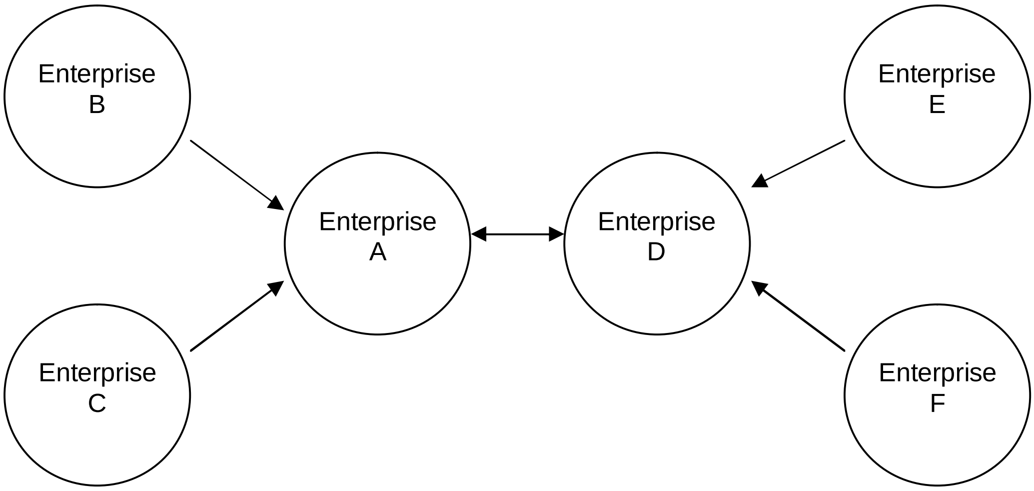
Flexibility

- To support dynamic **changes of virtual organizations**, e.g. mechanisms which allow partners/resources to join/leave a VEO or to be transparently replaced.

3.8 Value networks

Setup

Value networks are ways in which organisations interact with each other to drive increased business value. There are many of these in common existence today; the supply chain being the most familiar. The application of enterprise grids to this environment will allow more exotic business value networks to be formed and dissolved as business conditions change. The following shows a business value network around the joint development of a product:



In this example, Enterprises A and D are collaborating on the development of a new product. Enterprise A has sub-contractors or partners B & C, whilst Enterprise D has similar relationships with E & F. The network (shown here with arrows) consists of a series of virtual organisation relationships in both master-slave and peer-to-peer between the respective companies. By pooling resources, information, insight & capabilities a greater business value can be achieved than by a single organisation acting alone.
Special purpose service are likely to emerge to support the formation and management of value networks such as broker services, mediator services, marketplaces, accounting, authorization or authentication services etc.

Vision & Expected Impact

Business Grids are expected to become the trusted infrastructure for complex business value chains. They realize the SOKU vision in the following aspects:

- Provisioning of generic infrastructure and support services (service-orientation).
- Transparent and trusted management and accounting of resources according to value chain requirements (knowledge-assisted).
- Support for automated formation of value networks with dependable characteristics (utility).

Overall, Business Grids will ease the formation and management of complex business value chains which ultimately supports highly dynamic and innovative business interactions at the scale of complete economies.

Business requirements

Value networks reveal the following specific business requirements.

Functional & Commercial

- To provide or support basic services for the formation and management of value networks, such as mediation, brokering, accounting, authentication or authorization services.

Non-functional characteristics (dependability, security, performance)

- To **reflect non-functional characteristics** (dependability, security, performance, …) in service discovery and composition mechanisms.
- To allow for **awareness of non-functional characteristics** on the value network level, i.e. to provide prediction, monitoring and readjustment mechanisms.

Security

- To **secure accounting and management data** thus guaranteeing integrity and confidentiality of these data which partially needs to be shared between different organizations and partially needs to be kept strictly private.
- To support application domain specific end-to-end (infrastructure to user) security demands across whole value networks.
- To support a wide range of authentication mechanisms, including ad-hoc dynamic authentication linked to business trust, which can support pseudonyms but still allow application providers to manage risks.

Governance

- To reflect **legal policies on value network level** via appropriate logging, tracking and auditing of resources and services.

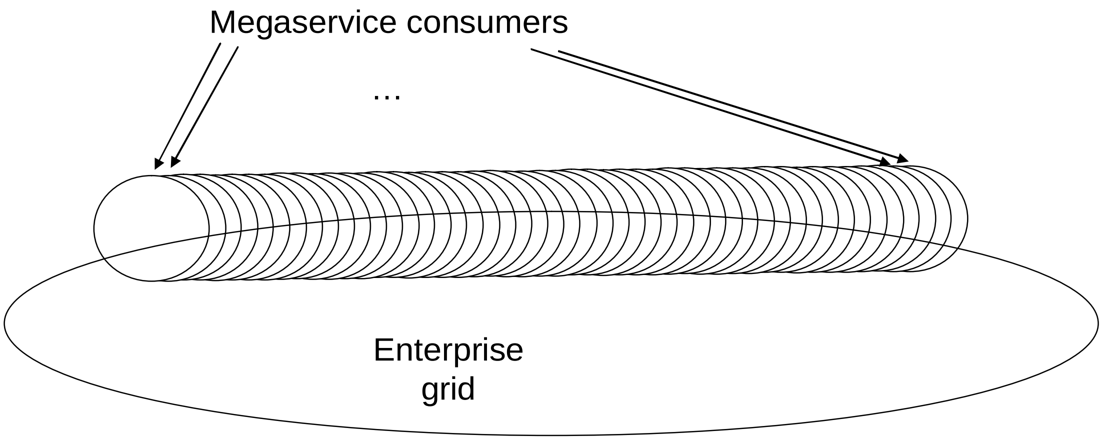
Flexibility

- To support **flexible setup and changes of value networks** which impact on the underlying infrastructure, such as dynamic allocation of additional resources or reassignment of existing resources.

3.9 Mega Services

Setup

Mega services are literally offered to millions of consumers possibly distributed over the globe. The type of services can widely vary, ranging from raw data storage / data caching, through virtual machine execution to high level software services, usually featuring personalized persistency and Web based access. The majority of the service consumers are considered to be citizens, whereas small and medium sized companies are also representing a significant customer segment.



Mega services are envisioned to be globally available, in many cases ubiquitous services, foreseen to actively shape the relationship of the population towards the information age services. Reaching extreme amount of customers mega services are expected to have significant volume, raising the most successful service providers among the largest global enterprises.

Vision & Expected Impact

Business Grids are expected to become a flexible and trusted infrastructure for mega services. They realize the SOKU vision in the following aspects:

- Homogeneous, dependable provisioning of globally distributed resources possibly across administrative domains (service-orientation).
- Transparent management globally distributed resources, customers and demands (knowledge-assisted).
- Support for seamless scalability of service offerings from small, localized to ubiquitous and very large scale (utility).

Overall, Business Grids will ease the provisioning of dependable mega-services in particular the dynamic provisioning of required resources. This will allow even small organizations to provide mega services in a simple and dynamic manner and also offers opportunities for intermediate resource providers to act with a global customer base.

Business requirements

The requirements listed in this section are extending those marked up in the hosting scenario in Section 3.3. The differences between the two scenarios stem from the very large customers base requiring a giant service infrastructure, potentially spreading over the world.

Functional & Commercial

- To support the adaptation of the business framework and the services to various country specific requirements (e.g. legislative restrictions).

- To provide a complete automation of the entire customer relationship management, service provisioning, service execution, SLA monitoring and charging. To restrict the need for human interaction to special (i.e. erroneous) conditions.

Dependability requirements

- To provide highly available and reliable operation on geographically distributed service infrastructures, possibly involving several service elements operated by 3rd party partners (typically network services). Although the availability and reliability requirements are dependent on the SLAs coupled with the actual services, it is a general requirement to avoid failures affecting large customer subsets or whole geographic regions.

Manageability requirements

- To support mass-scale automated remote infrastructure management. Providing services for extreme high number of customers require facilities composed by large amounts of potentially geographically distributed nodes, linked by vast networking infrastructure. This needs the utmost level of automation in the infrastructure management.

Flexibility requirements

- To support extreme large geographically distributed customer bases. To support high service volume changes happening in relatively short timescale in a geographically uneven manner.

3.10 Additional scenarios

Noteworthy, that the introduced scenarios can be combined in various ways in order to form more complex, compound scenarios. For example, resources from an extended enterprise might be offered via a hosting scenario. Multi-party federations or VEOs may become part of a value networks. However, these compound scenarios are not described in this document as we do not assume that they create requirements in addition to the core scenarios.

Various services might complement the usage and operation of a Business Grids infrastructure. For example discovery mechanisms might support the selection of appropriate service offerings according to technical requirements and offered SLAs. Rating systems may support the assessment of the trustworthiness of a service provider. Trusted brokers may support the matchmaking process between service providers and customers. However, this kind of additional services is not specifically addressed within this SRA.

A couple of additional scenarios have been suggested by the community, such as “federation of SMEs”, “resource brokering a la eBay” or “edge computing”. These scenarios will be worked out in collaboration with community stakeholders for future versions of this SRA.

4 State of the art

This section presents a high-level overview on the state of the art of technology relevant for the Business Grid vision.

The section is structured in three parts: in the first part, we survey the existing Grid infrastructures, from four different points of views: scientific grids, enterprise grids, virtualisation-based grids, and mobile grids. Then we discuss in more detail specific topics related to functional and commercial requirements of grid technologies, structured using the criteria presented in the introduction of Section 3. Last, we list some emerging technical trends that will have an important impact on future business IT systems in general and Business Grids in particular.

4.1 Infrastructures

Scientific Grids

In this section, we introduce some of the major scientific grid approaches. We concentrate on standards, middleware, and a collection of selected research projects. In this domain, "grid computing" is an extension of scientific cluster computing, merged with concepts from large-scale distributed computing, in particular Internet-based computing paradigms, such as peer-to-peer and service oriented architectures. The primary goals are either computing research oriented, or to facilitate scientific computing. The major themes deal with inter-organisation resource sharing, heterogeneity, scalability, decentralisation and dynamicity. While issues such as stability, security, and performance are important, these are usually deemphasised in this domain.

Grid computing originated as an approach for dynamically sharing inter-networked computer systems for compute-intense scientific applications [4] in situations where dedicated computer clusters were not available or provided insufficient computing power. The term Data Grid was introduced to refer to an architecture primarily focusing on the distributed management and analysis of large scientific datasets [5]. With the increased commoditization of computing hardware and pervasive business computing, the management of distributed sets of dynamic heterogeneous computing resources has further emphasized the importance of specific grid computing techniques and research.

The existing scientific grid projects present, to various extents, a number of features related to the NESSI-Grid requirements. This is particularly true in the areas of ***Functional issues, Interoperability, Dependability*** and ***Security***, and naturally ***Performance*** and ***Flexibility***. In many cases the ongoing research efforts, software components, platform development, and standardisation, will be a valuable basis for either inspiration or usage in enterprise grids.

There are several national and international grid infrastructures now in place. The following provides a summary of the major initiatives.

The EGEE-II [16] project has inherited from EDG [11], LCG [12], and EGEE-I [15] a grid software stack consisting of a combination of Web Services, Globus middleware, Condor [13], and custom services to provide the gLite [14] software set which is used by hundreds of sites to provide the EGEE grid environment. The services of gLite are grouped in five main categories, namely access services, security services, information & monitoring services, data services, and job management services. Currently gLite (the grid software) and EGEE (the computational grid overlay network) form the world's largest general purpose scientific grid, incorporating over 240 sites in 45 countries with 41,000 computing cores available, and running over 100,000 jobs per day. It has a community of over 100 virtual organisations and through these a network of thousands of users.

Other significant academic grid environments include DEISA (Distributed European Infrastructure for Supercomputing Applications) [41], which is a collaboration of 11 European supercomputing sites to link their resources (35,000 processors, of various kinds) through common interfaces to enable shared computing. The UNICORE software [22] has been adopted as a common portal to DEISA. Within France, the Grid5000 [42] research grid provides 3900 cores across 9 sites for a private grid computing network, emphasizing inter-site and parallel grid computing research. The Nordic countries in Europe have collaborated in producing the ARC software which is used within NorduGrid [18], a system deployed at several dozen sites, primarily in Northern Europe, and providing over 9000 cores. The CERN LCG project primarily makes use of three underlying grid infrastructures: EGEE, NorduGrid, and OSG.
The US provides both the OSG (Open Science Grid) [43] and TeraGrid [44]. OSG has a similar background to EGEE, and is based on a combination of Globus, Condor, and VDT (Virtual Data Toolkit) software released by NMI (NSF Middleware Initiative), and links computing clusters from over 60 sites, mostly located in the US. There are 30 virtual organisations within OSG. TeraGrid links together a dozen supercomputing centres across the US, with a total capacity of over 1 PetaFLOPS, and a planned increase of 1 PetaFLOPS per year. TeraGrid federates classical supercomputing applications, mainly using Globus and Condor job managers, but now has more focus on multi-site synchronised tasks, based on VO structures. Within Japan, the NAREGI [45] grid testbed utilises project-specific middleware across 11 sites providing over 3000 computing cores and over 17 TeraFLOPS. China has also started a national grid project ChinaGrid [46], initially involving 9 sites with over 4000 cores.

In general these scientific grid environments are generic distributed job management systems with federated data access and a common security infrastructure for access control and accounting. There are many domain specific grid environments, such as the International Virtual Observatory Alliance (IVOA)[108] and the US National Earthquake Engineering Simulation Grid (NEESGrid)[107], however it is beyond the scope of this work to report on the domain-specific features which have motivated the development of these special purpose grid environments.

Enterprise Grids

Enterprise grid deployments generally rely on the same kernel technologies as the academic/scientific grids mentioned above, but due to the limited security mechanisms (or at least their perceived limitations or difficulty of use) are usually restricted to a single administrative domain within an enterprise. It is not uncommon to find deployments that span several administrative domains within an enterprise, i.e. between different arms of the internal IT organization (such as engineering, business and general desktop divisions), but there is little support for interactions between these administrative groups in the deployed infrastructure.

Commercially available solutions for these environments feature policy based scheduling and workload management on heterogeneous infrastructures made out of desktops, servers and clusters. These systems contain basic resource control and mechanisms for fault tolerance as well as analysis tools for performance and debugging. Due to the lack of standardization in this space these solutions typically support a variety of *de facto* standards, translated to solution-specific formats. Finally, these solutions often contain their own billing and user management solutions, partially integrating with common security infrastructures prevalent in enterprises.

For most commercial organisations, their highest value electronic assets are the data stored in files and data bases. The reliability of the systems which manage enterprise data is of the utmost importance, and enterprise computing priorities dictate continual improvement to the management of these assets, in terms of access performance, reliability, scalability, and integration. While several research grid projects have addressed issues to do with data management on the grid, very few have looked at the issues typically found in commercial data management systems, or attempted to achieve the level of reliability required from enterprise applications which are central to the daily business activity of large organisations.

The formation of the EGA (Enterprise Grid Alliance) in April 2004 provided a forum for discussing the requirements of enterprise grid computing and interoperability between the major grid computing technology providers (for software, hardware, and services). The EGA produced a Reference Model aimed at providing a common view of an enterprise computing environment and the requirements which should be addressed by Grid technology. Since the merger of the EGA and OGF, this work is now being continued by the OGF Reference Model Working Group [70].

In 2006 ETSI (European Telecommunications Standards Institute)[109] created the Grid Technical Committee to produce interoperability standards specifically related to telecommunications. TC-GRID is responsible for ETSI's contribution to Information Technology and Telecommunications (ITC) convergence standards in the grid domain starting with grid interoperability drafting test specifications, and should progress towards a range of grid standards to integrate telecommunications infrastructure with networked computing.

Examples of such commercially available solutions (without weighing their market share or importance) are IBM Grid Computing solutions [25], SUN's "N1 Grid Engine" [26], Platform Computing's "Enterprise Grid Coordinator" [27], DataSynapse [28], Univa Globus Enterprise [29], and Cluster Resources's "Moab" suite [30].

Virtual machine based grids

The paradigm of Grid computing targets the execution of increasingly large workloads on heterogeneous infrastructure landscapes including computing, network and storage resources which span multiple administrative domains. Recently virtualisation has re-gained popularity as a way to increase system utilization and to reduce cost of ownership by providing a common (virtualized) environment. The concepts of Grid computing and virtualisation can be considered as complementary: Grid computing allows for sharing resources through services, virtualisation technologies allow flexible resource partitioning with a high degree of isolation between services sharing these virtualized resources.

A variety of IT vendors, both large and small, are offering solutions that can transparently manage arbitrary collections of applications in execution containers that can be dynamically deployed onto a heterogeneous pool of physical hosts. These containers then are treated like jobs in a “normal” grid but can be dynamically controlled and the resources allocated to them adjusted.

Typically these containers are implemented using virtual machine monitors, and hence contain an operating system together with the deployed application. Examples of this technology are XEN [31], VMWare [32], Virtual Iron [33], SUN’s Solaris Container [34], KVM[111], OSGI containers[112], and VJSC containers[113], or Virtual System Center Manager from Microsoft [35]. Containers can also be implemented without hypervisors inside the operating system (i.e. SWSoft’s Virtuozzo [36] or IBM’s virtualisation suite [37]).

Given the growing popularity of virtualisation, many commercial products (and research projects) are being developed to dynamically overlay virtual machines over physical resources. These efforts, in general, try to simplify the use of virtualisation to provide the enterprise with the potential benefits this technology may offer, like server consolidation, virtual machine isolation, performance partitioning, or legacy application execution, among others. Examples of this kind of product or service are: Platform VM Orchestrator, IBM Virtualisation Manager, Sun Hedeby[114], Xen Enterprise, the open source project GridHypervisor[115], Data Synapse Grid Server [28], VM Ware ESX Server [32], and Amazon EC2 [93]. Features of these solutions are similar in capability to what enterprise grid solutions offer, but add “steady state” resource management, additional security mechanisms by encapsulating applications in containers and by adding control of network partitions through setting up of virtual VPNs.

Previous research efforts investigating the usage of virtualisation in grid environments mainly consider two aspects: the management of virtual machines on the grid or grid-like management of virtual machines. The EU projects RESERVOIR (FP7) and XtreemOS (FP6) follow a more holistic approach. RESERVOIR intends to enhance virtualisation infrastructures to be grid-aware, e.g. by supporting live migration of virtual containers across administrative domains. Furthermore, RESERVOIR will allow the grid to be virtualisation-aware, e.g. by exploiting low level metering information, provided by virtualisation technology, for the purpose of metering or billing. The project XtreemOS aims at developing an open source grid operating system and examines how far virtual machines or containers can support performance and the very strict security requirements for the isolation of data.

Mobile Grids

Extending enterprise grids to integrate mobile devices (from PDAs through mobile phones to sensors and actuators) introduces pervasiveness and ubiquity. Generally speaking, there are two main approaches to integrate mobile devices in enterprise grids, (a) using mobile nodes to gain ubiquitous access to the grid and (b) taking advantage of resources in mobile devices to create new mobile services.

There are a number of challenges to be faced when deploying grids in this sort of pervasive environment. From an infrastructure point of view two main issues arise: mobility support and context-awareness [75]. Mobility support enables ubiquitous access to the grid as well as provision of mobile nodes resources, whilst context-awareness allows an appropriate management of grid resources according to execution environments. Integrating these features of real time event driven behaviour will provide new service opportunities to enterprise grids and increased business flexibility.

Though most current grid deployments are intended to provide resources for computationally intensive scientific and enterprise applications, several approaches to integrate mobile devices into the grid have been proposed. Some of these initiatives have been based on gateway solutions, that is, using a proxy server connected to the grid [75][76]. GridLab [77][79] and InteliGrid [80] projects integrate mobile user access to distributed data, applications and business processes using this approach [81]. There are also other initiatives that address access to the grid from mobile nodes without proxies [82], however these have suggested implementation complexities which reduce operational efficiency.

Nevertheless, an extended enterprise grid will benefit from mobile devices not only by allowing them to access the grid, but also by sharing mobile device resources [83][84]. Based on OGSI / WSRF specifications and, more concretely, on the implementation of such specifications over Microsoft .NET framework, OGSI.NET (subsequently WSRF.NET), Mobile OGSI.NET was developed to extend grid computing to mobile devices specifically addressing the challenge of resource limitations and intermittent network connectivity [85].
This new framework might be referred to as the Mobile Grid, a full inheritor of the Grid with the additional feature of supporting mobile users and resources in a seamless, transparent, secure and efficient way [86]. A Mobile Grid consists of resources that are not centrally controlled and supports various kind of mobility (user, terminal, session).

The EU funded Akogrimo project [73] dealt with the problem of incorporating mobility awareness into the grid, providing a framework where network and service operators can develop new business activities in a distributed, mobile and pervasive environment. One of the main inventions achieved in Akogrimo is the mobile collaborative business service provision by workflows, which has been demonstrated in three different scenarios (eHealth, eLearning and disaster handling and crisis management).

Another EU funded project, XtreemOS [74], aims at extending Linux operating system services to provide users with all the grid capabilities associated with current grid middleware. Besides PC and cluster XtreemOS flavours, XtreemOS-MD, a fully compatible implementation for mobile devices, is currently being specified and developed, targeting first PDAs and then mobile phones. XtreemOS-MD is focusing on providing access to the Grid from mobile nodes without the need for a proxy or gateway. To that aim XtreemOS-MD is currently being designed and developed with no middleware layer but providing directly the operating system with grid capabilities.

4.2 Business requirements

The business requirements listed in section 3 are naturally steering many research activities, which range from fundamental research to tool development, with a specific interest in standardisation and interoperability.

One difficulty in reporting the state of the art and the ongoing research on these topics is that research and development is conducted in terms of functionalities that will ultimately be made available at infrastructure and middleware level, not matching the “business view” expressed in the requirement. One typical example could be the new trends on autonomic computing, which methods and tools will contribute to solve open problems in the areas of dependency (fault tolerance), performance (load-balancing, dynamic allocation of resources), but also flexibility (run-time discovery of new services, combined with component adaptation).

Nevertheless, we have structured this section following the classification and the taxonomy defined in the business requirements section.

Functional and Commercial issues: Data Management

File handling and data management are key issues in academic and scientific grid computing, and were one of the original motivators for the EDG project. This work can be broadly divided into grid file systems and grid databases. The latter has been covered by OGSA-DAI [56], and is of particular interest to the health and bioinformatics domains where data is generally stored in relational databases rather than files. To date, most grid systems have implemented their own bespoke interfaces to expose database resources.

Regarding grid-enabled file access, there is a much greater level of experience and convergence around the requirements. These systems require access control facilities, replication, data staging from mass storage systems (e.g. tape silos), meta-data catalogues to locate relevant files, and systems to map from logical (grid) file names to physical (on-disk) file locations. Furthermore, these systems must integrate with task scheduling and task placement, and for large data sets must consider pipelining the staging of task data (both in and out) to avoid leaving processors idle while data is moved across network links and to disk. gLite already takes data requirements into consideration during task scheduling, and systems such as Condor Stork [13] allow for explicit scheduling of data movement. SDSC has developed the widely adopted SRB [57] system which is now being redeveloped as iRODS [58]. GridFTP and RFT from Globus provide mechanisms for high-throughput reliable file transfers, including third party transfer orders. The LCG project has critical requirements related to data management, staging, replication, and provenance, therefore a number of file catalogues and meta-data catalogues have been developed. In certain grid environments advanced network file systems are used, such as GPFS [59] or AFS [60].

Functional issues: Databases, replication, transactions

Scientific grids have generally emphasized the use of “flat files”, in other words standard files stored on typical file systems. In an enterprise domain, data bases are a much more common source (and destination) for application data. While life-sciences typically source data from databases, this is usually in a “read-only” mode, which can be easily parallelized. In contrast, the high volume of data base write operations in an enterprise context requires the use of atomic transactions, to ensure database consistency. This is coupled with a need for high performance access to geographically distributed servers and users. The result is a requirement for replication, load-balancing, and transaction management across an organisation’s distributed data assets and applications.
Current enterprise data management systems (e.g. Oracle 10g [94], Sybase Avaki EII [97], Data Synapse Grid Server [28]) are custom solutions with limitations on scope, interoperability, scalability, and user control. To date, there has been limited effort to develop standards for access and control of data base assets in a grid environment, outside of the work of OGSA-DAI which provides a common interface to different structured data sources. The OGF working group DAIS (Data Access and Integration Services) has supported the development of OGSA-DAI, and has also been responsible for the development of the WS-DAI (Data Access and Integration) standard. Enterprises require greater standardisation to encourage system vendors to provide systems which can easily be interfaced or replaced, and additional research into the area to push forward the state of the art for large scale distributed transactional data management. It is expected that the BEinGRID project [101] will provide feedback relating to the detailed needs of enterprises from their databases when deployed in a grid environment.

Functional issues: Application development

Application development methodology and tools are critical to support fast and competitive development, maintenance, and evolution of enterprise systems. This is a very rich area of research, particularly active in the domains of Service Oriented Architectures (SOA), of workflow languages and engines, of component-based applications.
Service Oriented Architecture provides both the user view of services (clients can discover the services that fit their quality requirements), and the infrastructure technology enabling a scalable and safe building of complex software (components are composed from smaller components, and export the services they provide).

A key area for grid adoption is a rich set of tools, libraries, and application development environments. The Globus Toolkit [10] has provided grid libraries for almost a decade and is now the premier provider of WSRF and “classic” grid services and libraries. OMII (Open Middleware Infrastructure Institute) [66] has acted as a clearing house for grid software, providing an environment for testing, quality control, packaging, and release of numerous grid services and libraries. The NMI VDT (Virtual Data Toolkit) [67]

specifically provides “hardened” versions of Globus, Condor, and other key grid software packages and is used by OSG and EGEE.
The European CoreGrid research network has worked on the GCM (Grid Component Model) [68], which extends the Fractal component model to grid computing, and provides a basis for software component composition in a grid environment. The INRIA project ProActive [69] has provided a rich Java-based grid application development environment which simplifies the development of parallel distributed applications to allow them to execute equivalently on a single system or with distributed/remote objects across a grid. The GridCOMP FP6 [68] project takes the Grid Component Model as a first specification, and use the ObjectWeb ProActive Open Source implementation as a starting point and reference implementation. ObjectWeb ProActive Grid middleware ensures interoperability with other standards: EGEE gLite, UNICORE, NorduGrid, Globus, LSF, PBS, SGE, Loadleveler, Web Services, etc.

Moreover, several OGF working and research groups are developing APIs to grid enabled applications, like the DRMAA (implemented by SGE, GridWay and Torque among others), SAGA (Simple API for Grid Applications) or GridRPC.
Several ongoing European projects are contributing to this effort. A significant example is from the g-Eclipse project: g-Eclipse delivers at the end of the project a middleware independent access tool to Grid infrastructures for grid end users, grid operators and grid developers. As such a generic toolbox, the g-Eclipse framework has the potential to provide enhancements to all mentioned business scenarios. G-Eclipse does not offer services to the middleware or scenarios, but a toolbox to use and to contribute to. With its middleware (and therefore vendor independent) approach, g- Eclipse is not limited to one middleware nor to one infrastructure. By following the fat-client approach g-Eclipse is not limited to one protocol like HTTP.

A different approach, well installed in business service infrastructures, but also in some kinds of scientific grid environments, is to specify the combination of tasks required to provide a service a workflow. This is less powerful, but simpler and easier to use and to maintain, than direct programming using a grid middleware. There exists approximately 40 different workflow languages today, and as many workflow engines (see a good classification of grids workflows in [124]). Workflow interoperability standards are certainly a major task in the forthcoming years.

Commercial issues: cost and revenue management

Monolithic commercial systems all provide some degree of resource metering and charging, at times with features such as variable rates and rich account models. These allow an enterprise to monitor and control resource usage, and charge-back IT infrastructure costs to the users (or departments). What is lacking is a common model for specifying charges for a service and a standard accounting system which allows a collection of services to be linked into a common infrastructure. These requirements have been touched on in WS-Agreement where the SLA contains charging details; however the specific mechanism for accounting and billing is not covered.
Some specific projects have included economic models in the scheduling framework, or provided services for charging resource usage against accounts however none of these

have gained wide-spread adoption. The OGF Resource Usage Service Working Group (RUS-WG) [116] have developed a service interface for resource usage accounting, but do not consider charging. EGEE have developed the Distributed Grid Accounting System (DGAS) [117] which analyzes server logs to determine what usage has taken place and to provide a rough means for managing quotas, albeit not in real-time. The GRIA project [119] has specifically looked at the accounting and billing facilities provided by different grid middlewares in a report for NextGrid [118]. The Gridbus project [121] has, for many years, been developing grid services which utilise economic models and consider cost and charging. More recently, the EU SORMA Project [120] has been started to investigate economic models for resource acquisition and composition, taking into account resource usage costs and market-driven resource pricing. Part of this work has included the Tycoon [122] market-driven resource allocation service developed by HP and this is now being incorporated into the gLite software used by EGEE.

Dependability and Performance: self-healing, self-optimisation

Autonomic Computing [88] aims at offering features for self-healing, self-protection, self-configuration and self-optimization for service-based or component-based applications. It is currently a very active research area, in particular within EC funded projects, and that will definitely constitute in the medium term important foundations for addressing enterprise grids dependability, performance, manageability, and flexibility requirements.
Self-healing systems adapt themselves at runtime to handle such things as resource variability, changing user needs, and system faults. More and more systems have this requirement, including e-commerce systems and mobile embedded systems.

The GridComp project [68] aims at providing complete autonomic management of non-functional features of software components related to dependability (e.g. fault tolerance / fault handling), or to performance tuning and optimization (including policies and methodologies adapting service execution, in terms of resources used, to the current service request rates).

The K-Wf project provides permanent monitoring and performance prediction (based on online measurements and historic data, made accessible through ontologies) and consequent scheduling and/or re-scheduling of workflows execution.
In a similar goal, the Grid4All project provides mechanisms where risk indicators warn the provider that failures could occur in the near future, so the provider can start negotiating alternative resources or activate self-organising fault tolerance mechanisms

Security: access control, authentication, infrastructures

The use of X.509 [98] certificates as a common mechanism for identity management and system/service security is dominant in grid computing and largely an outcome of the Globus X.509 proxy certificate concept, which has since been standardised by the IETF in RFC 3820 [47]. The complexity of generating and working with proxy certificates has been mitigated by the MyProxy [48] project which allows remote delegation and management of X.509 certificates. The use of Attribute Certificates [49] is also gaining in popularity, as users find they need to distinguish between different roles and VOs within a grid domain. The VOMS [50] system from EGEE [15] and PERMIS [51] provide VO and Attribute Certificate management functionality.

Many academic grid projects work within trusted domains, relying on PAM/LDAP based shared authentication between systems, or the use of DSA/RSA keys (from SSH), while other web-portal based grid environments hide the grid behind the portal and only require web username/password for access control.

Regarding security policies and access control management, these have generally been done with per-VO task queues and manual configuration and synchronisation of which VOs are permitted within a site, and which users (via X.509 DNs) are part of a particular VO. Some work has been done to control access via XACML [52], SAML [53], and GACL [54], and via the OpenPERMIS tools. XACML and SAML suffer from a complex notation and lack of tools for constructing, checking, and interfacing to security policies. GACL has been developed out of the GridSite [55] project which has also provided X.509-based wiki system, HTTPS file system, and a modification of the Apache HTTPD project's suexec program named "gsexec" which provides *NIX account switching based on X.509 certificates. This combination of tools encounters increased usage within the EGEE community.

While the tools and standards listed above are used widely in the services and grids infrastructures, there is still a large area open to research when it comes to guarantee security across all levels, from the network and the software infrastructure to the business level. In this domain, The EU GridTrust[123] project has provided a detailed study of various issues and possible solutions around security, virtual organizations, and trust management in grids from the perspective of industrial requirements. It is clear from these reports and other studies that more work is required in this domain to enable enterprise-grade secure grid systems to be used, in terms of policy management, roles, service access, and service composition.

Interoperability: standards, bodies, data and application deployment, software management, task management.

The Open Grid Forum [6] is the main standards development organization devoted to grid computing, with a focus on standardizing a service oriented architecture based on OGSA (Open Grid Services Architecture) [7] [8]. OGSA defines relevant services such as logging, reservation, and workflow, and formed the basis for the initial OGSI (Open Grid Services Infrastructure) proposal, which was an extension to WSDL and existing Web Service standards. This evolved into WSRF (Web Service Resource Framework), providing compatibility with existing Web Service standards. The open source *Globus Toolkit* [10] is a widely used middleware package used to build scientific grids. It provides mechanisms for remote process invocation (GRAM – Grid Resource Allocation and Management), data transfer (RFT – Reliable File Transfer – and GridFTP), scheduling (GridWay, that implements several OGF standards, namely: DRMAA (Distributed Resource Management Application API) and JSDL (Job submission description language), monitoring (MDS – Monitoring and Discovery Service), and security (via X.509 proxy certificates). A few Web Service standards from outside of OGSA are being explored and developed by the grid community, specifically WS-

Agreement [39], which provides SLA functionality for service negotiation, and WS-Context [40] which provides a mechanism to track session state between WS invocations.

One of the key issues in grid operation is the deployment and management of software, both in terms of persistent grid middleware services, and short-term services or applications required for a particular user group (VO) or individual task. Several different standards have been proposed to deal with this. WSDM (Web Services Distributed Management) [61] from OASIS presents standard interfaces to be implemented by conformant Web Services to allow them to be controlled via a common system. The OGF standard CDDLM (Configuration Description, Deployment, and Lifecycle Management) [62] provides a model for configuring Web Services, then managing their deployment and overall life cycle. This is a standard specific to Web Services.
The ETSI TC GRID has started work to standardize the Grid Component Model (GCM), developed by the CoreGRID NoE project for describing grid software components. The first part of this standardization will cover interoperable deployment of grid software components or applications, via a standardized format for different computing environments. For computational grids, the OGF standard JSDL (Job Submission Description Language) [63] is being widely adopted as a common task description language. It provides an XML syntax which identifies the key properties and requirements of a grid "task", thus allowing it to be appropriately scheduled and then subsequently launched. Task/application execution and scheduling are being standardized via the OGF BES (Basic Execution Service) [64] Standard, and the OGF GSA (Grid Scheduling Architecture), respectively.

Governance: Service Level Agreements

For the enterprise grid vision to be realized, it is necessary to establish, either statically or dynamically, some kind of service level agreement (SLA) between interacting services. The leading work in this area is the development of the WS-Negotiation and WS-Agreement standards from the OGF Grid Resource Allocation Agreement Protocol Working Group (GRAAP-WG) [102], and is being evaluated by the AssessGrid project [104]. A number of major industry players (IBM, Microsoft, BEA, SAP) have also formed a Web Services SLA framework named WS-Policy.

The HPC4U project [106] works towards an implementation of the WS-Agreement protocol for realizing SLA negotiation over the Globus Toolkit. The CCS resource management system has been enhanced for being able to drive such an SLA negotiation and respect the terms of negotiated SLAs in the system management and scheduling.

The FP7 EU project SLA@SOI follows a holistic approach to SLA management by supporting SLAs at business, service and infrastructure level also supporting the translation and mapping between different levels. Furthermore, it investigates SLA-aware infrastructures that exploit virtualisation techniques for dynamic readjustment.

Manageability and Flexibility: reconfiguration, autonomicity

The *Open-World* [89] assumption refers to systems running in an unknown and changing environment, demanding techniques that let software react to changes by self-reorganizing its structure and self-adapting its behaviour. The basics tools addressing this challenge can already be seen in the current trends of Web-services, service description standards (WSDL), service composition by workflows (e.g. BPEL), discovery, dynamic selection (substitutability notions), publish-subscribe infrastructure, etc.

Self-reconfiguration appears in current projects in various context, e.g. in component-based infrastructures like the CoreGRID Network of Excellence. There are also advanced workflow approaches, e.g. with WS-BPEL, where monitoring and reaction rules are specified to oversee the execution of Business Process Execution Language workflows and adapt to changes in the environment [78]; or with GWES, developed by the K-Wf project, providing dynamic configuration/re-configuration of workflows during run-time depending on results of the execution so far, or incoming new/additional information.

The EU AssesGrid project works on the integration of risk assessment and risk management modules as well as adaptation of the WSNegotiation manager. These developments address manageability by detecting bottlenecks in the provider infrastructure. The risk indicators are used to find out, which parts of the systems or networks are crucial for SLAs fulfilment and should be replaced. They also address flexibility in the sense that risk indicators warn the provider that failures could occur in the near future, so the provider can start negotiating alternative resources or activate self-organising fault tolerance mechanisms.

However, the current technology doesn’t support expression, at design time, of the requirements and constraints to be fulfilled at runtime in the discovery and selection phase to identify the services to be bound. Specifications of substitutable components can be classified in three layers: typing (of interfaces), protocols (dynamic compatibility), and semantics (ontologies). Afterwards verification of substitutability can be performed off-line using model-checking based engines, therefore ensuring the correctness of service composition; alternatively it can occur at runtime, as in the case of continuous testing [91], or even be performed by dynamic service adaptability. Very few commercial offers are available yet in this area.

4.3 Technology Trends

This section presents some emerging technology trends which are supposed to significantly impact future business IT systems in general and Business Grids in particular.

Storage and Data Management

Storage provides Business Grid applications with the ability to maintain persistent data and retrieve it as and when required. There are several distinct kinds of storage system used today, with different levels of performance, capacity and cost. Direct attached storage (DAS) has a storage device essentially as part of a host system, the simplest example being a server with an internal hard drive. Network attached storage (NAS) offers greater scalability and flexibility by separating the storage from the servers. It provides file-based storage to a number of servers from a specialised storage device over the same IP network used to interconnect the servers. Storage area networks offer block-level storage and generally use a special dedicated network to completely separate data storage and data processing. This separation and the use of optimised protocols give high performance and reliability.

Storage demands and capabilities are increasing very rapidly. Terabyte capacity hard drives are already available. Applications with requirements for petabyte storage are beginning to emerge in science, medical imaging and digital media. The need to comply with regulation (e.g. Sarbanes-Oxley, Basel II) is driving businesses towards large-scale storage of all relevant financial records, including email.

Data volumes in typical companies are growing at around 50% every year, and budgets for storage are running ahead of those for other IT hardware. In addition there is a clear need for storage and data management systems so that organisations can keep pace with the rapid growth in volume and complexity of usage. Approaches such as Hierarchical Storage Management (HSM) and Information Lifecycle Management (ILM) are becoming widely adopted.

Processor Technology

As an approach to maintain the rate of improvement in available computing power, major vendors (Intel, AMD, and IBM) have all recently moved towards multicore processors where several processor cores are linked within the same chip. Dual-core x86 processors are already commonplace with quad-core available and 8-core expected in 2008-9. Sun is already shipping the 8-core Niagara chip. Tens or hundreds of cores are anticipated in the future. Multicore processors offer benefits in processing power with very low latency, and also in power dissipation, which is becoming increasingly important.

Virtualisation

The last few years have seen the development of a range of distinct approaches to improving the flexibility and manageability of ICT infrastructure which can all be described as virtualisation. They all involve breaking the tight coupling between an application and the specific resources it requires, but this can be achieved in a number of different ways, involving innovation in processor architecture, operating systems and software.

- Network virtualisation in the form of virtual private networks (VPN) and virtual local area networks (VLAN) allows different users or applications to be isolated from one another even though they share the same physical connectivity.
- Storage virtualisation supports hierarchical storage management (HSM) and information lifecycle management (ILM) which are becoming increasingly important as data volumes explode in both commercial and scientific applications.
- Server virtualisation includes both clustering (building a large virtual computer from a number of smaller ones) for horizontal scaling and partitioning (running several independent instances of an operating system on the same computer).
- Component containers (such as application servers, EJB containers, web servers etc) can also be regarded as offering virtualisation – components written for a

particular kind of container can be deployed to any instance of the container, regardless of the underlying server architecture or operating system.

All of these approaches can contribute to building flexible business Grid infrastructure that can be adapted to the changing requirements of applications. There are many examples of enterprises obtaining significant business benefit from more efficient use of their internal resources by exploiting virtualisation technologies.

Network Connectivity

As Business Grids develop, the role of networks to interconnect resources in different locations becomes increasingly important. Wide-area fixed networking is increasingly being dominated by the use of packet-based technologies such as IP. Access networks are relatively diverse and include a range of copper, optical fibre, radio and satellite based technologies. The balance between these types of network varies considerably within and between countries. Local area networking, within a single datacentre for example, typically uses Ethernet (with Gigabit speeds now common), although there are specialised protocols used for high performance cluster interconnects or storage systems. These networks can have very varied characteristics for bandwidth, latency, jitter, reliability etc., with significant implications for application performance.

Mobile Devices

Mobile devices such as PDAs, mobile phones, portable media players, sensors etc. are becoming increasingly important constituents of the global networked ICT ecosystem. Mobile phone users are far more numerous than Internet users on a global scale. Despite relatively slow advances in battery technology, mobile devices have increased rapidly in their computational capabilities (processing, memory, local storage) and are often able to use multiple approaches to network connectivity (2G/3G mobile, WiFi, Bluetooth, for example). Incorporating these into Business Grid scenarios typically means that their context (e.g. geographic location, connectivity characteristics) have to be taken into account.

Sustainability

Environmental issues, and energy consumption in particular, are rapidly becoming very important for ICT infrastructure. Demand for computing capacity continues to increase rapidly with a corresponding increase in energy requirements for IT hardware and associated cooling systems. This has implications for operational costs and also for sustainability and climate change. Businesses are beginning to take corporate social responsibility seriously, with regulation and legislation expected to grow in importance.

The Green Grid is a consortium which has recognised the importance of energy efficiency in data centres both for business and the environment – identifying it as the most significant issue facing technology providers and their customers today. The cost of energy is projected to exceed 50% of total IT budget within the next few years and security of supply for large datacentres could also become a problem in many locations. The solution to this will involve both the introduction of more efficient hardware and data centre architectures, and integrated management solutions that take into account ambient temperature and energy consumption. Heat reuse for offices and homes could be included as part of an integrated solution. Combining power-hungry IT facilities with sustainable generation (hydroelectric, wind, solar etc) could also prove to be effective. Control of energy usage will be an important feature of an IT infrastructure and being able to demonstrate high efficiency should encourage use of Business Grids over standalone deployments.

5 Research challenges

In this section we describe the main research challenges resulting from the business scenarios and technology trends presented above. We follow the categories introduced in Section 3 and describe research challenges based on work underway in current research projects as well as on the views of experts working in the area.

5.1 Functional & Commercial Issues

Traditional grid systems focus on isolated batch jobs while virtual-machine based grids address complete systems to be submitted. However, Business Grids focus more on interactive and transactional applications with a low response time. Furthermore, individual applications are not completely decoupled, in the sense, that they are executed as part of business processes or coexist with other collaborating applications. Overall, transaction processing is a central design concept for many business systems, and hence should be directly supported.

The research challenges introduced in this section are most important for the range of hosting scenarios but also increasingly important for the (hierarchical and extended) enterprise scenario.

Data Management

Data needs to be managed throughout its lifecycle. In outsourcing or collaborative scenarios, flexible access controls must be offered. Many of these considerations will have implications on the type and configuration of the underlying storage systems. Data management systems need to be able to handle heterogeneous underlying resources and apply them effectively to support business processes and applications.

In large-scale storage systems, there is a high probability that individual hardware components will fail. Likewise, network partitions may render resources temporarily of permanently unreachable. These events must be regarded as normal and data safety guaranteed regardless.

Different policies for retrieval, assured immutability, confidentiality, privacy and compliant deletion may need to be applied at the level of individual items within a collection of many billion, and may change over time. Data and storage access is an essential part of many applications. Business Grids must be able to offer well-defined storage and data management services that can support these applications. Data must be available in a timely way to applications to avoid performance bottlenecks. Data management, application execution management and networking are highly interdependent. Location, partitioning and replication of data must be coordinated with the deployment of software components onto heterogeneous computing resources, and comply with any legal or regulatory constraints. File systems which can accommodate high and variable latency, complex failure modes and cross-organisation operation (including VOs) will be required. Distributed file systems will need to support multiple writers. This leads to particular challenges in dynamic and distributed version management.

Applications Development

The trend towards multicore processor architectures is clear. However, programming multicore systems to take full advantage of their capacity requires changes in the way software is written. Software needs to be written to exploit the concurrency inherent in new processors. Similar changes in software development are also required to ensure that applications run efficiently on distributed Business Grid infrastructure as this becomes the standard deployment environment for enterprise applications. Extending enterprise Grids to include mobile devices (PDAs, mobile phones, sensors, cameras etc.) and pervasive scenarios involving embedded processing capabilities means that software must be able to make use of a wide range of resources. This adds considerable complexity to the applications development activity. A major challenge is to support developers in producing effective, robust applications for this environment while minimising their awareness of the underlying complexity. One goal of moving towards an open infrastructure is to make it easier for a wide constituency of developers to write and deploy applications and services. The ability to use accessible programming and scripting languages (Python, PHP, Ruby etc.) is important in achieving this, so interfaces to Grid capabilities in a range of languages, via Web Services or otherwise, will be required.

Making Grid applications and services available in a dynamic, distributed infrastructure involves considerably more than the normal development and installation process. Effective, automatic mechanisms for testing, distribution and deployment are required. Applications development will need to include specification of preconditions and policies that can be used to configure these mechanisms.

Network Connectivity

Applications will increasingly be built from a number of concurrently executing components – whether these are Web or Grid services or parts of an application built for multicore or multiprocessor hardware. Communication patterns between these components play a vital role in determining end-to-end performance. Relevant networks may include high-speed cluster interconnects, local area networks ranging from megabit to gigabit speeds, wide-area networks with a wide range of bandwidth and latency and eventually intermittently connected, low bandwidth wireless networks.

Handling mobility in a dynamic infrastructure represents an important challenge. Different kinds of mobility must be considered: terminal mobility between different networks or coverage areas, personal mobility between different terminal devices and session mobility so that a user can change terminal and network while retaining service continuity. Adapting application behaviour to make best use of the available network (and other ICT) resources needs to be supported.

In addition to physical network characteristics, communication middleware plays an important role. This must be appropriate both for the application and also for the underlying networks. An effective Business Grid needs to be able to match the requirements of a particular application to the resources available so that acceptable performance can be achieved. Heterogeneous network connectivity and communication patterns have played a relatively minor role in most current Grid deployments.

Accounting

Though many grid systems provide basic means for accounting, this typically does not allow associating accounting data with related business activities or decisions. This significantly restricts the cost transparency of IT operations which should be achievable at individual project level. Solutions for precise accounting across all layers of an IT infrastructure (including hardware and software environments and licensing models) may be required at individual project or customer level. This will support the customer in deciding whether to use a particular Grid services in a business process. Appropriate accounting at business process level will allow the business value to the customer to be made clear. In addition, the provider can use TCO (total cost of ownership) calculations as an effective tool in capacity planning. In a global environment, accounting services will need to deal with multiple currencies.

Service Level Agreements

SLAs in current Grids typically relate to technical properties of the solution rather than being expressed in terms of the business requirements of the customer. This makes it difficult for the customer to know whether a particular SLA will actually meet his needs, as well as limiting the flexibility of the service provider and hence increasing operational costs.

SLAs need to include all the information that defines the shared context between service provider and service customer. In particular, they should include metrics that are relevant to (and measurable by) the customer and associated target values, pricing for the service and any penalties or compensation to be applied in case of SLA violation.

Reliable translation of business properties to technical properties and vice versa is required. Ideally this translation within the service provider domain will be automated so that service instances can be configured to meet the needs of the customer while making efficient use of resources. A prerequisite for this is that SLAs and policies are represented in an unambiguous syntax suitable for computer processing. Standard ways of describing Grid infrastructure and software components are also needed. These information models will need to be flexible and extensible so that they can accommodate a wide (and constantly changing) range of different usage patterns and infrastructure components. Semantic-based approaches for advertising and finding suitable services will be required, extended later to include automated negotiation.

Effective monitoring of parameters relevant to the SLA terms are important both to demonstrate compliance with the SLA and as input to management systems which can take action to maintain SLA terms within acceptable limits.

5.2 Dependability

In today’s systems dependability requirements are typically satisfied with a static solution and most often only at the level of infrastructure resources (e.g. by assigning a number of redundant nodes or network links to a particular activity). The dependability characteristics of the solution are not directly linked to application (or business process) level requirements and consequently the economic costs and benefits of a particular approach to providing dependability are hard to quantify.

Business Grids aim for dependability levels appropriate to the needs of individual customers. Consequently, the most important research challenge in this area is the automatic derivation of hardware configuration from customer requirements and the aggregation of the resulting costs so that different approaches to providing dependability can be compared. While cost awareness is highly relevant for all business scenarios, complete automation of the requirement mapping is of particular relevance for hosting scenarios where the provider needs to offer business value to his customers as efficiently as possible.

An important characteristic of a Business Grid is that it should be highly resilient. This implies that node failures (or withdrawal of nodes/sites, congestion or reconfiguration of network connectivity etc) have minimal impact on the overall performance.

Automated system management based on the so-called self-* properties (e.g. self-configuring, self-healing, self-tuning) aim at ensuring maximum service availability and performance when some components of the infrastructure or application fail or become disconnected (the normal state of any large Grid deployment). Exploiting both physical and virtual machine based technologies for supporting these properties is required. Flexible monitoring across multiple resource types should be used to respond rapidly to failures or even to identify impending failures before they happen. Once identified, mechanisms for automated fault mitigation – making good use of resources that are still functioning correctly – and recovery are needed.

Special attention must be paid to resources provided by mobile devices for the extended enterprise scenario. These are inherently less dependable components as a result of restricted or intermittent network connectivity and limitations associated with battery power. Predicting the behaviour of mobile nodes could be supported by the ability to monitor connectivity and other properties of mobile nodes in a consistent way.

5.3 Security

At the moment, Grid security mechanisms target traditional infrastructure resources (e.g. compute nodes, files, devices, .etc.) in a rather static way and do not consider the overlaying organizational structure. Enterprise security mechanisms are focused on organisational and user roles but make (implicit or explicit) assumptions about the underlying infrastructure topology which is subject to change in grid environments.

Business Grids need to provide security mechanisms for possibly virtualised infrastructure resources (e.g. a virtual node or a distributed grid job) with a granularity and flexibility that allows the reflection of higher-level security zones such as a process group, a business process or organizational entities including departments, customers and partners. Consequently, Business Grids require advanced security models and policies that can accommodate variability in both organisational structures and infrastructure topologies. This is most important for hosting scenarios, where systems process applications on behalf of competitors within a single operational domain. In this context, the ability to achieve complete isolation at any level (application to infrastructure), wherever required, is necessary. The dynamic nature of Grids means that the separation of resources into clear security layers is a major challenge. Possible conflicts between policies relating to local and remote resources albeit within a single logical layer will need to be addressed.

The security properties and performance experienced by a customer should be independent of any other services the provider may be offering to other customers. These challenges are synchronized with the ones stated by the European Security Forum for Web Services, Software and Systems (ESFORS).

The scale and the scope of Grids necessitate security models that can provide certain security assurances even in the presence of adversaries. They should be reactive and protective in nature yet adaptive to the situations where menaces are unavoidable. Development of low-level security mechanisms at the infrastructure level could be an effective way of addressing external threats (e.g. from the Internet). Such deeply integrated mechanisms may support traditional techniques such as network-level firewalls or might even make them obsolete.

Access control in dynamic virtual organisations and virtual hosting environments is an important issue. Considerable work has been done on authentication mechanisms for Grids. Effective authentication solutions are now available for Grids. However, little research has been done on improving authorisation mechanisms for dynamic virtual organisations, i.e. where users and resources come and go dynamically. Most of the existing mechanisms can be classified as static and/or coarse grained. One promising approach for improving authorisation in virtual organisations is usage control [71]. Usage control provides a model for fine grained and continuous control in Grids. The ability to check compatibility between user and provider usage control policies is required in order to support resource brokering.
Data security is an important element of any Grid service. It may make use of encrypted data handling and require scalable key management solutions for access control. Security features are required even for the protection of non-confidential data. Such data still requires its integrity to be maintained.

Research is also needed to develop forensics models for Grids to identify the causes of failures. This aspect of security modelling is important not only for the determination of the losses and liabilities but also for the determination of the preventive steps to be taken to avoid the similar breaches in the future.

Business applications require quantitative measures of the quality of protection. It is important to develop suitable metrics so that quality of security services (QoSS) can be included in SLAs.

Reputation and trust are important issues in any scenario that involves more than one administrative domain. Dynamic, lightweight interactions will need to be based on trust relationships, possibly supported by reputation rather than complex contracts. Mechanisms for distinguishing between providers on the basis of the quality of their delivery are required. These could be based on the shared experience of others as a means of establishing reputation at least for an unknown provider. Techniques for quantifying and managing the risk associated with dealing with unknown entities represent an important challenge. Solving this could allow appropriate security mechanisms to be selected at runtime for each interaction, balancing the risk associated with low levels of security against the performance overhead associated with highly secure operation.

5.4 Performance

Virtualisation involves hiding the differences between heterogeneous resources, across different resource types, between vendors and across time. It allows decoupling of computing hardware from applications and management systems. Hardware and software can then follow independent evolution paths since changes to one can have minimal effect on the other. The ability to rapidly configure and reconfigure virtual computing resources can be used to support a flexible resource pool for a wide range of applications. These are clearly attractive features for a service oriented infrastructure as envisaged for Business Grids. There is currently still a performance penalty associated with the use of virtualised computing resources although this is reducing as the technology improves. A more significant issue is that use of a virtualised infrastructure increases management complexity because the way applications are mapped to hardware can now be variable. This makes it much more difficult to reliably predict specific performance metrics of the system in advance. This results in unpredictable costs for a given performance level, as the resources required to meet a SLA can only be estimated very roughly.

Scheduling in current Grid systems is typically carried out for a number of different kinds of resource (e.g. processing, storage, network transport) with a low degree of coordination between them. The consequence is that decisions made by one scheduler may conflict with the decisions of another. An integrated approach which enables coordinated scheduling would reduce execution time, balance workload more efficiently across available resources and provide a wider range of management options in case of failure or degradation of performance. Dynamic adaptation and scalable scheduling are particularly relevant to several of the more complex enterprise Grid scenarios.

Performance prediction is necessary in order to apply advanced scheduling techniques. This needs to take into account the issue of predicting the performance of the overall system based on the characteristics of its constituent parts and also the logical boundaries in hierarchical and extended enterprises and in hosting scenarios, particularly where resources are shared by many independent applications.
Interactions between applications sharing the same resources are unpredictable in general and in a virtualised environment it will be infeasible to test all possible combinations in advance. Support for predictable isolation at the resource level is therefore required to achieve predictable performance. Combining monitoring with performance prediction and using the results to determine (re-)scheduling decisions will allow dynamic changes in workflows (or business processes) to be supported.

5.5 Interoperability

Today, organisations are adopting Grid solutions and achieving significant benefits even without having all the standards in place for interoperability. This is particularly true for organisations running a single version of one of the popular grid middleware software products. As Grid solutions become more widely adopted, the need for interoperability and standards increases. Interoperability becomes more critical as organisations connect Grids within their organisations or with other organisations that utilise different Grid middleware.

Current Grids offer restricted interoperability for two main reasons. Many applications, particularly business applications, will not run on a Grid without modification. Even if the executable software components are compatible with Grid (e.g. batch jobs), schedulers generally require an associated description of each component to be provided so that each component can be run on a compatible computer with the right parameters. Each scheduler typically has its own format for this description. Second, there are several different Grid middleware systems (and research infrastructures) which cannot easily be combined within a single Grid as they are based on different standards or have inconsistent behaviour.

Dynamic reconfiguration is another area where the heterogeneity of current systems limits interoperability. In the context of live migration of virtual machines, for example, the challenges include migration between subnets, between hosts that do not share access to common storage, and between hosts with different configurations (e.g. configuration of virtual switches).

Business Grid research should address the prospects running business applications on a Grid infrastructure without modification. In addition, development and adoption of open standards that support the interoperability of Grids based on different middleware is required. This is especially important for the hierarchical and extended enterprise scenarios where independently developed Grids based on different middleware may need to be integrated.

A particular challenge associated with interoperability is the need to go beyond harmonisation of message exchanges, which should be relatively straightforward. Mediation between different semantic models for non-functional aspects such as security, dependability, management information (e.g. monitoring data and policies) and performance is essential if the behaviour of the combined Grid is to be predictable and acceptable.

Open standards will be an important means of achieving interoperability between different Grids. There is already considerable work underway in a number of standards development organisations aimed at solving technical problems and defining standards for Grid interoperability. This needs to be better coordinated and supported by robust test methodologies so that users can have confidence that interoperability between different Grids across organisations can be achieved. Reference implementations will also have a valuable role to play in promoting the use of open standards and interoperable solutions.

5.6 Manageability

Management and system administration are still major cost factors in today's business systems. In particular, many management tasks are not yet extensively automated. Grid systems provide some support for automated management of infrastructure resources but typically follow a narrow perspective that simply covers the management of resources as assets rather than looking at how they fit into overall business solutions.

Infrastructure architectures and their management systems need to have the flexibility to incorporate new kinds of server and special-purpose devices alongside existing hardware.

Virtualisation technologies applied across many resource types can provide the abstractions and interfaces required for an integrated and automated approach to management. Realising this promise is a significant research challenge. In particular, a critical research area is in management of the deployment contexts of both individual virtualized resources and sets of resources. A common approach to this contextualization will be a powerful tool for management.

The traditional approach to systems management has relied on human intervention to manage complex situations. However, this is now becoming untenable in large enterprise deployments and will be even more so in future Business Grid scenarios.

There is a need to monitor and control on timescales relevant to individual application sessions whose duration could be on the order of seconds. Control decisions, particularly in a virtualised infrastructure must be made more rapidly than manual systems can achieve. This may be particularly important in extended enterprise scenarios where the environment of mobile nodes is subject to change on short timescales.

Some improvements are anticipated in management of a single administrative domain – lightweight mechanisms for collecting information about the state of the managed Grid and controlling it as an integrated infrastructure.

However, management information will not, in general, be completely shared across administrative boundaries. Dealing with this fact represents a major research challenge. There will be no central point where management decisions can be made with full knowledge of the system state. Automated and decentralised management decision-making is required to achieve the required speed of response, particularly in a large system. Decisions will need to be made based on incomplete information – whether driven by the need to respond quickly or because access is restricted. Linkage to SLAs is important so that the impact of a particular management decision on customers can be determined.

Tools for modelling and simulating large Grids are required in order to ease the development and evaluation of automated management mechanisms and policies.

Appropriate analysis and presentation tools are required to provide visibility of the business performance on appropriate timescales for human decision makers.

These challenges apply to all of our business scenarios. A specific research challenge for hierarchical enterprises involves concepts for mediation and conflict management in distributed environments. Hosting scenarios require special attention on remote management and mass automation support (e.g. cloning of customer systems).

5.7 Governance

Enforcement of governance policies, whether defined within a business or imposed by regulation or legislation, is poorly supported in current systems and requires manual intervention by administrators. In particular, support for complex organisational structures (such as hierarchical enterprise) is almost non-existent.

An important research challenge is therefore the enforcement and tracking of business policies at the infrastructure level. To achieve this, techniques for logging, tracking, and auditing of resources are necessary. In addition, (automatic) translation of business policies to configuration of the infrastructure needs to be addressed. A third issue is on automated mechanisms for enforcement and conflict resolution.
These requirements are essentially the same as those for automating management driven by SLAs. The major difference here is the need to demonstrate compliance via audit trails. Robust audit trails in turn rely on secure, reliable data management so that actions taken to comply with governance policies are recorded in a durable form and it can be demonstrated that records have not been tampered with.
These challenges apply to all business scenarios. However, the hierarchical enterprise scenario requires additional research on how to derive models and techniques to specify and manage a hierarchical structure of policies. The corresponding models must be very flexible such that they can easily be adapted for different enterprises. Finally, hosting scenarios require governance support for both providers and customers.

5.8 Flexibility

Typically, enterprise IT systems are implemented to meet specific requirements with little attention paid to building in the flexibility to support unanticipated future changes. This is understandable if the focus is on satisfying known requirements at the lowest cost. The result is a close association of software with dedicated servers and network infrastructure. This leads to inefficiency across the IT estate and also means that modifying IT systems to reflect changing business circumstances are expensive and time consuming.
Business Grids aim to address these issues by making ICT infrastructure flexible and dynamically configurable. This requires research in two main areas: the automation of information flows and management processes from business process to infrastructure, and the dynamic reconfiguration of infrastructure resources. As management automation issues are discussed in other sections we focus on flexible resources here.

All of the scenarios identified require research in dynamic allocation and release of resources. Migration of software between servers is also required to support load balancing and scaling, both horizontal (changing the size of a resource pool) and vertical (migration to a resource with different capacity). Business applications are often multi-tier (e.g. presentation, business logic, data) so migration is more complex than moving a self-contained software component, since relationships between different parts of the application need to be managed. Migration of an application or its data may be subject to licensing, regulatory or legal constraints which need to be enforced by any automated management system. Management of the scheduling and migration of applications within a flexible infrastructure is a major challenge. An application may be represented as an abstract workflow, independent of the details of the deployment environment. This is invariant to changes in the underlying infrastructure. Mapping of abstract workflow to a concrete workflow (configured on specific resources) will need to be done when the application is first deployed and subsequently whenever a change is required. Semantic descriptions of applications and infrastructure are expected to be important aspects of automating this mapping process.

Autonomic or intelligent control cycles will be very important for distributed infrastructures with decentralized control mechanisms, particularly for resilient large-scale grid. No entity in the system will have full knowledge of the overall state. For many types of operations, each site will have to infer relevant aspects of the state based on observable events. This indicates that event-driven approaches will be required for distributed workflows, cooperating agents in a grid, distributed data management etc.

There are particular challenges associated with large scale flexible infrastructures. Business Grids may support very large and geographically distributed populations of users. They will also support very large numbers of applications on shared resources. Users and applications will be added and removed on short timescales. The dynamic nature of the infrastructure will make it more difficult to keep track of the association between user sessions and physical infrastructure, with implications for fault handling and impact assessment.

Managing security in a dynamic environment is also a challenge. Mechanisms such as firewalls which rely on protecting a boundary between physical resources will not be appropriate. Isolation of individual application sessions from one another is required, even if the resources used change during the session. This includes management of dynamic VOs or VPNs where users or resources can join or leave on short timescales and security policies must be maintained.

5.9 Complex System Modelling and Simulation

Business Grids aim at supporting a range of collaborative scenarios. Eventually we can anticipate the creation of complex and dynamic business ecosystems based on Business Grid technology. Representing (with suitable abstraction levels), modelling and simulating these ecosystems (including both technological and socio-economic issues) will be very important to build an understanding of the new business environment they represent. Failing to do this could mean that businesses do not have the confidence to engage in these ecosystems or find that they are unable to manage and control their involvement effectively.

Research challenges include the development of techniques and languages for systems modelling integrated in advanced context sensitive (domain-dependent) and adaptive development environments enabling a multi-disciplinary approach. This will include operational analysis, operational research and optimisation techniques, stochastic and fractal processes.

5.10 Overarching challenges

Finally, there are some overarching research challenges which cannot be directly derived from the listed business scenarios and business requirements. Those research challenges cover:

Direct contribution to business value

Users of Business Grids should not care about the details of the supporting technology. Enterprise-wide Grids supporting multiple applications (and still more so, Grids that span multiple enterprises) need to offer well-defined functionality and performance, described in terms that make sense to users and at a predictable price. This will allow costs and benefits to be clearly quantified.

Reduction of the complexity of Grid technology

There is a need for the complexity of Grid technology as perceived by users to be significantly reduced. Ideally, the infrastructure deployed to support their applications should become invisible, allowing them to concentrate on their main concerns. Grids should be usable by people who are not aware that they are using a Grid. On the other hand, more expert users should be provided with appropriate levels of abstraction so that they can express their requirements more precisely.

The inherent complexity of the technology will undoubtedly increase as Business Grids evolve and this complexity will need to be managed by system administrators or service providers. Management systems and tools will need to offer high levels of automation, appropriate levels of abstraction and user friendly interfaces if this is to be tractable.

Semantic technologies will be required to handle large scale Grids. The diversity and heterogeneity anticipated means that a single ontology is extremely unlikely. Scalable and efficient techniques for reasoning will be required to improve interoperability in the absence of universal standardisation.

To raise the level of abstraction of a Grid system is, moreover, a key factor to allow also the development of Grid application and, to this purpose, we recognise the need for Domain Specific Languages, tools and development environments that have to hide Grid complexity and simplify the development of business applications for a specific domain. Approaches like software factories, Model Driven Architecture and software product lines and their applicability to Business Grid have to be investigated. Understanding software lifecycle issues in an open service oriented architecture will also be important.

Architecture driven solutions

Both domain-specific and inter-domain enterprise grid scenarios require well architected platforms that create a fabric for business flows, the services that power them, and the software components that enable the services. Interworking within a well-governed, coherent and consistent framework is required to deliver characteristics such as security, dependability and performance end-to-end. The framework should ensure that the ‘service ecosystem’ is:

- well-governed (policy and rules driven, monitored by SLAs)

- well-balanced between the need of flexible business flows and mandatory technology processes
- well-orchestrated at the level of complex 'end-user' services
- well-integrated at the level of software systems powering core and derived services
- well-instrumented at the level of the technology stacks powering the software systems
- well-supported at the level of infrastructure.

Grid architectures will need to address these requirements, and include both business and technical aspects. Interoperable Grid service architectures specific to the business scenarios described above will need to be developed. Accessible reference implementations, possibly based on a reliable, basic open source code infrastructure will also be an important contributor to the wide adoption of Business Grids.

6 Evaluation of Results on Relevance and Timeline of the Business Scenarios

Annex A describes in detail the implemented process for involvement of the community in the second version of the NESS-Grid SRA. The call for contribution has been highly successful, receiving feedback from the 70% of the contacted groups of interest**:**

- Most of groups have submitted the survey
- Other groups of interests have answered the invitation to contribute showing their interest in the results of the initiative and explaining that they are not contributing because the project has just ended, it is not concerned with the business scenarios of Grid applications or the developments are not yet mature enough to provide useful views
- Many in-depth reviews from Grid experts have been received

The surveys for the groups of interest (research projects, infrastructures and middleware solutions) were designed to extract relevant information according to the type of project. The answers to the “open questions” have provided relevant feedback from the community that has been incorporated into the three main sections of the SRA:

- Business scenarios: Questions 2 in all surveys and question 4 in the project survey
- State of the art: Question 5 in the project survey, question 4 in the infrastructure and middleware surveys
- Research challenges: Question 6 in project survey

The answers to the “closed questions” evaluate the relevance and timeline of the business scenarios. The following subsections provide a high level analysis of those results.

6.1 Relevance of the Business Scenarios

Concerning the **relevance of the business scenarios** described in the SRA, Figure 3 represents the percentages of groups of interest which ticked the four levels of relevance for each business scenario. Enterprise, Hierarchical Enterprise, Hosting and Virtual Organization are considered to be the most important scenarios (more than 50% of the groups of interest ticked them as high relevance). However, even the lowest rank scenarios (e.g. Merger&Acquisition) are still considered as medium or highly relevant by at least 40%.

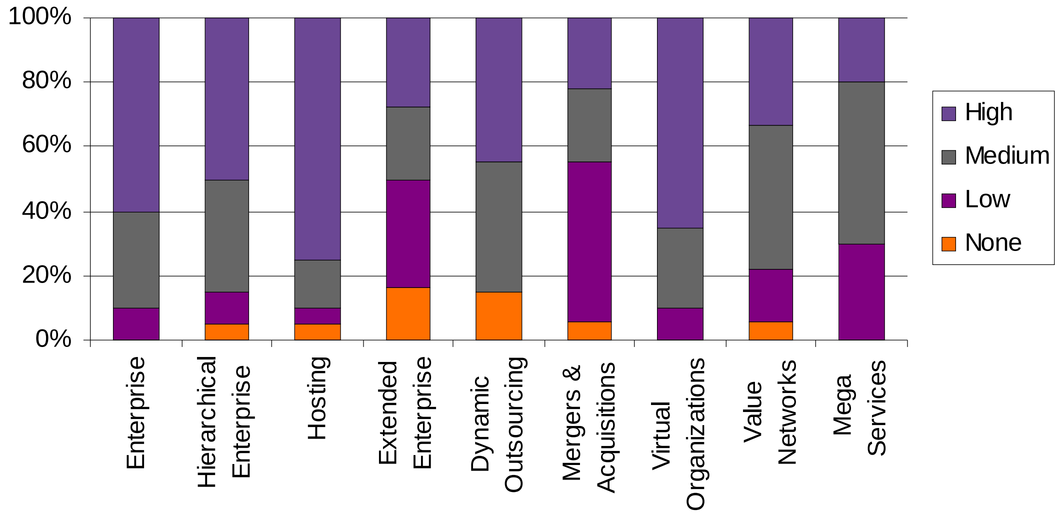


Figure 3: Relevance of the business scenarios described in the SRA

In order to sort by relevance the business scenarios, their relevance can be quantified between 0 (none) and 3 (high) by calculating the weighted average of the results. Figure 4 shows the business scenarios sorted by their average relevance. All scenarios are ranked above 1.6 (medium-low).

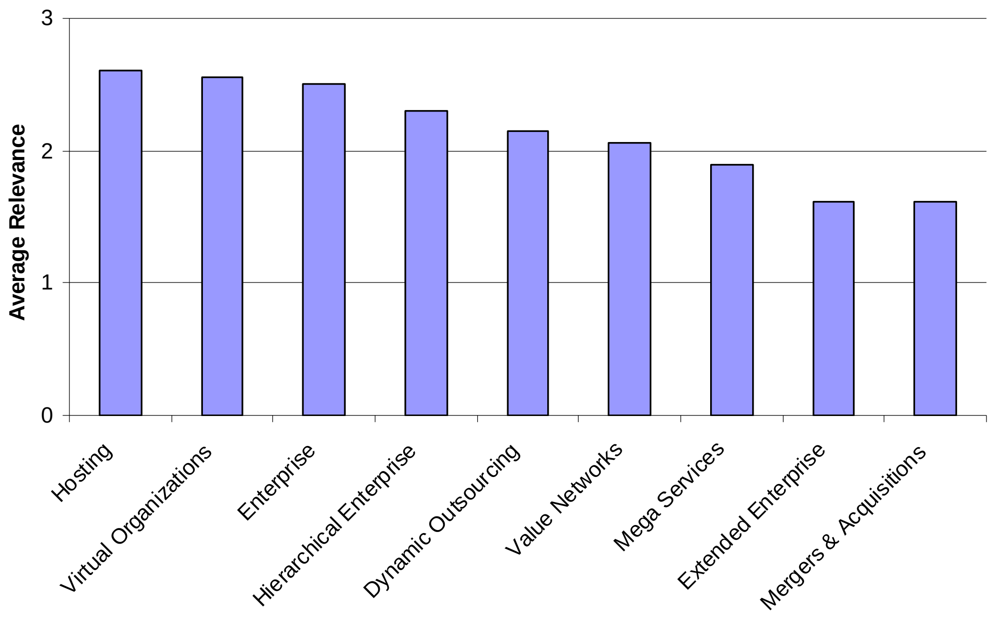


Figure 4: Average relevance of the business scenarios described in the SRA, from 0 (none) to 3 (high)

Concerning the **specific relevance for the surveyed groups of interest**, Figure 5 represents the percentages of groups of interest which ticked the four levels of relevance for each business scenario. Enterprise, Hierarchical Enterprise, Hosting and Virtual Organization are the scenarios with a higher relevance for existing projects (more than

50% of the groups of interest ticked them as high relevance). The rest are scenarios with none to low relevance for the existing projects (more than 50% of the groups of interest ticked them as low or none relevance).

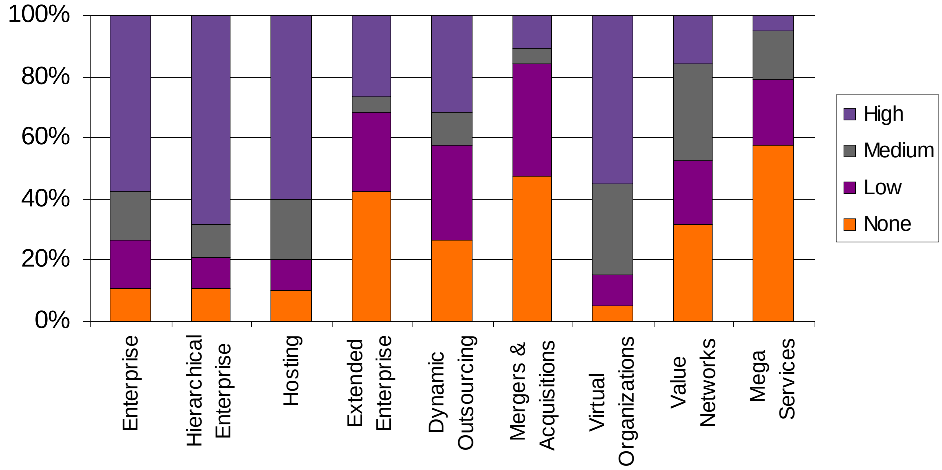


Figure 5: Relevance of the business scenarios in the projects.

Concerning the **percentage of use cases under development in the research projects that could be represented by the different Business Scenarios**, Figure 6 represents the percentages use cases under development in the projects that could be represented by the different business scenarios. Existing projects are developing solutions mainly for Enterprise, Hierarchical Enterprise, Hosting and Virtual Organizations scenarios.

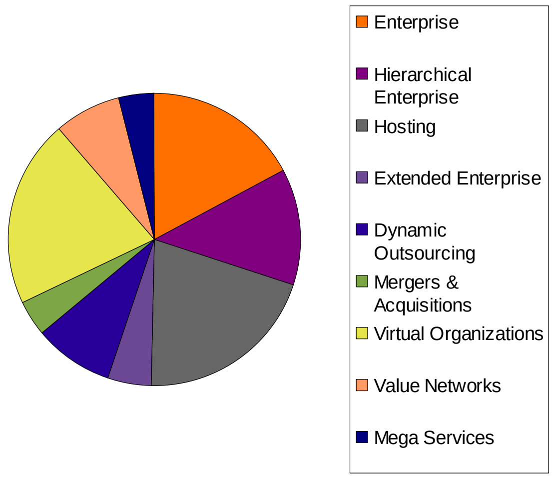


Figure 6: Percentage of use cases under development in the projects that could be represented by the different business scenarios

6.2 Timeline of the Business Scenarios

Concerning the **timeline of the business scenarios** described in the SRA, Figure 7 represents the percentages of groups of interest which ticked the four levels of timeline for each business scenario. Hierarchical Enterprise and Hosting are the shorter-term scenarios (more than 50% of the groups of interest ticked them as short timeline). Mergers & Acquisitions and Mega Services are the longer-term scenarios (more than 50% of the groups of interest ticked them as long timeline).

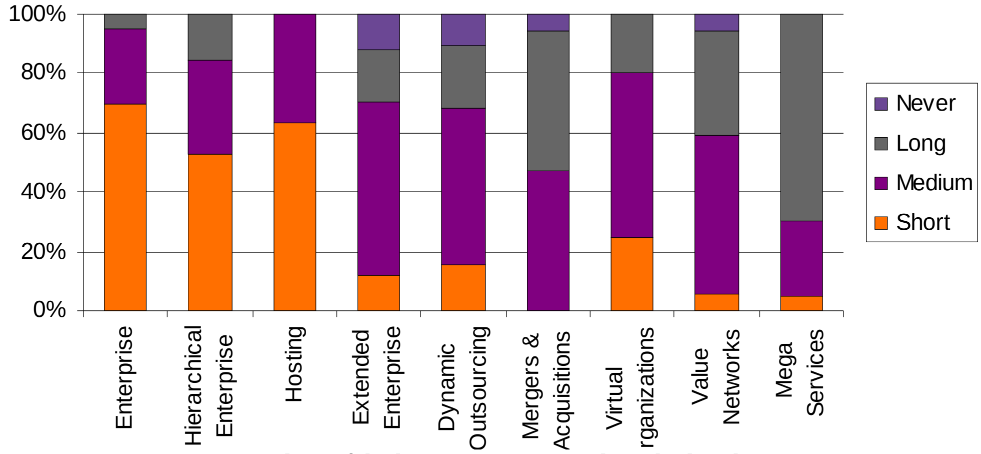


Figure 7: Timeline of the business scenarios described in the SRA

In order to sort by timeline the business scenarios, their timeline can be quantified between 0 (never) and 3 (short) by calculating the weighted average of the results. Figure 8 shows the business scenarios sorted by their average timeline. All scenarios are ranked above 1.4 (medium-long).

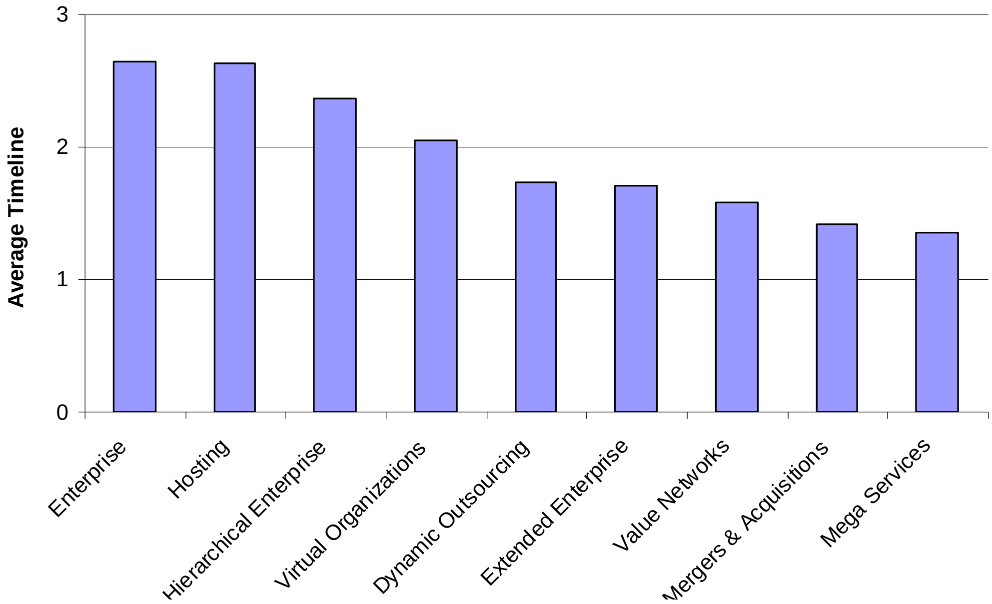


Figure 8: Average timeline of the business scenarios described in the SRA, from 0 (never) to 3 (short)

Finally, Figure 9 represents the average relevance against the average timeline, providing a summary of the feedback of the community on relevance and timeline of the business scenarios.

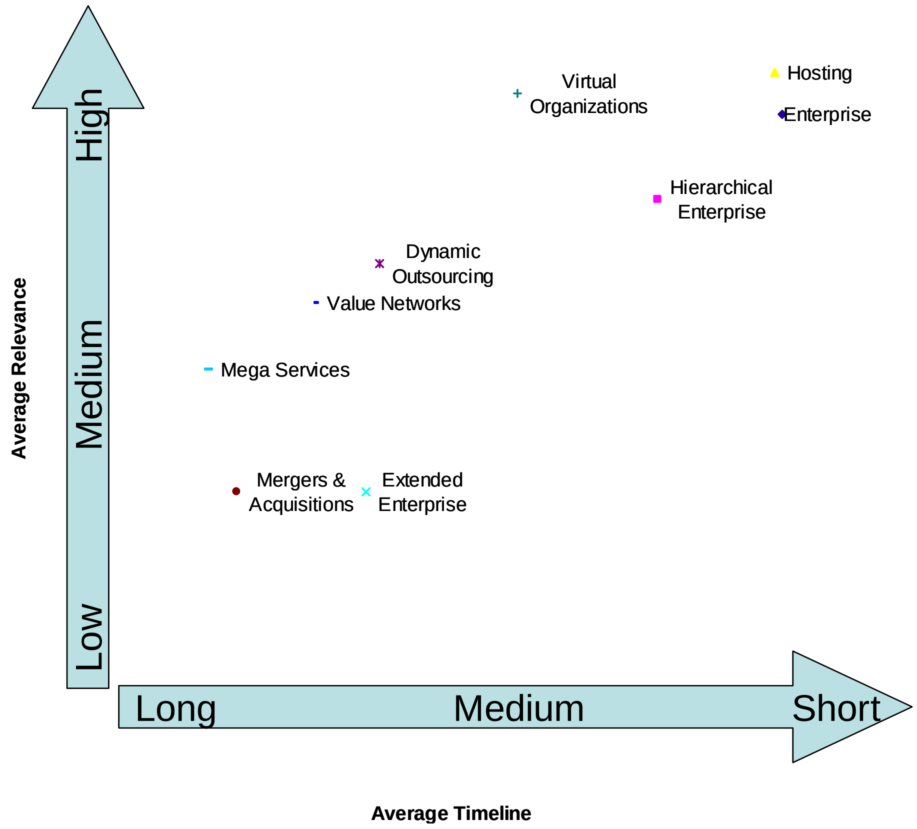


Figure 9: Average relevance against average timeline of the business scenarios described in the SRA.

7 Business and Market Indicators

7.1 Introduction

This section is going to be fully developed in version 3 of the SRA; this version will introduce its future contents. The aim of this chapter is to define a collection of market and business indicators, related to the Grid field, that help defining the SRA challenges and roadmap as well as monitoring the impact of its implementation.

This section is structured in two main parts:

- To find the right indicators the first premise is to define the target market from where we will extract them. Once determined, we will be able to identify the characteristics of this target market: the major players, the relationship among them, their functional needs, their priorities and get their statistics related to the use of grid-related technology.
- The definition of the business indicators, that is, those indicators which aim is to measure the impact of the SRA on the use of Grid technologies by the target market players.

Next subsections will elaborate each of these parts.

7.2 Target market and main players

Generally speaking, target market would be composed of all those stakeholders who may obtain benefit from deploying an infrastructure that allows looking for more computing resources to execute applications/services, either internally or externally to an administrative domain. Some stakeholders will ask for additional resources to execute their applications; other will provide/share them.

Identification of these players, the business/academic sector they belong and thus their activities, allows gathering their needs and derive the functional requirements how business grids can help to achieve them. This is reflected in the business scenarios section of this document.

Target market - sectors

Target market could be classified in three main sets of stakeholders:

- **Academia and research:** emphasis is on collaborative working and utilization of resources for discrete processes that require considerable processing power.
- **Public Sector:** a major user of high power computing. Some of the applications they use with grid technology are predictions and forecasts (whether models, population growth, etc.) in applications combining statistical and quantitative analysis, military applications and e-Government and e-Health
- **Commercial enterprises:** Main drivers in Grid adoption are improving performance -which allows a significant reduction of computing time and thus the

possibility of doing new things, new services and reduce time to market-, cost reduction and also competitive concerns..

Drilling into target commercial enterprises, main functional requirements can be obtained from the following sectors and applications:

- Energy (for oil and gas for exploration)
- Finance/insurance/real estate (securities and brokerage –especially for stock/ portfolio analysis and risk management)
- Aerospace and Automotive (for collaborative design and modelling)
- Life sciences (particularly in pharmaceuticals)
- Media/entertainment (to generate digital animation)
- Architecture (engineering and construction)
- Electronics (design and testing)
- Manufacturing (inter/intra-team collaborative design, process management)
- Utilities (to improve efficiency while dealing with peaks and valleys in utilization)
- Telecommunication Operators (mainly processing and analysis of big amount of data)

An important input for the SRA is the identification of how SOI-related technology can improve the solution for the main needs of the above sectors (i.e. for manufacturing, how SOI can influence in the speed of the product chain and in the product margin; for services, how can improve the capacity utilisation; or for knowledge, how can improve image and reputation, competence and know-how).
Furthermore, looking to the short, mid and long-term needs and expectations of these different sectors will provide the basis for establishing priorities in the SRA and adjust the roadmap accordingly.

Target market - Main players

Looking to the different sectors that are sensible to benefit from grid technology, we can abstract the different generic players who have an active role in the market and are present in one or more of the above sectors:

- **Middleware providers**, in particular lower layer towards the infrastructure and grid application and utility
- **Application providers**, in particular modularized applications that can be used on a grid
- **Hardware providers**
- **Grid Systems Integrators**
- **Network providers,** e.g. offering servers for hiring, or in case when the mobile world is involved, the Telco company is providing its infrastructure
- **Content providers -** Content creator, content aggregator, content distributor
- **Service providers -** Grid service provider, retail service provider, access provider, …
- **Payment providers**

- **Users –**Business and private users, government and science who can use the grid , provide resources for the grid or even both.
- **Regulatory bodies,** like standard bodies, and governments who provide the rules
- **Other Influential entities** as media, consultants, other information sources, (press, Web portals, professional events)
- **Identity providers** as Certification Authorities.

These players and the sectors they belong are the basis to define the business and market indicators that are introduced below.

7.3 Indicators

The purpose of these indicators is to have a snapshot of the use of grid technology in certain moment in time and to follow its evolution by getting their value periodically. This evolution can show us the impact on the market of the implementation of the SRA. In what follows we introduce the list of proposed indicators that will be refined in further versions of the SRA.

Adoption Dynamics

Different industrial sectors adopt technologies at different rates. There has already been large-scale adoption of grid technologies by many financial and engineering companies. Penetration into public sector, retail and general consumer sectors has been much less so, if at all. The adoption dynamics of each sector will differ because their business requirements differ, but this indicator will reflect a number of these dynamics such as the number of sectors adopting business grids, the speed of adoption with individual sectors, the increasing speed of adoption in sectors adopting later, etc.

Emergence of New Vendors

The belief is that the development of business grids will give rise to new companies selling hardware, software and services to both implement the grids themselves as well as implement new services using these business grids. This indicator will reflect the number of new commercial organisations that start trading in areas pertaining to business grids.

Interest from Current Vendors

Just as new companies will form to take advantage of the business opportunities around business grids, so existing vendors of hardware, software and services will adopt technologies and business models to take advantage of these opportunities too. This metrics will reflect the number of new offerings of hardware, software and services offered by existing companies (those existing prior to the formation of NESSI) to address the opportunities arising from business grids.

Emergence of new Grid Middleware products

New requirements from the business world can result in new grid middleware products in the market which allow that requirements to be fulfilled. This new business can be dealt by current or new vendors that found that gap in the market. In addition, new grid middleware can come from research projects that are more focused on business aspects.

One of the sources of new requirements for these products is the list of challenges depicted in NESSI-GRID SRA. This indicator will reflect the number of new middleware products that have appeared in the market, regardless of whether they are commercial or open source.

Behaviour of Big Companies

Big company behaviour basically falls into one of two main categories. There are those big companies that will be providing business grids – whether this is providing the components to build them, providing the grids themselves as utilities, etc. - and those who are using them internally. Indications related to the former will be covered by the previous indicator (Interest from current vendors). This indicator will reflect the latter. A possible measurement for the success of this might be the percentage of their internal infrastructure that is in a business grid.

Standards and Ontology Resources

Interoperability of business grid components from different providers will only be possible through the widespread adoption of standards. One of the reasons that commercial grid implementations are much wider than those grids implemented with a 'standards-based' approach is predominantly down to a lack of a set of standards and consequently a lack of software that implements these standards. This indicator will reflect the number of agreed and ratified standards for business grids that have been implemented by products from multiple vendors.

Emergence of New Consulting Companies

Just as the development of business grids will give rise to new companies selling hardware, software and services to both implement the grids themselves as well as implement new services using these business grids, there will be business opportunities for organisations consulting in this area. This indicator will reflect the number of new commercial organisations that start trading to provide consultancy pertaining to business grids.

Emergence of New Applications, Services, Architectures and Components

The belief is that the development of business grids will give rise to new applications, new commercial services, architectures and components (be they hardware, software or services). These will be used to undertake new business challenges or to address business challenges in new and innovative ways. This indicator will reflect the number of these business solutions that simply could not have been addressed without the implementation of business grids.

Publications

There are many academic publications today with grid-related content. There are very few, if any, non-academic publications directly aimed at the business user of these technologies. As business grids become more established as platforms, the number of publications will undoubtedly increase. This indicator will reflect the number of business-centric, non-academic publications around business grids, their implementation, use and management.

Seminars and Conferences

There are many academic seminars and conferences today with grid-related content. There are very few, if any, non-academic conferences / seminars directly aimed at the business user of these technologies. As business grids become more established as platforms, the number of these will undoubtedly increase. This indicator will reflect the number of business-centric, non-academic conferences around business grids.

Analyst and Media Attention

There is already a considerable amount of analyst and media attention on grid computing. More correctly, this interest has existed in the past and is now starting to tail off as grid has been somewhat overtaken by newer technologies which are currently perceived as being 'hot'. This indicator will reflect the number of media and analyst articles that are written about all aspects of business grids.

Amount of grid-enabled applications

Business grids are only of interest if they can run applications. Some of these applications will be new. Others will be implementations of existing (legacy) applications that can both run on, and take advantage of a business grid. This indicator will reflect the number of business applications, both commercial and open source that can run on a business grid.

Evolution of users (nodes) of grid middleware, grouped by business or academic

As business grids are adopted by commercial organisations, the number of nodes running grid middleware or the number of users of grid middleware (whether directly or indirectly) will increase. This metrics is a simple reflection of the number of nodes with “business grid middleware” installed on them or the number of (direct and indirect) users of “business grid middleware”.

Evolution of other technologies related to “SOI”: SOA, WS, …

Business grids build upon many technologies, standards etc. from other areas of IT. Without these items being in place to be ‘reused’ by business grids, there will be a requirement for business grids to ‘reinvent the wheel’ and provide similar functions themselves. This is both wasteful and time-consuming. As such, these other technologies and standards need to evolve to support the needs of business grids. This indicator will reflect the progress made towards delivering these underlying and complementary technologies and standards, as well as reflecting the modification or extension of existing technologies and standards to better support the needs of business grids.

Usage of any common platform, joint testbeds, repositories etc. that are created for the purpose of fostering the implementation of business grids

Whilst most commercial organisations are likely to either purchase the components and build their own business grids or buy commercial services from other commercial organisations, there may be a proportion of organisations that will consider an open source or ‘community’ approach to building a business grid. To facilitate these, repositories of open source business grid software, testbeds, common platforms, etc. are going to be needed. This indicator will reflect the number of these entities and the numbers of commercial companies actively using them.

Emergence of new Business Models for Grids [125]

This indicator would consider business models related to how to apply grid-based applications on a commercial scope. Then, we will mainly consider in this category business models that describe network-centric applications, because they are candidates to get benefit from grid technology. As example, variants of ASP model (Application Service Provider) would enter into this category. This indicator will reflect the number of these new business models.

New application areas for Grid Technology [126]

As grid technology progresses in its development and new functionalities are being achieved, there will be new application areas that will benefit from this technology. Aspects of security, latency, complexity and autonomous maintenance are some of the barriers that prevent its use on particular areas. Some promising areas are i.e. sensor networks or remote control of complex and distributed instrumentation. This indicator will reflect the number of new application areas where grid technology has been adopted.

8 Conclusions and Outlook

This document presents the second version of a vision and a strategic research agenda for grid-like infrastructures to be used as general purpose infrastructure in business environments. It is largely driven by the ambition to define a research agenda that matches industrial needs and finds the right balance between visionary outlook and the practicality of the presented issues.

The SRA has been made subject to a broad community review process which resulted in several extensions of the identified requirements and challenges. However, this process also validated the specified SRA and its chosen approach as highly relevant and largely complete.

By following a multi-perspective approach, an extensive set of requirements has been identified based on societal trends (e.g. compliance, sustainability), business scenarios (e.g. enterprise, value networks), technical trends (e.g. multi-core), as well as technical constraints (e.g. nature of business applications). The requirements themselves are structured via functional, commercial and non-functional categories. Based on a thorough analysis of the state of the art requirements are translated into a detailed set of research challenges.

Looking on the extensive set of identified research challenges it is pretty obvious that these cannot be tackled separately from each other. Instead, we see a need for advancing the state of the art in 3 major dimensions, namely architectures, lifecycle management and infrastructure technologies.
The following 3 key challenges represent these dimensions:

KC 1: **New system architectures** that harmonize service architectures (SOA) and infrastructure architectures (SOI), advance the structure of multi-tier, federated and Internet scale architectures, support all kinds of business models, applications and emerging hardware environments and provide transparent and integrated access for all relevant stakeholders (architects, engineers, operators, customers, …).

KC 2: **Advanced system lifecycle approaches** including engineering, deployment, composition, provisioning, operation and decommissioning phases that support transparent knowledge tracking, feedback loops, prediction and simulation, allow for a clear separation of concerns between different stakeholders (business vs. IT, developers, providers, customers, …) and support the full variety of business scenarios (from traditional data centres to complex service value networks) while adhering to overarching sustainability requirements.

KC 3: **Advanced infrastructure technologies** in terms of hardware (energy efficient, flexible allocation, virtualization …), middleware (new multi-tier system design, flexible storage systems, harmonized virtualization on all layers) and related programming models (parallel programming, multi-core) that meet the required flexibility of the networked economy.

The next version of this SRA will feature a detailed roadmap for the identified challenges, including recommendations for researchers, industry and policy makers. Furthermore, continued community involvement will assure the further complementation on all aspects of this SRA, such as business scenarios, state-of-the-art and research challenges.

As a first indication and input to the planned roadmap, the following figure represents the discussed business scenarios and their dependencies. The Enterprise scenario is set as basis out of which other scenarios can be stepwise developed according to three dimensions: organizational complexity, technical reach and degree of service enablement.

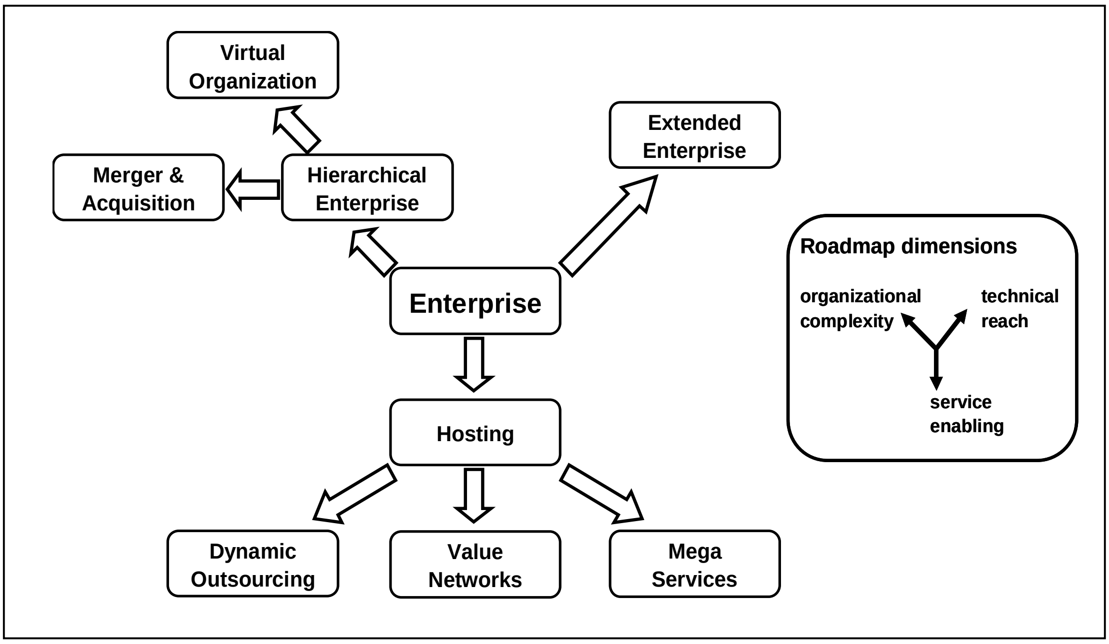


Figure 10: Roadmap for business scenarios.

Though Business Grids do not directly create new application types or business models, their impact on the further evolution of the service economy cannot be underestimated. As a general-purpose service-oriented infrastructure they will become the backbone for highly flexible, service-based interactions, where business needs and IT capabilities are transparently reconciled.

9 References

[1] NESSI: *SRA – Volume 1 – Framing the future of the service oriented economy*. Jan.2006. http://www.nessi-europe.com/documents/NESSI_SRA_VOL_1_20060213.pdf
[2] NGG3 Expert Group Report: *Strategic Future for European Grids: Next Generation GRIDs based on SOKU - A new paradigm for service delivery and software infrastructure*. December 2005
[3] Jim Gray: *Distributed Computing Economics*. Technical Report, MSR-TR-2003-24, Microsoft Research, March 2003. http://research.microsoft.com/research/pubs/view.aspx?tr_id=655
[4] I. Foster, C. Kesselman, S. Tuecke: *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*. International J. Supercomputer Applications, 15(3), 2001.
[5] The Data Grid: *Towards an Architecture for the Distributed Management and Analysis of Large Scientific Datasets*. http://www.globus.org/alliance/publications/papers/JNCApaper.pdf
[6] Open Grid Forum. http://www.ogf.org/
[7] I. Foster, C. Kesselman, J. M. Nick, S. Tuecke: *The Physiology of the Grid: An Open Grid Services Architecture for Distributed Systems Integration* v2002-06-22 http://www.globus.org/alliance/publications/papers/ogsa.pdf
[8] I. Foster *et al.* : *GFD-I.030 The Open Grid Services Architecture, Version 1.0* 2005-01-29 Global Grid Forum http://www.ogf.org/documents/GFD.30.pdf
[9] S. Tuecke et al.: *Grid service specification*. June 02. http://www.gridforum.org/ogsi-wg
[10] Globus. http://www.globus.org
[11] EDG http://eu-datagrid.web.cern.ch/eu-datagrid/
[12] LCG http://egee-sa1.web.cern.ch/egee-sa1/joining.html
[13] Condor http://www.cs.wisc.edu/condor/
[14] gLite. https://edms.cern.ch/file/476451/1.0/architecture.pdf
[15] EGEE-1 http://public.eu-egee.org
[16] EGEE-2 http://www.eu-egee.org/
[17] GRIA - Grid Resources for Industrial Applications. http://www.gria.org/
[18] NorduGrid http://www.nordugrid.org/
[19] NextGRID: *NextGRID Architecture White Paper*. December 2005. http://www.nextgrid.org/white_paper.htm
[20] G. Allen et al.: *Cactus-G: Enabling High-Performance Simulation in Heterogeneous Distributed Computing Environments*. http://edoc.mpg.de/3477
[21] NetSolve. http://icl.cs.utk.edu/netsolve/
[22] Unicore Forum. http://www.unicore.org/
[23] GrADS - Grid Application Development Software Project. http://hipersoft.rice.edu/grads/
[24] Datagrid: http://eu-datagrid.web.cern.ch/eu-datagrid
[25] IBM: IBM Grid Computing. http://www-1.ibm.com/grid/
[26] SUN: Sun N1 Grid Engine. http://www.sun.com/software/gridware/
[27] Platform Computing: *Enterprise Grid Coordinator*. http://www.platform.com/Products/Platform.Enterprise.Grid.Orchestrator/
[28] DataSynapse products. http://www.datasynapse.com/products/default.asp
[29] Univa Globus Enterprise. http://www.univa.com/products/
[30] Cluster Resources: MOAB. http://www.clusterresources.com/pages/products/moab-cluster-suite.php
[31] XEN *Virtualization Software*. http://www.xensource.com
[32] VMWare: *Virtual Infrastructure 3*. http://www.vmware.com/products/vi/
[33] Virtual Iron: *Virtual Iron*. http://www.virtualiron.com/

[34] SUN: *Solaris Container*. http://www.sun.com/software/solaris/utilization.jsp
[35] Microsoft: *System Center Manager*. http://www.microsoft.com/systemcenter/scvmm/default.mspx
[36] SWSoft: *Virtuozzo*. http://www.swsoft.com/products/virtuozzo/
[37] IBM: *Virtualization Suite*. http://www.ibm.com/virtualization
[38] ESFORS: ESFORS Position Statement. Version 2, June 2006. http://www.esfors.org
[39] WS-Agreement http://www.ogf.org/Meetings/GGF11/Documents/draft-ggf-graap-agreement.pdf
[40] WS-Context http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ws-caf
[41] DEISA http://www.deisa.org/
[42] Grid5000 https://www.grid5000.fr/
[43] Open Science Grid http://www.opensciencegrid.org/
[44] TeraGrid http://www.teragrid.org/
[45] NAREGI http://www.naregi.org/index_e.html
[46] ChinaGrid http://www.chinagrid.edu.cn/chinagrid/index.jsp
[47] RFC3820, X.509 Proxy Certificates http://www.ietf.org/rfc/rfc3820.txt
[48] MyProxy http://grid.ncsa.uiuc.edu/myproxy/
[49] Attribute Certificates http://www.ietf.org/rfc/rfc3281.txt
[50] VOMS http://hep-project-grid-scg.web.cern.ch/hep-project-grid-scg/voms.html
[51] PERMIS http://sec.cs.kent.ac.uk/permis/
[52] XACML http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=xacml
[53] SAML http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security
[54] GACL http://www.gridpp.ac.uk/authz/gacl/
[55] GridSite http://www.gridsite.org/
[56] OGSA-DAI http://www.ogsadai.org.uk/
[57] SRB http://www.sdsc.edu/srb/index.php/Main_Page
[58] iRODS http://irods.sdsc.edu/index.php/Main_Page
[59] GPFS http://www-03.ibm.com/systems/clusters/software/gpfs.html
[60] AFS http://www.openafs.org/
[61] Web Service Distributed Management http://www.wsdm.com/
[62] CDDLM http://forge.gridforum.org/projects/cddlm-wg
[63] JSD http://jsdl.sourceforge.net/, http://www.gridforum.org/documents/GFD.56.pdf
[64] Basic Execution Service http://forge.ogf.org/sf/projects/ogsa-bes-wg
[65] SAGA http://forge.ogf.org/sf/projects/saga-rg
[66] OMII http://omii-europe.org/OMII-Europe/
[67] VDT http://vdt.cs.wisc.edu/
[68] GridComp and GCM http://gridcomp.ercim.org/
[69] ProActive http://proactive.inria.fr/
[70] OGF Reference Model Working Group: http://www.ogf.org/gf/group_info/view.php?group=rm-wg
[71] SCF Associates ltd. (25 November 2006) Commercial Exploitation of Grid Technologies and Services. In: Contract no.30-CE-065970/00-56 Simon Forge and Collin Blackman
[72] The 451 Group (April 2006) Grid Computing Adoption in the Financial Sector. An investigation into User experiences and Vendor Strategies.
[73] AKOGRIMO http://www.akogrimo.org
[74] XTREEMOS http://www.xtreemos.org
[75] T. Guan, E. Zaluska and D. De Roure. “A Grid Service Infrastructure for Mobile Devices” Semantics, Knowledge and Grid, SKG’05
[76] D. Millard, A. Woukeu, F. B. Tao and H. Davis “Experiences with Writing Grid Clients for Mobile devices” ELeGI Conference on Advanced Technology fro Enhanced Learning 2005

[77] GRIDLAB http://www.gridlab.org/about.html
[78] L. Baresi and S. Guinea,"Towards Dynamic Monitoring of WS-BPEL Processes," Proc. Int'l Conf. Service-Oriented Computing (ICSOC 05), 2005
[79] P. Grabowski and B. Lewandowski "Access from J2ME-Enabled Mobile Devices to Grid Services" White Paper from GridLab Project – 2004 http://www.gridlab.org/WorkPackages/wp-12/res/docs/Mobility_Paper_GridLab.pdf)
[80] INTELIGRID http://www.inteligrid.com
[81] P. Grabowski, K. Kurowski, J. Nabrzyski and M. Russell "Context Sensitive Mobile Access to Grid Environments and VO Workspaces" Mobile Data Management, MDM2006
[82] O.M. Rendon, F.O.M. Pabon, M.J.G. Vargas and J.A.H. Guaca "Architectures for Web services access from mobile devices" Web Congress, 2005 LA-WEB 2005
[83] S. Isaidis and V. Getov Chapter "Integrating Mobile Devices into the Grid: Design Considerations and Evaluation" from the book "Euro-Par 2005 Parallel Processing" Springer Berlin / Heidelberg 2005
[84] L. dos S. Lima, A. T. A. Gomes, A. Ziviani, M. Endler, L. F. G. Soares and B. Schulze "Peer-to-Peer Resource Discovery in Mobile Grids" ACM International Conference Proceeding Series; Proceedings of the 3rd international workshop on Middleware for grid computing, 2005
[85] D.C. Chu and M. Humphrey "Mobile OGSI.NET: Grid Computing on Mobile Devices" in: Proceedings of the 5th IEEE/ACM IEEE/ACM International Workshop on Grid Computing (associated with Supercomputing 2005), November 2004, Pittsburgh, (http://www.cs.virginia.edu/~humphrey/papers/MobileOGSI.pdf)
[86] A. Litke, D. Skoutas and T. Varvarigou "Mobile Grid computing: Changes and challenges of resource management in a mobile grid environment", in: Access to Knowledge through the Grid in a Mobile World Workshop, held in conjunction with 5th Int. Conf. on Practical Aspects of Knowledge Management, PAKM 2004, Vienna (http://www.mobilegrids.org/docs/pakm2004/papers/3-Mobile-Grid-Computing.pdf)
[87] KVM http://kvm.qumranet.com/kvmwiki
[88] J.O. Kephart, "Research Challenges of Autonomic Computing", Proc. Of 27th Int. Conf. Software Eng (ICSE 05), ACM Press, 2005
[89] L. Baresi, E. di Nitto, C. Ghezzi, Towards Open-World Software: Issues and Challenges, Computer, oct 2006
[90] G. Canfora and M. di Penta, "Testing Services and Service-Centric Systems: Challenges and Opportunities", IT Professional, Mar/Apr 2006
[91] A. Bertolino et al, "Audition of Web Services for Testing Conformance to Open Specified Protocols," J. Stafford et al, eds, Architecting Systems with Trustworthy Components, Springer, 2006
[92] Amazon Simple Storage Service (S3) http://aws.amazon.com/s3
[93] Amazon Elastic Compute Cloud (EC2) http://aws.amazon.com/ec2
[94] Oracle 10g http://www.oracle.com/database/index.html
[95] dCache http://www.dcache.org/
[96] SRM http://sdm.lbl.gov/srm-wg/
[97] Sybase Avaki EII http://www.sybase.com/products/allproductsa-z/avakieii
[98] RFC2510: X.509 Public Key Infrastructure http://tools.ietf.org/html/rfc2510
[99] ITU X.509 Standard http://www.itu.int/rec/T-REC-X.509-200508-I
[100] WCAMG'07 *Context Awareness in Mobile Grids* http://wcamg2007.inf.puc-rio.br/
[101] BEinGRID Project http://www.beingrid.eu/
[102] OGF Grid Resource Allocation Agreement Protocol Working Group (GRAAP-WG) http://forge.ogf.org/sf/projects/graap-wg
[103] WS-PolicyConstraints, Sun Research, http://research.sun.com/projects/xacml.
[104] AssessGrid http://www.assessgrid.eu/

[105] WS-Policy http://www.ibm.com/developerworks/library/specification/ws-polfram/
[106] HPC4U http://www.hpc4u.org/
[107] National Earthquake Engineering Simulation Grid http://it.nees.org/
[108] International Virtual Observatory Alliance http://www.ivoa.net/
[109] European Telecommunications Standards Institute http://www.etsi.org/
[110] ETSI Grid Technical Committee http://www.etsi.org/WebSite/Technologies/GRID.aspx
[111] KVM http://kvm.qumranet.com/kvmwiki
[112] http://www.osgi.org
[113] Java Enterprise Edition http://java.sun.com/javaee
[114] Hedeby http://hedeby.sunsource.net/
[115] GridHypervisor http://www.gridhypervisor.org/
[116] OGF Resource Usage Service Working Group http://forge.ogf.org/sf/projects/rus-wg
[117] Distributed Grid Accounting System http://www.to.infn.it/grid/accounting/main.html
[118] GRIA Accounting Services Survey for NextGrid http://www.gria.org/downloads/publications/NextGRID%20P5.4.3%20v1.0.pdf
[119] GRIA http://www.gria.org
[120] SORMA Project http://sorma-project.org/
[121] GridBus Project http://www.gridbus.org/
[122] HP Tycoon Project http://tycoon.hpl.hp.com/
[123] Grid Trust Study http://www.gridtrust.eu/
[124] Workflow in Grid systems, G. Fox and D. Gannon, Concurrency and Computation: Practice and Experience vol. 18, issue 10, aug. 2006
[125] GRIDIPEDIA – Emerging Business Models, http://www.gridipedia.eu/index.php?id=699
[126] GRIDIPEDIA – New application areas for grid technology: http://www.gridipedia.eu/index.php?id=362
[127] NESSI: *NESSI Strategic Research Agenda - Vol. 3.FP7-1 (NESSI Roadmap: NESSI Input for FP7)*; Public -Version 2006-7-26 – Revision 1.0; http://www.nessi-europe.com/Nessi/Portals/0/Nessi%20repository/Publications/Flyers/NESSI_SRA_VOL_3_2006_07.pdf

ANNEX A: Overview of the Community Involvement Process

The aim of this annex is to describe the implemented process for involvement of the community in the second version of the NESS-Grid SRA. The first section provides an overview of the community involvement process. The second section lists the relevant members of the grid community that have contributed. The third section evaluates the answers to the closed questions on relevance and timeline of the business scenarios. All the inputs received as result of the call for contribution are publicly available at the SOI-NWG WIKI (http://www.soi-nwg.org).

Definition of the Relevant Members of the Community

In the context of NESSI-Grid, Grid is defined as the adaptive service-oriented utility infrastructure for arbitrary business applications, primarily aimed at providing business flexibility in the most efficient way. Typically, business applications do not exist as independent executables but involve a complex technology stack, containing for example application servers, databases, other middleware components and other co-existing and cooperating business applications. In addition, usage patterns are often interactive or session-based. The scope of research and development in business grids covers the development of middleware that provides an ITC infrastructure to support business applications; from the basic services of the infrastructure to the support for process provisioning, monitoring, integration and outsourcing. That includes research on service computing, utility computing, virtualisation, SLAs, stable and safe architectures data management, knowledge and analysis services, service flow description, dynamic VOs...

Contribution to the SRA is open to any individual (*individual contribution*) with relevant experience in the business grid area, i.e., it is not restricted to NESSI members. In fact its promotion to any valuable player is one of the aims of the community involvement activity. Participation of experts from Industry and Academia is of special importance and so invitation for participation was submitted to business oriented working groups and experts with a clear view of the state-of-the-art and the challenges in the major technological components of a Grid solution. Additionally, there are groups of interest (no legal entities) in the grid community; such as grid projects, infrastructures and middleware solutions, whose feedback has provided great value to the document. It is understood that in this case (*groups of interest contributions*) the contributors acted as representatives of their corresponding groups of interest.

Contribution Models

Contributions could have any extension. Accepted contributions imply that the corresponding contributor gets cited in an acknowledge section. Three approaches were defined to provide feedback and contribute to the NESSI SRA:

- **Individual Contributions**: Experts from Industry, Research and Academia contributed by submitting comments and/or change proposals completing the form provided in Annex B.

- **Group of Interest Contributions**: Representatives of the main groups of interest (research projects, infrastructures and middleware solutions) contributed by submitting one of the surveys provided in Annex C, D and E, which have been designed to extract relevant information according to the type of project.
- **Related Analysis**. Projects and working groups also contributed by submitting related analysis, surveys and interviews.

Implementation of the Open Participation Process

NESSI-Grid followed an open process of consultation to collect inputs to the SRA from relevant members of the grid community.

PHASE A: Design and Development of the Open Process

- **June 18**: Development of a draft describing the open process for community involvement and distribution to the participants of the Service Oriented Infrastructure NESSI Working Group meeting held on June 21.
- **June 21**: Presentation of the draft at the NESSI Working Group meeting and open discussion. Representatives from main projects on grid computing participated in the meeting and provided feedback to the community involvement process.
- **July 5**: Incorporation of the definitive open process for community involvement into the SOI-NWG web page and into the draft of the second SRA version (v1.5).
- **July 15**: Release of the draft of second SRA version (v1.5).

PHASE B: Implementation of the Open Process

- **July 15**: Call for Contribution, which was open to any individual, company or group of interest with relevant experience in the business grid area, i.e., it was not restricted to NESSI members. The NESSI-Grid project putted the resources involved in promoting and monitoring participation of individuals and groups of interest in this call for contribution. The following relevant players of the Grid community were personally contacted.
    - **Research Projects**: Coordinators of large Grid-related FP6 projects: AKOGRIMO, ARGUGRID, ASSESSGRID, A-WARE, BEINGRID, BREIN, BRIDGE, CHALLENGES, CHEMOMENTUM, COREGRID, DATAMININGGRID, DEGREE, EC-GIN, ECHOGRID, EDUTAIN@GRID, G-ECLIPSE, GREDIA, GRID@ASIA, GRID4ALL, GRIDCOMP, GRIDCOORD, GRIDECON, GRIDTRUST, HPC4U, INTELIGRID, KNOWARC, KWF GRID, NEXTGRID, ONTOGRID, PROACTIVE, PROVENANCE, QOSCOSGRID, SIMDAT, SORMA, UNIGRIDS and XTREEMOS
    - **Infrastructures**: Coordinators of European and regional Grid infrastructures: EGEE, DEISA, D-Grid, UK NGS, NORDUGRID and BALTICGRID
    - **Middlewares**: Coordinators of main Grid OSS communities and middleware projects: GRIA, GLOBUS ALLIANCE, GLITE, UNICORE and OMII. Other middleware solutions for utility computing, virtualisation, SLAs, management, knowledge and analysis services, service flow description, dynamic VOs, etc. will be invited to contribute.

  - **Individual Contributors**: Participation of experts from Industry and Academia was of special importance and so invitation for participation was also submitted to (1) business oriented working groups such as the Enterprise Grid Requirements Working Group at OGF, and the Service Oriented Infrastructure NESSI Working Group; and (2) the coordinators of the FP6 Grid Projects Technical Groups and Concertation Tasks.
- **September 6**: First reminder of the call for contribution.
- **September 20**: Second reminder of the call for contribution.
- **September 26**: Presentation of the Call for Contributions to the NESSI-Grid SRA at European Service, Software and Grid Technology Days 2007, Brussels.
- **October 1**: First release of the results of the call for contribution at the SOI-NWG WIKI and extension of the deadline to October 10th.
- **October 2**: Presentation of the Call for Contributions to the NESSI-Grid SRA in the Grid in Industrial Contexts session at EGEE 07, Budapest.
- **October 10**: Deadline and publication of the inputs at the SOI-NWG WIKI. Some of the contributors prefer to keep their contributions private.

PHASE C: Analysis of Contributions

- **October 10 - October 12**: Evaluation of the response of the community.
- **October 12 - November 13**: Incorporation of the contributions in the SRA v2.0.
- **November 13 - November 23**: Publication of the last draft of the SRA for public comment.
- **November 30**: Release of SRA v2.0.

SOI-NWG as Instrument for Open Participation

The Service Oriented Infrastructure NESSI Working Group (SOI-NWG) has provided the umbrella under which the open edition and promotion of a widely agreed NESSI-SRA for business grids have taken place:

- The SOI-NWG mailing list (wg-service-oriented-infrastruture@nessi-europe.eu) has been used for open discussion on the contents of the SRA and to notify updates in the web.
- The SOI-NWG WIKI (http://www.soi-nwg.org) has described the open process adopted to develop the SRA and incorporates collaborative tools which support distribution of the subsequent versions of the SRA and workshop presentations as well as the contribution and participation of members of the community in the SRA development process.

Contributions Received from the Community

Research Projects

The following projects have submitted the survey form for research projects, acting the contributor as a representative of the research project and not as an individual contributor himself:

- ASSESSGRID, by Odej Kao (TU Berlin)
- A-WARE, by Sergio Bernardi (CINECA)
- BRIDGE (SIMDAT & GRIA)
- EDUTAIN@GRID, by Thomas Fahringer (University of Innsbruck)
- G-ECLIPSE, by Harald Kommayer (NECLAB)
- GRID4ALL Project, by Vladimir Vlassov (KTH)
- GRIDCOMP, by Marco Danelutto (UNIPI)
- GRIDTRUST, by Phillippe Massonet (CETIC)
- HPC4U, by Géry Schneider (IBM)
- KWF GRID, by Steffen Unger (FIRST)
- ONTOGRID, by Oscar Corcho (UPM)
- SIMDAT, by Mike Boniface (IT Innovation)
- SORMA, by Erel Rosenberg (Correlation Systems Ltd)
- XTREEMOS, by Bernd Scheuermann (SAP)
- ProGRID, by Peter Stoll (SCIENCE COMPUTING)
- UNIGRIDS (UNICORE)
- BEinGRID, by Elies Prunes (ATOS Origin)

Infrastructures

The following infrastructures have submitted the survey form for infrastructures, acting the contributor as a representative of the research project and not as an individual contributor himself:

- CESGA, by Carlos Fernández (CESGA)
- EGEE, by Gabriel Zaquine (CERN)

Middleware Solutions

The following middleware systems have submitted the survey form for middleware systems, acting the contributor as a representative of the research project and not as an individual contributor himself:

- GRIA, by Mike Boniface (IT Innovation)
- gLite (EGEE)
- UNICORE, by Achim Streit (Forschungszentrum Jülich GmbH)
- GRID4ALL Middleware, by Vladimir Vlassov (KTH)
- Globus GridWay, by Eduardo Huedo (UCM)
- KNOWARC, by Peter Stoll (SCIENCE COMPUTING)

Individual Contributions

The following relevant experts from Industry, Research and Academia have contributed by completing the survey form for individual contributors: Craig A. Lee (The Aerospace Corporation/OGF President), Yvon Jégou (INRIA), Oscar David Sánchez (INRIA), Bernd Scheuermann (SAP), Daniel Galindo and Luis Pablo Prieto (TI+D), Daniel Vladusic (XLAB), Gregor Pipan (XLAB), Marjan Šterk (XLAB), Matej Artač (XLAB), Silvana Muscella (Metaware), Steve Crumb (OGF), Mark Linesch (OGF), David Snelling (Fujitsu Labs UK), Ali Anjomshoaa (Mobability Limited), Ian Osbourne (Intellect UK), Rubén S. Montero (UCM), Antonio Puliafito (University of Messina), Syed Naqvi (CETIC), Philippe Massonet (CETIC), Pierluigi Ritrovato (CRMPA), Björn Kolbeck, and Jan Stender (Zuse Institute Berlin).

Related Analysis

The following projects and working groups have generated related analysis, surveys and interviews:

- CoreGRID European Project: Roadmap version 2 on Knowledge and Data Management, Roadmap version 2 on Programming Model, Roadmap version 2 on Architectural Issues, Roadmap version 2 on Grid Information, Resource and Workflow Monitoring Services, Roadmap version 2 on Resource Management and Scheduling, and Roadmap version 2 on Grid Systems, Tools and Environments
- Challengers European Project: "Research Agenda" and "Roadmap on Challenges and Visions" documents (to be available in December 2007)
- 3S ECSS European Project: Joint European White Paper for R&D and Industry in the Service and Software Architectures, Infrastructures and Engineering field (under development)
- DEGREE European Project: ES family of applications and their Grid requirements (focused on e-Science)
- Concertation Task for Collaboration on Research Inventory and Roadmaps: European Grid Roadmap V1.1 and First input for European Grid Roadmap V2.0
- Technical Working Group on Grid Architecture: Scientific and Technological Challenges Section in Third White Paper
- Open Grid Forum Working Group on Enterprise Grid Requirements

High-level Plan for Community Involvement in SRA v3.0

The goal of version 3 of this SRA is to provide a detailed roadmap analysis for the identified research challenges, including recommendations for researchers, industry and policy makers. This section describes the open process of consultation to collect inputs to the third version of the SRA from relevant members of the grid community:

PHASE A: Design and Development of the Open Process

- **March 1-April 30, 2008**: Design and development of the questionnaires for community involvement in the edition of SRA v3.0.
- **April 30, 2008:** Release of SRA v2.5.

- **May 15, 2008:** Presentation of SRA v2.5 and community involvement process in 4th SOI-NWG meeting.

PHASE B: Implementation of the Open Process

- **May 15-June 30, 2008**: Call for participation in the edition of SRA v2.5.
- **July 15, 2008**: Publication of contributions at SOI NWG site.

PHASE C: Analysis of Contributions

- **July 15-September 15, 2008**: Evaluation of the response of the community and incorporation of the contributions into the SRA v2.9.
- **September 16-September 30, 2008**: Publication of the last draft of the SRA for public comment.
- **October 30, 2008:** Release of SRA v2.5.

ANNEX B: Sample of Contribution Form for Individual Contributors

The Service Oriented Infrastructure NESSI Working Group (www.SOI-NWG.org) is defining the NESSI Strategic Research Agenda on new generation of ITC infrastructures (NESSI-Grid SRA). In this context, business Grid is envisioned as the adaptive service-oriented utility infrastructure for business applications, which will become the general ICT backbone in future service oriented economies.

The aim of this template is to collect the feedback of relevant experts in the Grid community on the draft of second version of the NESSI-Grid SRA (v1.5). Thank you for participating in this survey. Your feedback is important in helping us improve future versions of the NESSI-Grid SRA. Please be concise and specific.

CONTRIBUTOR

Name: Organization: E-mail:

1. Overall Comments on the SRA

2. Specific Comments and Amendments on Sections

Copy the next table for each section. Comments have to be accompanied by an amendment proposal. An amendment proposal will be produced by copying the amended text and introducing changes under change control.

Section Number:
Comment
Amendment

Section Number:
Comment
Amendment

ANNEX C: Sample of Survey Form for Research Projects

The Service Oriented Infrastructure NESSI Working Group (www.SOI-NWG.org) is defining the NESSI Strategic Research Agenda on new generation of ITC infrastructures (NESSI-Grid SRA). In this context, business Grid is envisioned as the adaptive service-oriented utility infrastructure for business applications, which will become the general ICT backbone in future service oriented economies.

The aim of this template is to collect the feedback of relevant grid application projects on the draft of second version of the NESSI-Grid SRA (v1.5). Please notice that you, the contributor, act as representative of a research project and not as an individual contributor himself. Thank you for participating in this survey. Your feedback is important in helping us improve future versions of the NESSI-Grid SRA. Please, be concise and specific.

CONTRIBUTOR Name: Organization: E-mail: Project:

PART A: BUSINESS SCENARIOS

1. **What is your opinion on the relevance and timeline of the business scenarios described in the SRA?**

Scenario	Relevance	Timeline
Enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hosting	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Extended enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Virtual organisations	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Value networks	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mega Services	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long

2. **Does your project consider other scenarios apart from those described in the SRA?. Please briefly describe the new scenario.**

3. **What is the relevance of the business scenarios in your specific project?, What is the percentage of use cases in your project that could be represented by each business scenario?**

Scenario	Relevance	Percentage of Use Cases
Enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Hosting	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

Extended enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Virtual organisations	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Value networks	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mega Services	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

4. For those scenarios relevant to your specific project, please indicate if your use cases exhibit other requirements apart from those described in the SRA for the scenario

Scenario	Specify other Fundamental Requirements
Enterprise	
Hierarchical enterprise	
Hosting	
Extended enterprise	
Dynamic Outsourcing	
Mergers & acquisitions	
Virtual organisations	
Business value	
Mega Services	

PART B: STATE OF THE ART

5. Indicate the middleware solutions used to implement the scenarios and, if any, the infrastructures used to support them. Please also briefly describe if your project is developing enhancements to the middleware

Scenario	Middleware	Infrastr.	Enhancements
Enterprise			
Hierarchical enterprise			

Hosting			
Extended enterprise			
Dynamic Outsourcing			
Mergers & acquisitions			
Virtual organisations			
Value networks			
Mega Services			

PART C: RESEARCH CHALLENGES

6. Please indicate research challenges not described in the SRA resulting from your project.

Categories	Specify other Research Challenges
Functional & Commercial Issues	
Dependability	
Security	
Performance	
Interopeability	
Manageability	
Governance	

Flexibility	
Overarching Challenges	

ANNEX D: Sample of Survey Form for Infrastructures

The Service Oriented Infrastructure NESSI Working Group (www.SOI-NWG.org) is defining the NESSI Strategic Research Agenda on new generation of ITC infrastructures (NESSI-Grid SRA). In this context, business Grid is envisioned as the adaptive service-oriented utility infrastructure for business applications, which will become the general ICT backbone in future service oriented economies.

The aim of this template is to collect the feedback of relevant grid infrastructure projects on the draft of the version of the NESSI-Grid SRA (v1.5). Please notice that you, the contributor, act as representative of a infrastructure and not as an individual contributor himself. Thank you for participating in this survey. Your feedback is important in helping us improve future versions of the NESSI-Grid SRA. Please be concise and specific.

CONTRIBUTOR Name: Organization: E-mail: Project:

PART A: BUSINESS SCENARIOS

1. **What is your opinion on the relevance and timeline of the business scenarios described in the SRA?**

Scenario	Relevance	Timeline
Enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hosting	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Extended enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Virtual organisations	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Value networks	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mega Services	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long

2. **Does your infrastructure support other scenarios apart from those described in the SRA?. Please briefly describe the new scenario.**

3. **What is the relevance of the business scenarios in your specific infrastructure?, What is the percentage of use cases that your infrastructure is currently supporting in each business scenario?**

Scenario	Relevance	Percentage of Use Cases
Enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Hosting	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Extended enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Virtual organisations	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Value networks	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mega Services	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

PART B: STATE OF THE ART

4. For those scenarios supported by your infrastructure, indicate the middleware solutions used. Please also briefly describe if your project is developing enhancements to the middleware

Scenario	Middleware	Enhancements
Enterprise		
Hierarchical enterprise		
Hosting		
Extended enterprise		
Dynamic Outsourcing		
Mergers & acquisitions		
Virtual organisations		
Value networks		
Mega Services		

ANNEX E: Sample of Survey Form for Middleware Solutions

The Service Oriented Infrastructure NESSI Working Group (www.SOI-NWG.org) is defining the NESSI Strategic Research Agenda on new generation of ITC infrastructures (NESSI-Grid SRA). In this context, business Grid is envisioned as the adaptive service-oriented utility infrastructure for business applications, which will become the general ICT backbone in future service oriented economies.

The aim of this template is to collect the feedback of relevant grid middleware projects on the draft of the second version of the NESSI-Grid SRA (v1.5). Please notice that you, the contributor, act as representative of a grid research project as not as an individual contributor himself. Thank you for participating in this survey. Your feedback is important in helping us improve future versions of the NESSI-Grid SRA. Please, be concise and specific.

CONTRIBUTOR
Name:
Organization:
E-mail:
Project:

PART A: BUSINESS SCENARIOS

1. **What is your opinion on the relevance and timeline of the business scenarios described in the SRA?**

Scenario	Relevance	Timeline
Enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Hosting	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Extended enterprise	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Virtual organisations	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Value networks	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long
Mega Services	□ None □ Low □ Medium □ High	□ Never □ Short □ Medium □ Long

2. **Does your infrastructure support other scenarios apart from those described in the SRA?. Please briefly describe the new scenario.**

3. **What is the relevance of the business scenarios in your middleware solution?, Can the scenario be implemented by the last release of your technology?**

Scenario	Relevance	Grade of Implementation
Enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Hierarchical enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

Hosting	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Extended enterprise	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Dynamic Outsourcing	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mergers & acquisitions	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Virtual organisations	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Value networks	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%
Mega Services	□ None □ Low □ Medium □ High	□ 0% □ 25% □ 50%□ 75% □ 100%

PART B: STATE OF THE ART

4. **For those scenarios relevant to your middleware solution, please indicate the requirements that are not currently met by your middleware and if you are planning to provide such missing functionality.**

Scenario	Unmet Requirements
Enterprise	
Hierarchical enterprise	
Hosting	
Extended enterprise	
Dynamic Outsourcing	
Mergers & acquisitions	
Virtual organisations	
Value networks	
Mega Services	

ANNEX F: Contributors

This document has been provided by the NESSI-Grid consortium as well as numerous community contributions.

Expert	Country	Affiliation & Main contribution
Andrea Manieri	IT	Engineering, NESSI-Grid: recommendations
Daniel Fey	HU	Nokia-Siemens Networks, NESSI-Grid
Elies Prunés Soler	ES	ATOS Origin, NESSI-Grid: business scenarios and business indicators
Eric Madelaine	FR	INRIA, NESSI-Grid: state-of-the art
John Easton	UK	IBM, NESSI-Grid: Basic model, business indicators
Juan José Hierro	ES	Telefonica, NESSI-Grid: community involvement
Kumardev Chatterjee	BE	Thales, NESSI-Grid: architectures
Mike Fisher	UK	BT, NESSI-Grid: research challenges
Wolfgang Theilmann	DE	SAP, NESSI-Grid: overall lead and business scenarios

Editor-in-chief:

Ricardo Jimenez-Peris (Universidad Politecnica de Madrid, UPM).

Grid Community Liaison Coordinator for SOI-NWG:

Ignacio Martín Llorente (UCM).

Participants of the SOI-NWG Workshop (June 21, 2007):

- Mike Boniface (IT Innovation) representing SIMDAT and GRIA, András Boros (Nokia Siemens Networks) representing NESSI-Grid, Juan Cáceres (Telefonica) representing NESSI-Grid, Michel Cezon (INRIA) representing NESSI-Grid, Kumardev CHATTERJEE (Thales) representing NESSI-Grid, Miguel ESTEBAN (UPM) representing OntoGrid, Yu Feng (INRIA) representing NESSI-Grid, Mike Fisher (BT) representing NESSI-Grid and NextGrid, Francisco Garijo (Telefónica) representing NESSI-Grid, Claudio Grandi (NFN & CERN) representing EGEE and GLITE, Hans-Christian Hoppe (INTEL) representing SIMDAT, Ricardo Jimenez-Peris (UPM) representing NESSI-Grid, Ignacio Martín Llorente (UCM) representing NESSI-Grid and Globus Alliance, Andrea Manieri (Engineering) representing NESSI-Grid, Philippe Massonet (CETIC), Fco. Nieto (ESI), Veronique Pevtschin (Engineering), Ignacio Soler (ATOS ORIGIN) representing BEinGRID, GRIDCOMP and NESSI-Grid, Antonio Puliafito (Univ. Mesina), Konstantinos Tserpes (NTUA) representing NEXTGRID, Wolfgang Theilmann (SAP)representing NESSI-Grid, Vladimir Vlassov (KTH) representing GRID4ALL, Stefan Wesner (HLRS) representing AKOGRIMO and BREIN, Philip

Wieder (FZJ) representing UNICORE and Gabriel Zaquine (CSSI) representing EGEE.

Participation in the Call for Contribution:

- **Research Projects**: ASSESSGRID by Odej Kao (TU Berlin), A-WARE by Sergio Bernardi (CINECA), BRIDGE, EDUTAIN@GRID by Thomas Fahringer (University of Innsbruck), G-ECLIPSE by Harald Kommayer (NECLAB), GRID4ALL Project by Vladimir Vlassov (KTH), GRIDCOMP by Marco Danelutto (UNIPI), GRIDTRUST by Phillippe Massonet (CETIC), HPC4U by Géry Schneider (IBM), KWF GRID by Steffen Unger (FIRST), ONTOGRID by Oscar Corcho (UPM), SIMDAT by Mike Boniface (IT Innovation), SORMA by Erel Rosenberg (Correlation Systems Ltd), XTREEMOS by Bernd Scheuermann (SAP), ProGRID by Peter Stoll (SCIENCE COMPUTING), UNIGRIDS (UNICORE) and BEinGRID by Elies Prunes (ATOS Origin).
- **Infrastructures**: CESGA by Carlos Fernández (CESGA) and EGEE by Gabriel Zaquine (CERN).
- **Middleware**: GRIA by Mike Boniface (IT Innovation), gLite, UNICORE by Achim Streit (Forschungszentrum Jülich GmbH), GRID4ALL Middleware by Vladimir Vlassov (KTH), Globus GridWay by Eduardo Huedo (UCM) and KNOWARC by Peter Stoll (SCIENCE COMPUTING).
- **Individual Contributors (Experts)**: Craig A. Lee (the Aerospace Corporation/OGF President), Yvon Jégou (INRIA), Oscar David Sánchez (INRIA), Bernd Scheuermann (SAP), Daniel Galindo and Luis Pablo Prieto (TI+D), Daniel Vladusic (XLAB), Gregor Pipan (XLAB), Marjan Šterk (XLAB), Matej Artač (XLAB), Silvana Muscella (Metaware), Steve Crumb (OGF), Mark Linesch (OGF), David Snelling (Fujitsu Labs UK), Ali Anjomshoaa (Mobability Limited), Ian Osbourne (Intellect UK), Rubén S. Montero (UCM), Antonio Puliafito (University of Messina), Syed Naqvi (CETIC), Philippe Massonet (CETIC), Pierluigi Ritrovato (CRMPA), Björn Kolbeck and Jan Stender (Zuse Institute Berlin).