Project Number:	**FP6 – 2005 - IST5 - 033634**
Project Title:	Grid Economics and Business Models
Deliverable Number and Title:	Deliverable 1.1 Scenario Interim Report
Contractual Date of Delivery to EC:	October 2006
Actual Date of Delivery to EC:	December 2006
WP contributing to the Deliverable:	WP1
Nature of the Deliverable:	R
Deliverable Security Class:	PU
Editor:	Steve Wallage (451), Rob Blaauboer (LCMG)
Contributors:	AUEB, ATC, RTEL, GISP, IU, LCMG
Current Version:	0.9

Abstract: The main objective of this document is to give an overview of available scenarios in the area of Grid Computing.

Executive Summary

This document lists and provides basic classifications of scenarios from the GridEcon Technical Annex as well as a number of new scenarios from other EU Grid projects or consortium partners. It functions mainly as a repository of Grid scenarios, classifying them for subsequent analysis described within deliverables D1.3 and D1.2. This document will be the basis for the final deliverable D1.2.

List of Abbreviations

Abbreviation	Explanation
USP	Utility Service Provider
OMS	Order Management System
FZJ	Forschungszentrum Jülich
RZG	Rechenzentrum Garching
IDRIS	Institut du Développement et des Ressources en Informatique Scientifique
CINECA	Consorzio Interuniversitario del Nord est Italiano Per ilCalcolo Automatico
CSC,	Centre for Science Computing
BSC	Barcelona Supercomputing Center
HLRS	High Performance Computing Center Stuttgart
LRZ	Leibniz Rechenzentrum
ECMWF	European Centre for Medium-Range Weather Forecasts
DEISA	Distributed European Infrastructure for Supercomputing Application
SLA	Service Level Agreement
HPC	High-Performance Computing
RMS	Resource Management System
EPCC	Edinburgh Parallel Computing Centre

Table of Contents

Table of Figures

1 Introduction

Deliverable D1.1 describes the different scenarios that we consider for the GridEcon project. The scenarios are collected from the Technical Annex of the GridEcon project and from other projects like BEinGrid. In total, we describe 16 scenarios according to a pre-defined template. Any additional scenario can be added to this version or even in the deliverable D1.2 Final Scenario Report.

1.1 *Goal of the document*

The goal of this document is to present scenarios in a way that they can be compared to each other. We do not rank any scenarios (this is done in D1.3) but offer a high-level classification system.

1.2 *More Literature about Grid business models*

The Enterprise Grid Alliance maintains a document on “Reference Model and Use Cases” [1], in which basic components of Grid Business models are defined. The Grid GGF has also set up a repository on business models, which is collection of different cases that it considered [2].

1.3 *Model used to describe scenarios*

We describe the scenario according to the following template:

Field	Mandatory	Remarks
Scenario Meta data		
Scenario name	YES	Actual title of scenario
Origin of scenario	YES	Can be Technical Annex, other project, partner et cetera
Typical users	YES	e.g. SME, Corporates, Consumers
Kind of benefit	YES	Describe the benefit, short prose rather than key words
Alignment with EU goals	YES	Why and how does this align with EU goals (e.g. increased competitiveness)
Classification	YES	Classify scenario as resembling: • Interconnection of grid systems • Service Oriented Architecture • Software as a service

Scenario Information		
Synopsis	YES	Create a synopsis of the scenario identifying key points
SWOT analysis	YES	Fill in SWOT analysis
Complexity	YES	Explain the complexity (short prose)
Ambition	YES	Explain the ambition level
Actors / entities in scenario	YES	Describe which entities are involved in the scenario (e.g. end-user, resource provider, resource brokers)
Additional remarks	NO	
Schematic overview	YES	Draw high level overview showing actors, acts, movement of money etc
Long scenario description	NO	Please fill in as much information as possible to describe this scenario. We appreciate that little will be available

Meta Data about Scenario Record		
Author	YES	The author of the scenario
Filled in by	YES	The name of the scenario contributor
Date	YES	Date filled
Remarks	NO	Any comments you wish to share

2 Overview of the grid market for commercial users

2.1 *What is grid?*

Analysts, The 451 Group, have interviewed over 250 enterprise users who are using some form of grid computing. There has been much debate within both academic and commercial circles as to what exactly is meant by grid computing. Under the strict early definitions of grid computing which referred to 'multi-partner and multi-domain', there are no current grid implementations within commercial users. Although some grid deployments are now global in nature, they remain within the corporate firewall. We now have a wide range of different types of vertical users who have evolved their HPC requirements beyond simple clusters towards a loosely coupled grid infrastructure. Those at the leading edge of this charge, such as the investment banks, are now putting non-HPC applications on their grid infrastructures, and using their internal grids as a basis for a shared IT platform across their organizations. But there is no consensus among users as to what to call these deployments.

The 451 Group asked its user base what was the best term to describe their 'grid' deployments. Over 70% said there was a better term than grid, but again there was no clear consensus over what this term should be.

Best term	Percentage of respondents
High Performance Computing	23%
Virtualization	21%
Utility computing	19%
Clustering	19%
Service Oriented Architecture	15%

2.2 *What are the key drivers?*

The 451 Group also asked its user base what was driving their grid deployment forwards, and what had been the initial driver for adoption.

Key driver	Percentage of respondents
Improved performance	77%
Save money	57%
Do new things	41%
Competition	41%
Time to market	26%

It is clear that, to get management approval, a grid deployment needs to be able to make sense financially. Early deployments often talked about the distributed power of the grid, or the ability to use under-utilized resources, ensuring that they could do something in hours that had previously taken days to do. This greater utilization of resources, with some users reporting server utilization rates going from 10-20% to 80-90%, clearly had benefits in terms of reducing the expenditure needed on hardware.

It is also worth noting that hardware is often of lesser importance in terms of the total cost of ownership to commercial users. Studies have suggested that hardware can account for only around 9-11% of the total IT cost, with software and people costs the two main components. This has led some grid users to gain the largest financial benefits from better usage of their software licenses, particularly in those industries where the software licenses are still expensive and restrictive in their deployment.

However, as users have evolved beyond the first year of their grid deployments, it has quickly become apparent that the real benefit is in terms of greater performance and the ability to gain business benefits. Examples abound including:

- Financial analysts who are now able to add additional variants and run additional scenarios on exotic trades and derivatives, thus allowing them to make more accurate calls on financial instrument movements.
- Oil industry analysts are now able to do four dimensional modelling (the fourth dimension being time) when undertaking seismic evaluations. When it can cost over €100m to drill a well, the benefits of more accurate assessments are clearly meaningful.
- Insurance analysts who used the additional compute power of a grid deployment to work out new insurance quotations during Hurricane Katrina on an hourly basis. Competitors who did not have access to such compute resources, were not able to provide new insurance quotations until the end of the hurricane.

One further driver for certain industries has been compliance and regulation. Meeting new and increasingly complex regulation often requires much additional compute power. This has been a particular issue for the financial industry, with the new MIFID regulations seen as a key driver for European banks to further invest in grid infrastructures.

2.3 *Barriers to adoption*

A lot of commercial users have never considered deploying a grid. Reasons vary from ignorance and lack of awareness, no HPC applications and therefore less obvious drivers to deploying a grid, and to a lack of proven case studies. Indeed, many users tell us that they do not want to be early adopters of grid technology and want to see and hear from users in the same vertical who have gained clear benefits and overcome some of the challenges. This is particularly true for SMEs who do not have the internal resources to overcome the challenges faced by the early adopters, who tend to be large enterprises.

At The 451 Group we particularly found that many users had a grid infrastructure in place that was running one or two HPC applications such as Monte Carlo simulations, but were unwilling to either put additional applications onto their grid, or extend their grid infrastructure by adding additional hardware or scavenge from existing resources. The following table shows their reasons for not developing their grid deployments.

Issue	Percentage of respondents
Software licensing	48%
Cultural	43%
Data management	34%
Bandwidth	31%
Security	27%
Grid enabling applications	21%
Skills shortage	17%
Lack of standards	16%
Prove RoI	16%

Looking in a little more detail at each of these issues:

- Software licensing – as with many of these issues, the impact varies between vertical markets. This is a major problem in markets where software is expensive and the market is dominated by a few software vendors. A classic case is the electronic design automation market, where software licenses can cost up to €1m. The market is also dominated by three vendors, who have been able to ensure that each software license is tied to one CPU. In theory, if a user was running the software across a 1,000 node grid they would need to pay for 1,000 software licenses. There are some ways around this problem such as users buying unlimited enterprise usage versions of licenses, but often users have had to rely on putting pressure on software vendors or seeking out open source alternatives.
- Cultural challenges vary widely from fear of the unknown to losing control to concerns over a shared environment. Often, a grid deployment starts in one particular department or group, such as the seismic processing team within an oil company. While it may make sense to broaden the usage of the grid to other departments, there is often much suspicion from these other departments as to how well this would work, and whether they would lose control over their own resources. There are also often concerns at the individual engineer or researcher level who are loathe to allow their computers to be used as part of a grid, and fear that a standardized approach would weaken their own role in the company.
- As users move beyond embarrassingly parallel applications, then data management becomes a real challenge. This includes many aspects include storage, caching, data movement, data transfer and federation. Put simply, if you are running a job over multiple machines how can you ensure that the data is in the right place in the right format. This is a particular challenge for experienced grid users such as semi-conductor design companies who typically run thousands of jobs on their grids every day. It also creates major headaches for users, such as the banks, who have stringent regulatory pressure to know where and how each job is processed.
- Bandwidth concerns refer to cost and availability of high-speed bandwidth. Many applications require a low latency for data exchange between or data representation and, therefore, need very good computer network links.

- Security would be a much larger issue, were it not for the fact that users are still operating within their own firewalls. However, there are still concerns over the transmission of confidential data over a grid, and security is always the first concern of the senior management. Security can also be an issue with regard to different departments who are worried that other groups will be able to see their work, which could be either a problem in regulatory terms or simply in terms of internal competitiveness.
- Grid enabling applications is a problem as many large software vendors have yet to optimize their software for a grid or distributed environment. The software vendors usually claim they are waiting to see clear demand, although their own reticence is slowing user demand.
- Skills shortage will become an increasing issue as IT staff with specific 'grid' knowledge become harder to find. This is already been seen in the investment banking industry where individuals who have led grid architecture teams are becoming very sought after.
- Despite several initiatives, there is a real lack of standards in the grid world. Users are often stuck with proprietary offerings from vendors, and have to make a judgment as to whether that vendor will be able to continue to develop their products and be a long-term player in the market.
- Users clearly want to be able to prove the return on investment on their grid deployments. Commercial vendors have tried to help in this area with grid assessment tools, and in running proof of concept trials to try and show the benefits. A particular challenge is that many of the benefits of grid, such as improved performance or greater collaboration, are difficult to measure. A good example is the pharmaceutical industry where the key usage of grid technology is in drug discovery. Yet, no grid user has yet been able to claim that they discovered a new drug on their grid infrastructure.

2.4 *Vertical comparisons*

Within different vertical markets, there are particular applications that have tended to be the first to be deployed on a grid. These are typically HPC applications that may previously have been run on a cluster or a supercomputer. The secondary tier may also be traditional HPC applications but will start to include other types of applications. As the different verticals move on to a third tier, they start to look at how they can run more traditional and front-office software on their grids. The starting point for such an exercise is often data mining or business intelligence software.

Vertical market	Initial applications running on grid	Secondary applications running on grid
Investment banks	Monte Carlo simulations	Risk analysis, exotic trades
Pharmaceutical	Drug discovery	Clinical modelling
Digital media	Rendering	Animation
Insurance	Actuarial regulations	Risk management
Manufacturing	Crash and clash testing	Collaborative software

Healthcare	Patient records	Sharing data between hospitals and health centres
Telecom (internal usage)	Billing	Operations management

2.5 *Future developments*

The 451 Group has mapped user evolution of their grid deployments onto a five-stage model.

Stage	Description
1	Running trials
2	Single application
3	Siloed grids – single grids or grids in multiple departments that are not linked
4	Linked grids – with or between departments and with multiple applications
5	Internal shared utility or global platform

Most of the grid users in The 451 Group database are at Stage 3. Many are considering Stage 4 and facing the challenges listed in Section 2.3.

The leading adopters are the investment banks who all have plans in place for Stage 5 deployments. Some have already created an internal shared utility, which is based on a global grid infrastructure, and offers compute resources to other departments.

3 Overview of scenarios

3.1 *Accessible super computing for SME*

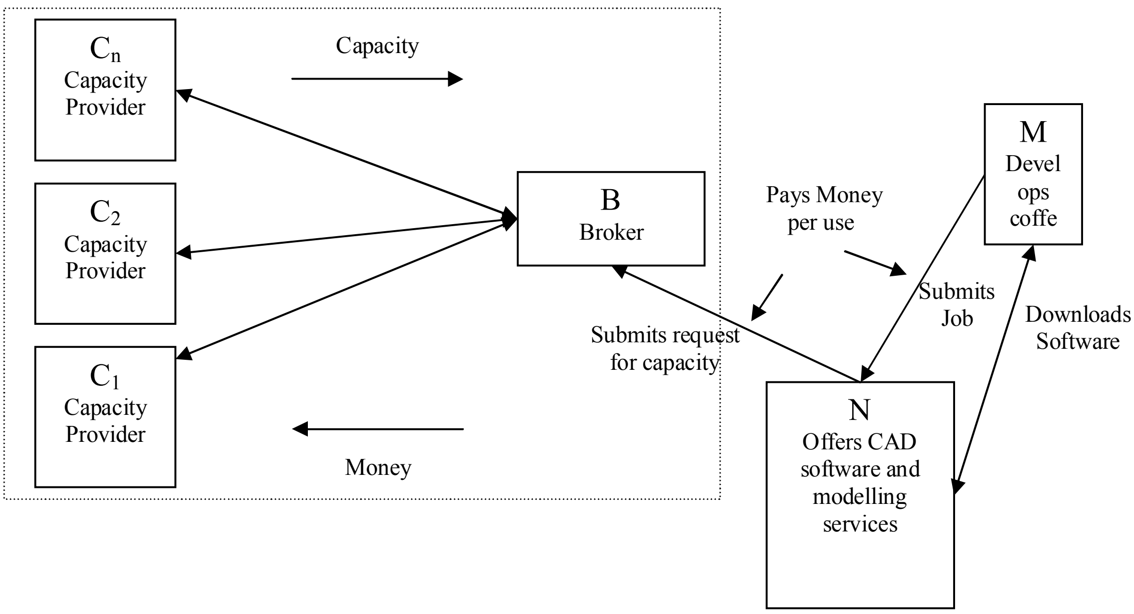
3.1.1 Scenario Meta data

Scenario name	Accessible Super Computing for SME
Origin of scenario	Technical Annex
Typical users	SME
Kind of benefit	New capabilities, originally not available to SME
Alignment with EU goals	Improves the competitive position of European SME by offering them the possibility to use super computing, something that was hitherto not possible. The offering is service-based, offering modelling tools and charges for software usage and super computer usage on a pay per use basis.
Classification	Interconnection of grids

3.1.2 Scenario information

Synopsis	A SME designs top-level equipment. It wants to improve the quality of its products but does not have the capacity (i.e. required supercomputing capacity) to do advanced modelling and meshing. A software companies offers design software for free and charges for the actual meshing and modelling tasks on a pay per use basis. The software company sends any modelling tasks to a broker that negotiates the best available execution platform (based on a number of parameters like time and price).
Strengths	The scenario employs the Internet as a front end / delivery channel, combined with 'free software' to model the products. The user only pays for the modelling services in a very transparent way (i.e. you only pay what you use). This approach reduces the cost of ownership.
Weaknesses	Relatively low number of users. Not all SMEs are involved in actually developing products; many of them are in the services business / sales. Since this kind of services is not available to SMEs, how many will have the skills (e.g. in terms of FTEs) to do this kind of work?
Opportunities	This is a new service currently only available to large corporations that now becomes available for SMEs.
Threats	This kind of service might fail if the set of potential customers (SME) is too small.

Complexity (technical)	The complexity is considered to be medium since the process is broken down in identifiable pieces (Capacity supplier – broker – software supplier – user) with clear boundaries and tasks. The broker however has the challenge to negotiate the best offer based on a number of criteria.
Ambition	The ambition level is considered to be medium. The scenario seems to be viable and realistic (it does not involve any paradigm shift or discontinuities)
Actors / entities in scenario	End Users = SME, Hardware Resource Providers, Software Resource Brokers, Brokers
Additional remarks	None
Schematic overview	



Long scenario description

Assumption, a specialist European company, M, designs and manufactures top-of-the-range kitchen equipment: blenders, mixers, coffee machines and toasters. It would like to use numerical modelling and simulation to optimise the performance of its products; e.g. to improve the noise, vibration and harshness of its mixers or to model the thermal properties of its toasters.

However, being a small enterprise, it cannot afford to have specialist staff to develop these modelling technologies and certainly cannot afford to purchase and support the high-performance parallel clusters that would be needed to model their designs to the required degree of fidelity.

Software Company, N, has developed a design optimisation service. Company M downloads the front-end CAD system for this service and uses this front-end freely to digitise its designs. These are then submitted to the meshing and modelling components of N's service that are available on a pay-per-use basis. When each modelling job is submitted, a broker finds the best currently available execution platform, utilising high-degrees of parallelism to achieve accurate prediction and rapid turn round. For each run, appropriate payments are remitted to the software provider, N, for the use of its modelling software and the operator of the selected execution platform.

Company M uses this optimisation facility as an integral part of its design process. As a result its products gain a reputation for efficient and pleasant operation and its sales flourish. Over time the software company, N, improves the performance, speed and accuracy of its modelling routines and the utility company vendors replace their equipment with new, faster processors. All these improvements are available, transparently, to M without it having to retool or even change its purchase arrangements for software and execution.

3.1.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Rob Blaauboer
Date	26/10/2006
Remarks	-

3.2 *BEinGrid experiment 15*

3.2.1 Scenario Meta data

Scenario name	BEinGrid Scenario 15 – Data Recovery System
Origin of scenario	BEinGrid TA / D0.2.1
Typical users	SME / SOHO
Kind of benefit	Data recovery services for small and medium sized enterprises, as well as for small office and home office (SOHO) using redundant peer-to-peer storage
Alignment with EU goals	Actual experiment to see if Grid technologies can be deployed commercially in SME / SOHO environment
Classification	Interconnection of Grids

3.2.2 Scenario information

Synopsis	Data Recovery Service (DRS) is a service, which provides SME (Small Medium Enterprises) and/or SOHO (Small Office Home Office) with recovery services for business critical applications, information and data. Basically, user's backups are distributed and stored as small encrypted packages on a number of peers (PCs) from other users / companies. Users are unaware of the nature / content of the packages that are stored on their PCs. In order to avoid a single point of failure (a PC that crashes with data on it) the packages are stored in a redundant fashion so if any PC fails, the recovery can still be executed. The percentage of redundancy will be determined by calculating chance of failure. Companies that want to use the DRS services contact a trusted third party who arranges contractual arrangements, authentication of users and possible financial settlement. The Grid service provider delivers the PCs.
Strengths	Novel concept, online storage / DRS is in high demand. Since there is no single point of failure it should be considered extremely safe (in theory).
Weaknesses	Peer to peer is associated with illegal activities. The method is complex especially when compared with current online storage offerings.
Opportunities	Since it is a peer-to-peer system (all share the burden) it can be a solution for companies who do not have the capacity / money.
Threats	Storage is quite cheap (in theory GMAIL offers already 2GB of storage for free). Other banks (e.g. ABN AMRO) also offer storage (secure with a challenge response security) to the same target audience and consumers (traditional).
Complexity (technical)	The DRS service is considered to be technically complex.

Ambition	The ambition level is high, both technically and as the market goes. Will users buy into the concept?
Actors / entities in Scenario	End Users = SME , Resource Providers = SME (including SUPERPEER) Resource Brokers = Trusted 3rd party
Additional remarks	None.

Schematic overview

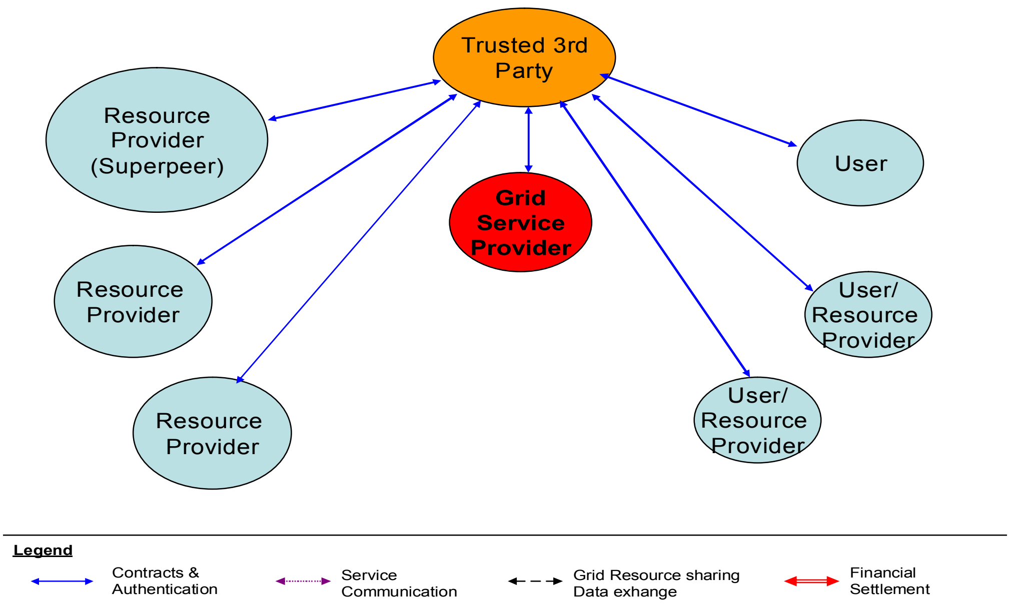


Figure 3.2.2-1 Contract Overview

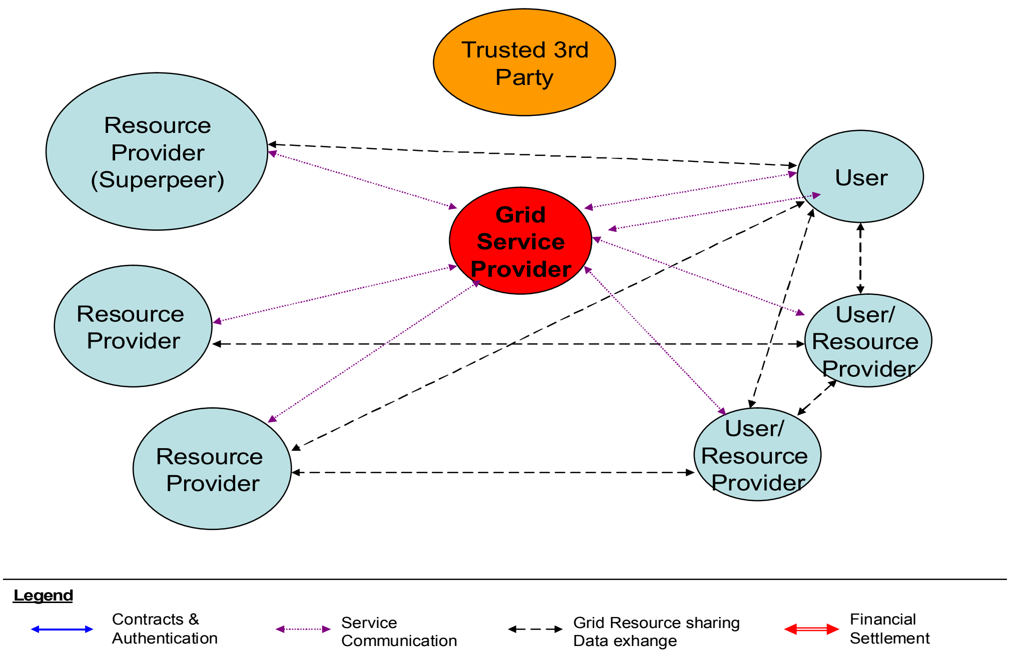


Figure 3.2.2-2 Operational DRS Overview

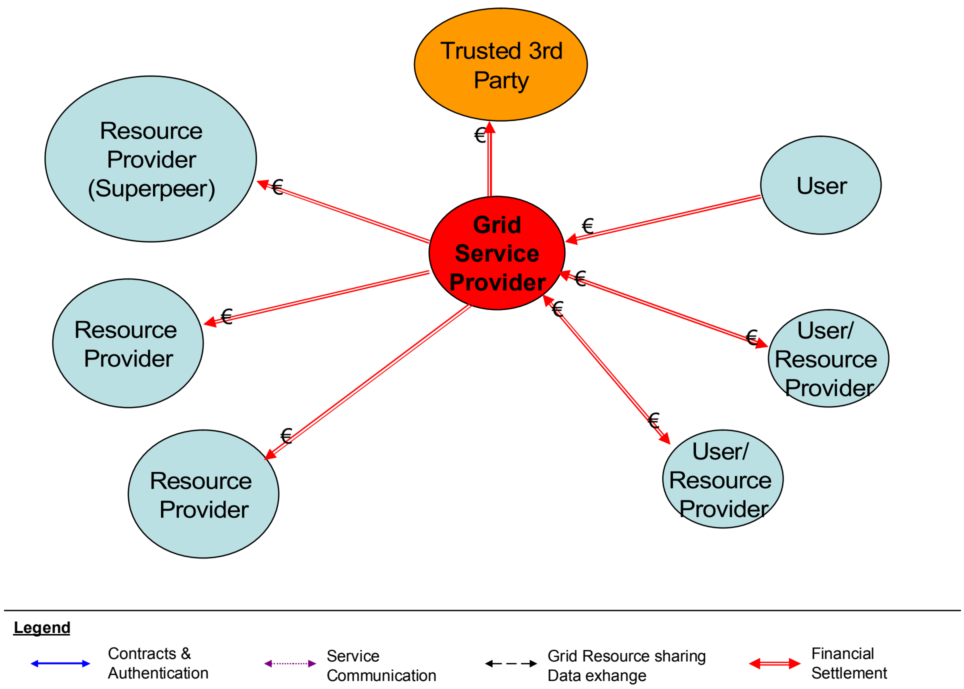


Figure 3.2.2-3 Financial Flow

3.2.3 Meta data about scenario record

Author	BEinGrid
Filled in by	Rob Blaauboer
Date	2-11-2006
Remarks	-

3.3 *DEISA scenario*

3.3.1 Scenario Meta data

Scenario name	Interconnecting High-Performance Computing Centres
Origin of scenario	DEISA - Distributed European Infrastructure for Supercomputing Application
Typical users	Scientist (physicist, biologist, mechanical engineering, etc), Industry users
Kind of benefit	Hiring computation resource and application
Alignment with EU goals	It promotes the application of HPC in industry and between HPC sites
Classification	Interconnection of Grids

3.3.2 Scenario information

Synopsis	This scenario allows industry users and scientist to access high-performance computing resources across several organizations. The sum of the resources is greater than the resources available at one single site. The interconnection requires to solve technical problems with respect of interconnection of heterogeneous systems as well as issues about policies for resource allocation.
Strengths	The scenario supports executing a HPC workflow as well as HPC application on reserved Grid resources within the scope of a business contract. The purpose of the SLA is to identify the shared goals and objectives of the concerned parties. A good SLA is important as it sets boundaries and expectations for the following aspects of a service provisioning. An SLA clearly defines what the user wants and what the provider promises to supply, which helps to reduce the chances of disappointing the customer. Provider's promises also help the system stay focused on customer requirements and assure that the internal processes move in the right direction. An SLA describes a clear, measurable standard of performance. Based on this description, internal objectives become clear and measurable. An SLA defines penalties. This criterion make the customer understand that the service provider truly believes in its ability to achieve the set of performance levels. It makes the relationship clear and positive.
Weaknesses	Not all users need this service. The targets of this service are industry users, who need execute HPC application or HPC workflow.
Opportunities	This is a new service, which has not been addressed in literature fully.
Threats	Here we can see the broker is a centralized service. If it is

	broken, the whole system may collapse. To avoid this problem, some methods like redundant equipment or data backup have to be employed.
Complexity (technical)	To realize this scenario, many issues, which are not considered in the present system, must be solved. Some of them are: An effective mapping mechanism to map each sub-job of the workflow to resources in a manner that can satisfy two main criteria: being able to finish workflow execution on time and being able to optimize the job execution cost. The first criterion is quite clear because it is the main reason for an SLA system to exist. The latter criterion is derived from the business aspect of a SLA. If a customer wants to use a service, he must pay for the service usage and has the right to receive it with an appropriate quality. An automated mapping, which considers economic parameters, is necessary as it frees operator from the tedious job of assigning sub-jobs to resources under many constraints such as workflow integrity, time condition, etc. Additionally, a good mapping mechanism will help users to save money and to increase the efficiency of using Grid resources. A billing stack system for accounting and charging. This system will be the basis for recording the performance parameters that were defined in the SLA. A mechanism to handle the error, which may occur during the execution of the workflow. Randomly appearing errors may damage the workflow completion as well as the negotiated SLA. Thus, it is demanding to build an error recovery mechanism for a workflow in order to eliminate the affection of error to users and to make the Grid system more stable and reliable.
Ambition	The ambition level is considered to be medium. The scenario seems to be viable and realistic (it does not involve any paradigm shift or discontinuities)
Actors / Entities in Scenario	end-users = corporations, HPC Centres Resource Provider = HPC Centres Resource Broker = DEISA consortium
Additional Remarks	None.

Schematic **Overview**

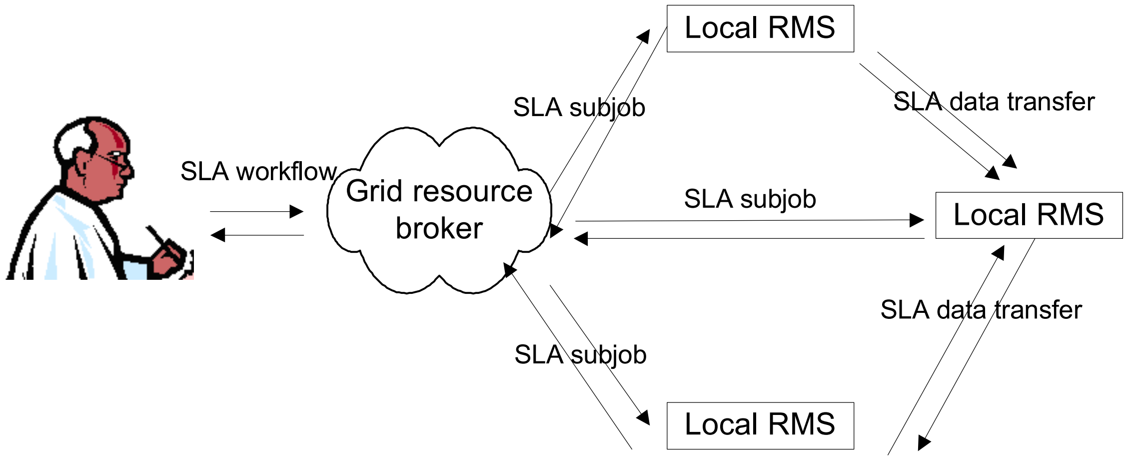


Figure 3.3.2-1 Job mapping

Long scenario description

Benefits and incentives for interconnecting HPC

The benefits of using interconnected HPCs have long been realized and led to the creation of the DEISA grid. The advantages of interconnecting HPC centres includes:

- Load balancing between high performance centres
- Persistent, reliable computational power on a continental scale
- Transparent operation of the grid
- Global data management
- Enabling scientific discovery

There are many incentives for opening the access to the DEISA grid. Some of these are:

- Access to a continental supercomputer
- Higher resource utilization by allowing other scientists to use idle resources
- Financial benefits from selling unused computing power to industry

The characteristics of resources in DEISA

The DEISA resources are supercomputers or super clusters, which are managed by different DEISA sites. The DEISA partners are individual organizations (FZJ, RZG, IDRIS, CINECA, CSC, BSC, HLRS, LRZ, ECMWF) with different financial schemes and funding models, which can make resource sharing a very complex matter. Each of the DEISA partners participating in the DEISA Grid provides a fraction of its computational resources (in general, more than 10%) to a DEISA resource pool, which is globally managed by DEISA Consortium.

The main objectives of the DEISA Grid enabled research infrastructure are:

- To deploy and operate a persistent, production quality, distributed supercomputing environment with continental scope
- To enable scientific discovery across a broad spectrum of science and technology. Scientific impact (enabling new science) is the only criterion for success.

The software architecture of the DEISA Grid is specific to the needs and the requirements of a virtual European supercomputing centre. They include Unicore to support the execution of single jobs and workflows, and GPFS to support a global data system. Other necessary components like co-scheduling services for distributed applications, Portals and Web services are under development.

Users of DEISA

Users of DEISA are only a small fraction of those users using High Performance Computing and they have a high demand for computational power. Those users include:

- Scientific users from countries joining DEISA
- Scientific users from third party countries having no resources contribution to DEISA
- Industry users

In theory, all those users can use the resources in DEISA but with different priorities and limitation.
The applications from those users include:

- Extreme computing demands for challenging projects requiring a large fraction of a single supercomputer
- Workflow applications involving at least two supercomputers
- Coupled applications involving more than one platform. These are applications that consist of separate modules, which only exchange moderate amounts of information, such as multi-physics, multi-scale applications.

Working flow of DEISA

A user having a project (a HPC application) will perform several steps, which are applied for all users:
Step 1: The user submits the project to the national evaluation committee. The committee will decide whether the research project may acquire the DEISA label or not. This decision is based on scientific relevance and scientific excellence factors.

Step 2: Next, a technical validation of the project is performed by the DEISA technical teams. The goal is to determine if the proposal requires the supercomputing Grid and that the application can be deployed with a reasonable amount of effort.

Step 3: The DEISA Consortium decides how and when the DEISA resources are used. It establishes priorities among the projects, which were validated by the national evaluation committees, and decides on how the resources will be managed in each case. If the user is from a third party country, he can use at most 3% of the resources at each site. If the user is from industry, he will be charged by each site that actually provides the resources. The Consortium will negotiate internally, on a per project basis where the resources will come from and how the project will be operated in the DEISA environment. Naturally, this negotiation is part of the selection process.

The DEISA Consortium makes decisions based on resource balance usage criteria. The usage of the resource pool must be balanced over a sufficiently long period of time, so that each partner has the opportunity to recover as much as he has contributed during the time period.

Step 4: The user moves the job to the specific resources at a specific time.

Some open issues with DEISA approach

From the working mechanism described above, we would like to make the following observations:

- The DEISA Grid is at the early phase of management and operates with too much interference from administrators.
- The mechanism does not encourage users to use the grid due to the complicated procedures and slow response time.
- The DEISA Consortium plays the role of a scheduler (or broker), which is only possible when the size of the Grid is very small. Once the Grid size increases, man-made scheduling is not an option due to its lack of speed. It could be the case that the policy at each site is too complicated for a simple broker program.
- The consortium provides services for users and charges them (in the case of industry users). However, there are no mechanisms to ensure the right of user. For example, if the job is late because of resource failure at a site, there is no mechanism in place to compensate the user. It seems that DEISA does not need to attract customers and users have no choice but to use DEISA.

A specific problem: Scheduling

We would like to give an example of a workflow. For an industry user, the main requirement is finishing the workflow on time according to the SLA. The present working mechanism at DEISA cannot meet this requirement. We will discuss only the scheduling problem here. It is easy to see that if the size of the workflow or the number of DEISA site is large (more than 6 sub-jobs and 10 sites), finding a solution that meets the requirement of the user is a very difficult task. The DEISA consortium has to consider the following factors:

- They have to consider the resource allocation policy at each site over a very long time period.
- They have to consider the dependencies among sub-jobs in the workflow.
- They have to consider the network connection between the DEISA sites.
- They have to consider a lot of combinations that are candidate allocations that solve problem.
- They have to consider the balance of resource usage among sites.

A specific problem: Policy limitation

A third party user who would like to use the DEISA infrastructure for his computation-intensive application cannot do it if it requires more than 3 % of the resources (which is a policy). This policy is even valid if it was known that the supercomputer was idle for the length of the entire run-time period of the application.

The same problem goes for partners within the DEISA consortium. The local national governmental administration will not allow other organizations (which do not belong to the governmental administration) to use more than a certain percentage of the total capacity, since the local governmental administration has to justify the spending of tax money. If the tax money does not contribute to the benefit of the citizen of the local government, the investment in supercomputing would be considered unnecessary.

Possible economic based solutions

Many of these processes mentioned above are dealt with by people even though they could easily be automated. While the evaluation and technical evaluation has to remain in the hands of committees, a number of issues can be simplified by using a middleware system. We propose a system as presented in the following figure.

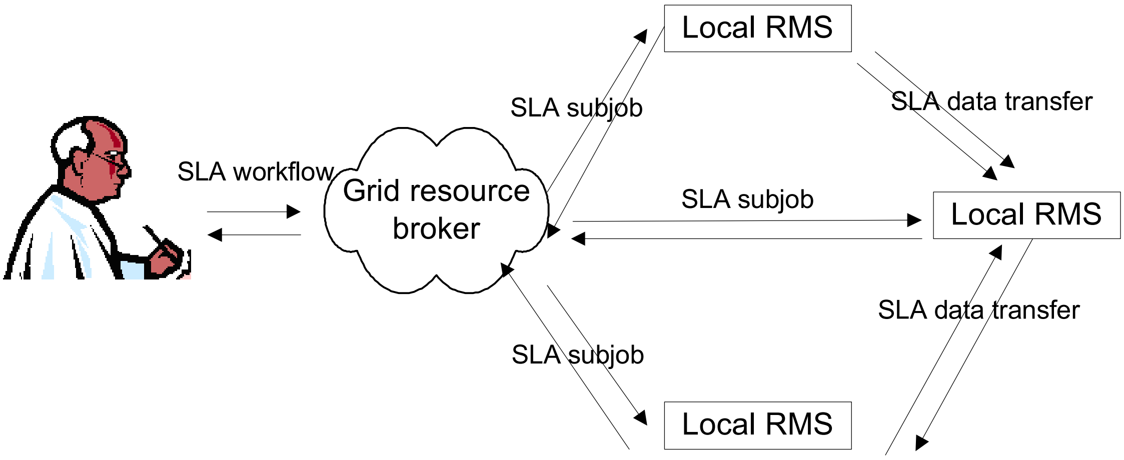


Figure 3.3.2-2 A proposed workflow running scenario

The resource broker will handle the task of the present DEISA committee. Each site contributing to DEISA can use the price of a service to reflect the local policies. For example, in a critical time period, when the resources at a site must be used to support local users, the price of the service will be very high.

For industry users, finishing the execution of the workflow within a specific period of time is a mandatory requirement. As they pay the cost, they have the right to receive high quality service. An agreement between the user and the Grid system is a possible solution. This agreement is called a Service Level Agreement (SLA).

The purpose of the SLA is to identify the shared goals and objectives of the concerned parties. A good SLA is important as it sets boundaries and expectations for the following aspects of a service provisioning. An SLA clearly defines what the user wants and what the provider promises to supply, which helps to reduce the chances of disappointing the customer. Provider's promises also help the system stay focused on customer requirements and assure that the internal processes move in the right direction. An SLA describes a clear, measurable standard of performance. Based on this description, internal objectives become clear and measurable. An SLA defines penalties. This criterion make the customer understand that the service provider truly believes in its ability to achieve the set of performance levels. It makes the relationship clear and positive.

To realize this scenario, many issues, which are not considered in the present system, must be solved. Some of them are:

- An effective mapping mechanism to map each sub-job of the workflow to resources in a manner that can satisfy two main criteria: being able to finish workflow execution on time and being able to optimize the job execution cost. The first criterion is quite clear because it is the main reason for an SLA system to exist. The latter criterion is derived from the business aspect of a SLA. If a customer wants to use a service, he must pay for the service usage and has the right to receive it with an appropriate quality. An automated mapping, which considers economic parameters, is necessary as it frees operator from the tedious job of assigning sub-jobs to resources under many constraints such as workflow integrity, time condition, etc. Additionally, a good mapping mechanism will help users to save money and to increase the efficiency of using Grid resources.
- A billing stack system for accounting and charging. This system will be the basis for recording the parameters that were defined in the SLA.
- A mechanism to handle errors, which may occur during the execution of the workflow. Randomly appearing errors may damage the workflow completion as well as the negotiated SLA. Thus, it is demanding to build an error recovery mechanism for a workflow in order to eliminate the affection of error to users and to make the Grid system more stable and reliable.

3.3.3 Meta data about scenario record

Author	DEISA, IU
Filled in by	Dang Minh Quan, International University in Germany
Date	1-11-2006
Remarks	-

3.4 *Brokerage of enterprise computing and storage capacity*

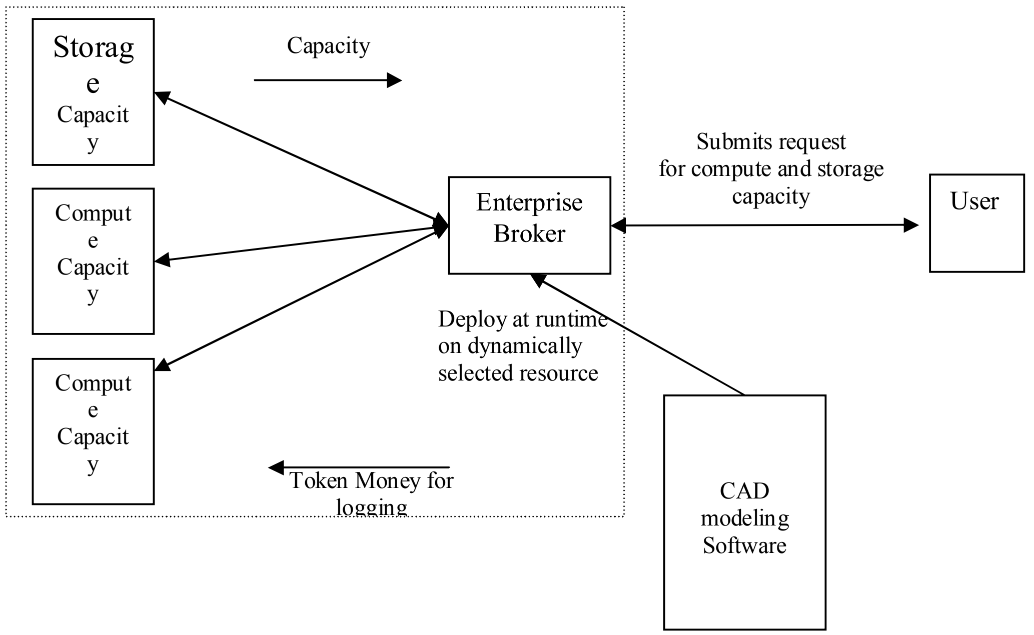
3.4.1 Scenario Meta data

Scenario name	Brokerage of Enterprise Computing and Storage Capacity
Origin of scenario	TA - Based on past experience of London e-Science Centre projects at ICL
Typical Users	Medium to large size Enterprises, End users
Kind of benefit	Improved performance, increased utilisation
Alignment with EU goals	It would lead to improved productivity with better use of existing investments and reduced wastage. It could also lead to cost savings and increase in competitiveness of firms.
Classification	Interconnection of Grids

3.4.2 Scenario information

Synopsis	Enterprises has over provisioned and departmentalized their computing resources in order to make sure that they have enough capacity to handle peak loads. Cost reduction / doing more with available resources / money is now a key driver for most companies. Virtualisation makes it possible to define a platform (that can consists of multiple physical platforms) where jobs / applications can run. This platform can scale up or down depending on demand ending the huge over provisioning (and reducing the associated costs) currently done by enterprises. Enterprises can negotiate and compare SLAs with both internal and external providers, reducing prices and creating transparency.
Strengths	This scenario envisages the deployment of brokers to utilise computational and data capacity. The user uses the broker to find optimal utilisation of his resources instead of dedicated allotment to projects/departments.
Weaknesses	It is only worth the effort if there are enough resources, which warrant the extra effort. It also makes sense only if the efficient allocation would lead to higher utilization, i.e. the resources are not already overstretched.
Opportunities	This could lead to higher return on existing investments. Cost savings due to more efficient use of resources could lead to spending on new product improvements.
Threats	If the brokerage is provided as a centralized service, then it could become the central point of failure. Some strategy to avoid this, such as provision of redundant and reliable backup facilities, could help avoid some of these threats. In some instances, the nature of existing licenses could limit

	the ability to broker resources due to licensing restrictions. Software applications, which require complex set-ups (i.e. are not off the shelf), might inhibit the scale of brokering.
Complexity (technical)	The complexity is considered to be medium since the process is broken down in identifiable pieces (Capacity – Broker – software – user) with clear boundaries and tasks. The broker however has the challenge to negotiate the best offer based on a number of criteria.
Ambition	The ambition level is considered to be medium. The scenario seems to be viable and realistic. However, testing and validation of this in a real world setting is challenging and requires considerable effort.
Actors / entities in scenario	None.
Additional remarks	None.
Schematic overview	



Long scenario description

Enterprises have been heavily consolidating its IT operations and infrastructure over the past decade, focusing on areas like email, messaging, ERP, supply chain, and other horizontal functions. They have typically been successful in standardizing the hardware platforms and applications being used, on centralizing data centres, disaster recovery sites and sharing a common communications infrastructure. However, you will find that most, if not all, hardware and applications are physically dedicated to single business units or specific use cases, resulting in huge overcapacity in terms of software licenses, CPU cycles, and storage capacity.

Virtualization technology allows these companies to define logical resources tailored to specific application's needs, running on and across the actual physical platforms. Offering seamless migration and expansion/contraction of these logical resources depending on the actual needs at any point in time, these 'virtualized' environments need not cater to the sum of the usage peaks of all applications. However, there are currently no market mechanisms embedded in these virtualized data centres.

Application administrators still negotiate a certain SLA, which is statically translated into a certain virtual configuration of processors, memory, storage and bandwidth. The SLA stated in terms of performance or availability is used as a benchmark for monitoring and sometimes penalty payments but not for setting the price. If the demand for resources grows, additional resources must be negotiated and purchased.

Imagine instead an environment where application administrators (the consumer) are able to define their SLA requirements with the enterprise computing and storage broker, these brokers dynamically assign the resources required to meet the SLA. Pricing reflects these dynamic adjustments, charging more during spikes and less when demand is low. If allowed, application administrators can now easily compare internal and external providers of the same service and make more informed purchasing decisions.

From the other side, the infrastructure manager can decide to build a minimal internal infrastructure and buy capacity during peeks externally. Moreover, he/she can sell excess capacity to the market when available. With today's technology, setting up such an environment would take weeks if not months for each application, and days to weeks to increase capacity when needed.

3.4.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Asif Saleem (London e-Science Centre)
Date	6th November 2006
Remarks	-

3.5 *University utility computing*

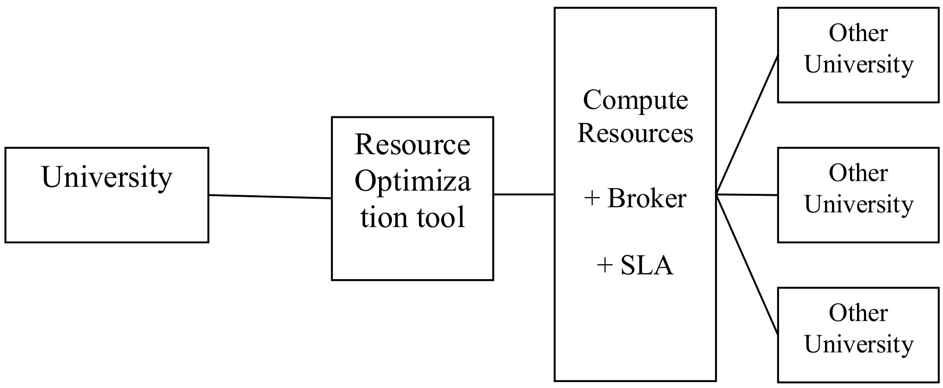
3.5.1 Scenario Meta data

Scenario name	University Utility Computing
Origin of scenario	Technical Annex
Typical users	Universities
Kind of benefit	Enables university to meet peak compute loads and more closely aligns costs with revenue and research projects.
Alignment with EU goals	Enabling European universities to become more globally competitive as they gain access to additional compute resources and align costs more closely with research. For some universities, the ability to offer unused compute resources could also become an important revenue source.
Classification	Interconnection of Grids

3.5.2 Scenario information

Synopsis	HPC users within universities sharing compute resources with similar users. At times of over-supply, the university would provide compute cycles onto the grid. At times of under-supply, it would buy resources on a pay-per-use basis. The university can create a model of predicted demand such that it can maximize the revenues it receives for spare cycles, and minimizes the cost of additional capacity.
Strengths	There are clear benefits for the universities in terms of access to compute resource and reducing costs. As universities are becoming accustomed to working in such a collaborative manner, such a project would not be a major change operationally.
Weaknesses	The model does assume that universities and research departments are able to make reasonable assessments of future demand, and that resources from other universities can provide the functionality they need. The latter is a particular challenge given the heterogeneity of resources and types of job.
Opportunities	There would be great opportunities to expand the scope of the program to include other research institutions and partners. There would also be the opportunity to outsource management of the compute resources to a third party, thus freeing up additional time and monies for the universities.
Threats	University research departments will start to rely again on their own resources if they do not receive the level of compute power and internal SLA that they need.
Complexity (technical)	High. The widely differing compute resources and types of

	job ensure that this will be a complex process.
Ambition	Medium – high. Sharing between HPC users at universities is a medium ambition. However, there is the opportunity to make this a high ambition project by including a broader constituency.
Actors / entities in Scenario	End Users = universities Resource Providers = universities Resource Brokers = Intra-university agreements
Additional remarks	None.
Schematic overview	



Long scenario description

A medium-size European university, U, purchases a high performance parallel cluster to support the computational requirements of its research workers. It is conscious of the fact that the resource is not of sufficient size to meet the peak demand from its workers but that the equipment purchasing cost, support cost, and maintenance cost are high and replacement costs are a continuing capital burden.

The university's HPC support group constructs a two-way gateway between their own resources and the global Internet market in compute cycles. The compute cycle market is provided by utility computing operators.

At times of peak load on the university cluster, the gateway is used to find appropriate resources on the open market and to deploy excess university jobs there. These external resources are used on a pay-per-use basis. At times of low internal load, the university uses the gateway to offer its resources on the open market and to receive revenue for the use of its resources.

A professor of the university specialising in financial computing helps the HPC support group to construct a stochastic optimisation package. It, over time, helps the university to optimise the combination of its internal and external resources. That package buys futures in external resources, sufficient to meet the university's anticipated peak demand, especially for high-priority jobs, while ensuring that the university's resources are offered externally when the spot price is high.

This arrangement allows the university to receive revenue from the use of its resources while catering for its peak demand without ever purchasing expensive excess resources. Importantly, this model changes the university's costs from a recurring capital expenditure to a more revenue-based regime. It is able to directly associate the marginal costs of meeting a particular research group's needs for computation with the results of that research. In this way the university develops a model that directly connects the marginal cost of supporting particular research computationally with the outcome of this work and thus optimises the input (money) and outputs (research results) of its activities.

3.5.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Steve Wallage
Date	1-11-06
Remarks	-

3.6 *Business intelligence*

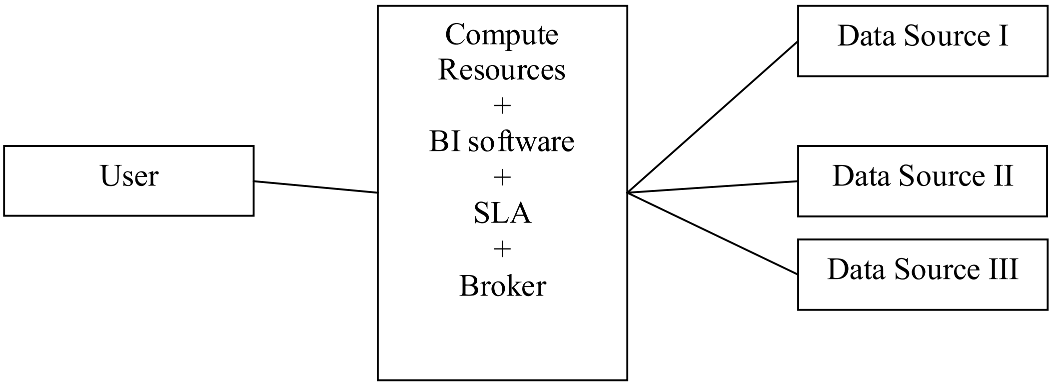
3.6.1 Scenario Meta data

Scenario name	Business Intelligence
Origin of scenario	Technical Annex
Typical users	Corporations
Kind of benefit	There is the potential to save costs but the far greater opportunity are for large corporations to use far more advanced data mining techniques, and for SMEs to gain access to the sort of business intelligence tools that they could never have previously afforded.
Alignment with EU goals	Improving the competitiveness of European corporations and, particularly, SMEs.
Classification	Software as a service

3.6.2 Scenario information

Synopsis	For commercial users to have the opportunity to run sophisticated data mining and business intelligence tools across their datasets.
Strengths	Sophisticated data mining and business intelligence tools can bring real competitive advantages to commercial users, particularly SMEs who have not been able to use such tools in the past.
Weaknesses	Such an ambitious project has some major technical challenges. These include the wide diversity of datasets and formats, security, cost on a pay-per-use basis, ability to use latest software and wide range of user needs. Semantics is also a big issue in business intelligence.
Opportunities	There would be the opportunity to extend this to other software tools on a pay-per-use basis, providing a significant cost and functionality advantage to European business.
Threats	Although the results from data mining and business intelligence tools can be difficult to quantify, users have a clear idea of what they want to achieve. Large corporations, in particular, will expect any service to be at least as good as their existing offering, or will not use a Grid-based service.
Complexity (technical)	High – a very diverse range of needs and data types.
Ambition	High – this fulfils a very specific need for corporate users.
Actors / entities in scenario	End-users = SMEs, larger corporations Resource Provider = SMEs, larger corporations, data providers, Business Intelligence specialists Resource Broker = Third party, independent body

Additional Remarks	None.
Schematic Overview	



Long scenario description

It creates the capability to analyse and generate reports from a huge amount of raw data coming from different data-sources, while reducing response time and keeping cost low. The technical challenge is to significantly improve the performance of the process loading data to the data warehouse and fast access to queries of this data, while allowing for easy scalability in the future, reliability, simplicity and cost-effective hardware. A possible solution may involve the use of a parallel processing Grid, utilizing the parallel processing, and pay-per-usage according to a sophisticated accounting mechanism.

3.6.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Steve Wallage
Date	1-Nov-2006
Remarks	-

3.7 *Museum – Learning and TV channels*

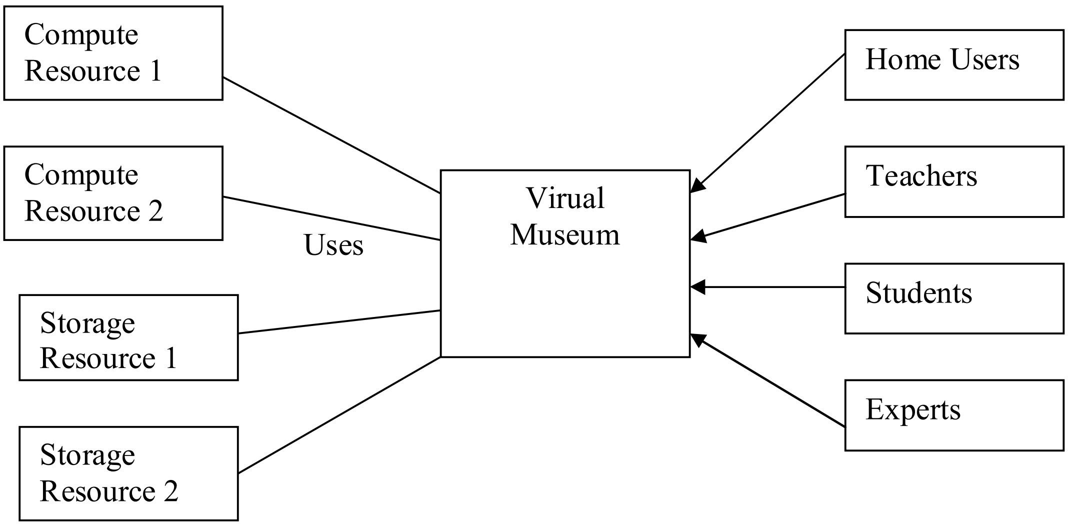
3.7.1 Scenario Meta data

Scenario name	Museum – Learning and TV channels
Origin of scenario	Technical Annex - ICL
Typical users	Teachers, historians, curators, school children, and home users
Kind of benefit	Information dissemination, public interest and awareness of science, technology and/or history among wider public
Alignment with EU goals	Raising public awareness about some of the public issues like climate change, global warming, globalization, poverty, which are debated at EU level among the public is very important. Similarly, developing public interest in science and technology is also considered very important for school children.
Classification	Software as a service

3.7.2 Scenario information

Synopsis	A consortium of European museums and libraries agrees to work together to make their material electronically available for a variety of uses, both social and commercial. Each institution worked to digitally archive its material, to semantically annotate these items, and to develop innovative ways of accessing and presenting their material electronically over the Internet.
Strengths	There are wide social implications of such initiatives in the medium to long term. Providing access to school children and the wider public would be of great benefit.
Weaknesses	Due to dispersed nature of target audience it is a challenge to provide information at local level.
Opportunities	Some of the recent advances through which computational and storage capacity are widely accessible have made it possible to scale such an initiative to national and EU level. In addition, there is potential for users to be able to interact in various ways e.g. by uploading their own material and views on various issues.
Threats	Capturing social dimensions of these issues in an accessible manner. Getting to target audience when there is already a proliferation of various media, sites and channels already. Producing content to drive such an initiative is a major effort beyond the scope of this project.
Complexity (technical)	This scenario is of reasonable technical complexity. Major issues tend to be around producing good quality relevant

	content and policy issues.
Ambition	The scenario seems to be of relatively low ambition technically. However, efficient large-scale information dissemination (non-technical nature) is still a considerable challenge.
Actors / entities in scenario	Universities, libraries, museums, storage provider, hardware utility providers, end-users = general public.
Additional remarks	Software as a service
Schematic overview	



Long scenario description

A consortium of European museums and libraries agrees to work together to make their material electronically available for a variety of uses, both social and commercial. Each institution worked to digitally archive its material, to semantically annotate these items, and to develop innovative ways of accessing and presenting their material electronically over the Internet. Working with a university Internet centre, the museums develop a set of over-arching semantically driven services that allow semantically related material, potentially in differing modalities and from different subject areas and institutions, to be linked and explanatory narratives derived. For example, Newton’s work on optics is linked to the practical development of the microscope (with virtual use) and hence to the discovery of the previously invisible creatures and the impact this had on society and religion. Storage space for the electronic archives is provided by commercial storage utilities that also provide the required level of security and backup. Computation facilities for the analysis routines, some of which are computationally intensive, are similarly provided by commercial Utility computing centres. The service space developed by the institutions and the Internet centre provides a flexible and open means of accessing and relating the collective material. The museums and the Internet centre then work with a consortium of schools. Together they develop a set of access and presentation services that allow the available material and services to be used to support e-Learning and knowledge discovery processes in schools. The result is a set of services that allow schools themselves, both teachers and pupils, to develop and share learning and education resources. The museum's material is made freely available but digital rights protection is attached to each item to ensure its proper use. The museums also work with a science TV channel, D, to develop a series of interactive programs exploring various aspects of the same material, interleaving interactive explorations with streamed content and narration. These offerings utilise the semantic analysis and presentation services developed earlier to relate this material in interesting ways. These "programs" are placed on the Internet as use-on-demand, pay-per-use media services. The museums automatically receive revenue from each use.

3.7.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Asif Saleem (London e-Science Centre)
Date	6th November 2006
Remarks	-

3.8 *Financial services engineering*

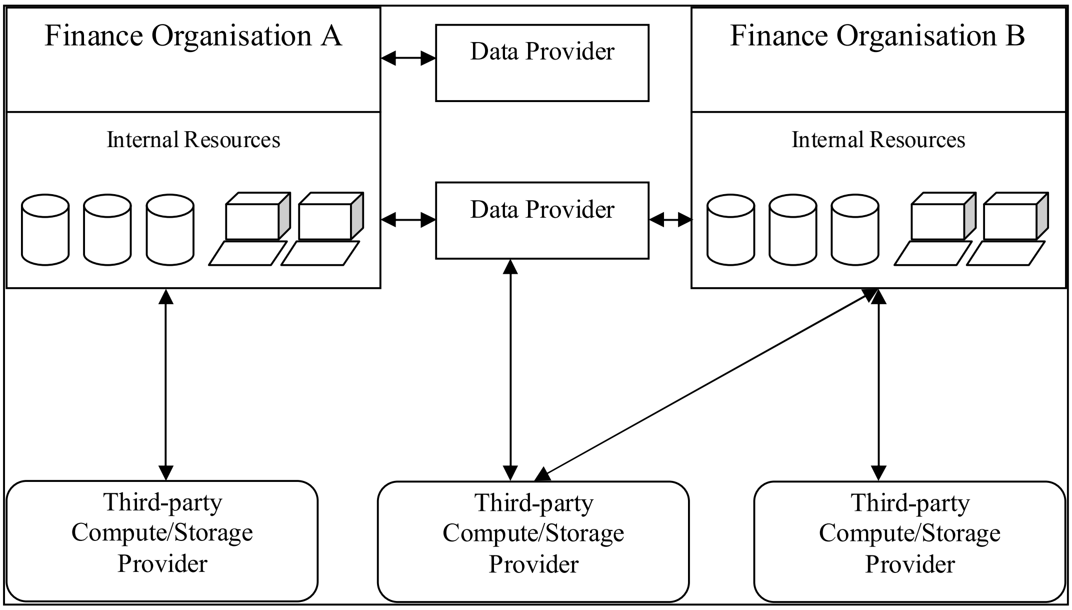
3.8.1 Scenario Meta data

Scenario name	Financial Services Engineering
Origin of scenario	Technical Annex
Typical users	Finance companies, particularly SMEs
Kind of benefit	Local and remote access to high powered computational resources enabling businesses to carry out more advanced modelling and simulation of financial scenarios.
Alignment with EU goals	
Classification	Interconnection of Grids

3.8.2 Scenario information

Synopsis	This scenario addresses access to high performance computational resources that are required by finance-related organizations. These organizations, including, banks, hedge funds, insurance companies etc, carry out large scale, computationally intensive financial modelling that consumes vast quantities of computing power. The more resources that are available, the more advanced the modelling that can be carried out. Since hosting the resources locally can be a very expensive and impractical proposition, this scenario also addresses remote access to High Performance Computing (HPC) resources.
Strengths	This scenario considers the requirements of financial businesses to carry out modelling and simulation using HPC resources. The more computing power that is available, the more complex the scenarios that can be modelled.
Weaknesses	Removing all local resources and relying on utility computing resource provision is impractical in an environment where reliability and security are critical.
Opportunities	The financial services market uses technology as a source of major competitive advantage and is aware of the need to experiment with latest technology,
Threats	Since some commercial vendors (DataSynapse, Gigaspaces, Tongosol) are already operating in this space. There is need to avoid overlap and target new areas previously unexplored. Due to the nature of the market it's difficult to collaborate with financial services organizations as they tend to be fairly secretive.
Complexity (technical)	This is of medium technical complexity and the market is huge. However, it provides a very good test case.
Ambition	It is of medium ambition and requires close collaboration with some financial organizations to tackle the challenge successfully.
Actors / entities in	End-users = Financial companies; hardware utility provider.

scenario	
Additional remarks	None.
Schematic overview	



Long scenario description

Utilisation of Grid-based computational resources accessed on a utility computing basis can allow Financial Services organizations to concentrate on their core skills without the need for the office space, administrative and technical staff, cooling and dedicated power feeds that are necessary for the provision of in-house high performance computing equipment. Additionally, the ability to dynamically scale the available resources to the business's requirements reduces cost since large quantities of resources are only paid for when they are required.

Retail banking

In Retail Banking, the most important aspects of interest are data mining of huge datasets, high performance and reliable messaging infrastructure to develop scalable systems able to handle millions of transactions per second. Risk management and reporting applications also tend to important drivers of technology in retails banks.

Insurance services

Insurance services providers need to constantly assess the risks they face based on constantly changing situations. Typically, this involves running mathematical models with lots of different parameters to study various scenarios.

Capital markets

In capital markets, various asset classes are traded in markets (exchanges, options and futures markets) around the world. These asset classes could consist of derivatives, equities, fixed income, commodities, and foreign exchange.

In capital markets, technology is a source of major competitive advantage. The capital markets sector has specific requirements for both computationally intensive and data intensive applications. These applications have huge processing demands to process large data sets and continuous data streams e.g. as required in securities trading applications. Storing and archiving data streams for regulatory and archival purposes produce large databases.

In other applications, like risk analysis and portfolio management, there is a need to calculate asset portfolio exposure to continuously changing market risk. Monte Carlo simulations and stochastic modelling techniques require use of server farms/clusters and are an ideal Grid computing application. High performance and reliable messaging are critical to ensure efficient operation in capital markets.

3.8.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Jeremy Cohen / Asif Saleem
Date	1st November
Remarks	-

3.9 *Exchange trades monitoring*

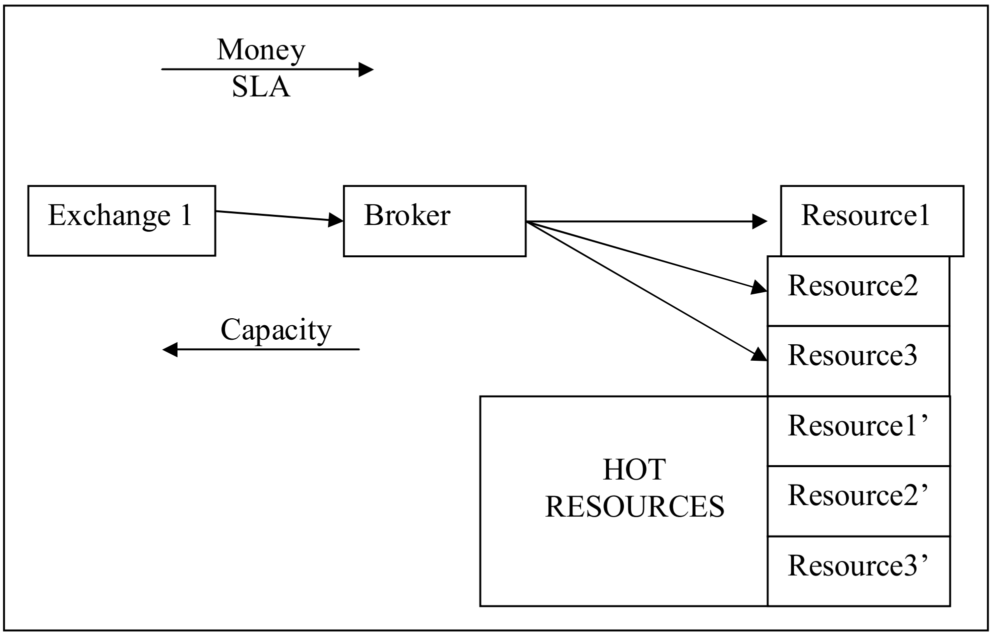
3.9.1 Scenario Meta data

Scenario name	Exchange Trades Monitoring
Origin of scenario	Technical Annex
Typical users	Stock Exchanges
Kind of benefit	Creation of a scaleable trades monitoring solution with ultra fast failover
Alignment with EU goals	Compliance
Classification	Service-oriented architecture

3.9.2 Scenario information

Synopsis	Many stock exchanges are experiencing unprecedented growth in trade volumes due to automated trading and increased usage of derivatives. It needs to scale the entire business, while keeping costs at check and avoiding vendor lock-in.
Strengths	Stock exchanges face growth in number of trades and increasing demand for compliance (e.g. money laundering, terrorism et al). Extremely high availability is required.
Weaknesses	Stock exchanges are highly advanced ICT users and are on the forefront of high availability / fast failover and reliable systems.
Opportunities	To offer a system that goes beyond current capabilities and reduces cost and increase transparency.
Threats	Stock exchanges are heavily regulated and favour proven technology and strongly penalized SLAs.
Complexity (technical)	Complexity is high especially given the demands in latency, availability, scalability, security, failover and no vendor lock in.
Ambition	Ambition level is high
Actors / Entities in Scenario	End users = exchange Resource provider = broker or certified 3rd parties Resource broker = broker
Additional Remarks	None.

Schematic Overview



Long scenario description

Many stock exchanges are experiencing unprecedented growth in trade volumes due to automated trading and increased usage of derivatives. It needs to scale the entire business, while keeping costs at check and avoiding vendor lock-in. Its technical challenge is to re-implement the monitoring application around Grid capabilities to be linearly scalable, high-performance with ultra-fast fail-over, while remaining standards-based (vendor neutral). A solution may involve the provision of a grid infrastructure that provides this application with a scaling solution for high-performance and reliability.

3.9.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Rob Blaauboer
Date	2-11-2006
Remarks	-

3.10 *Order management system (OMS)*

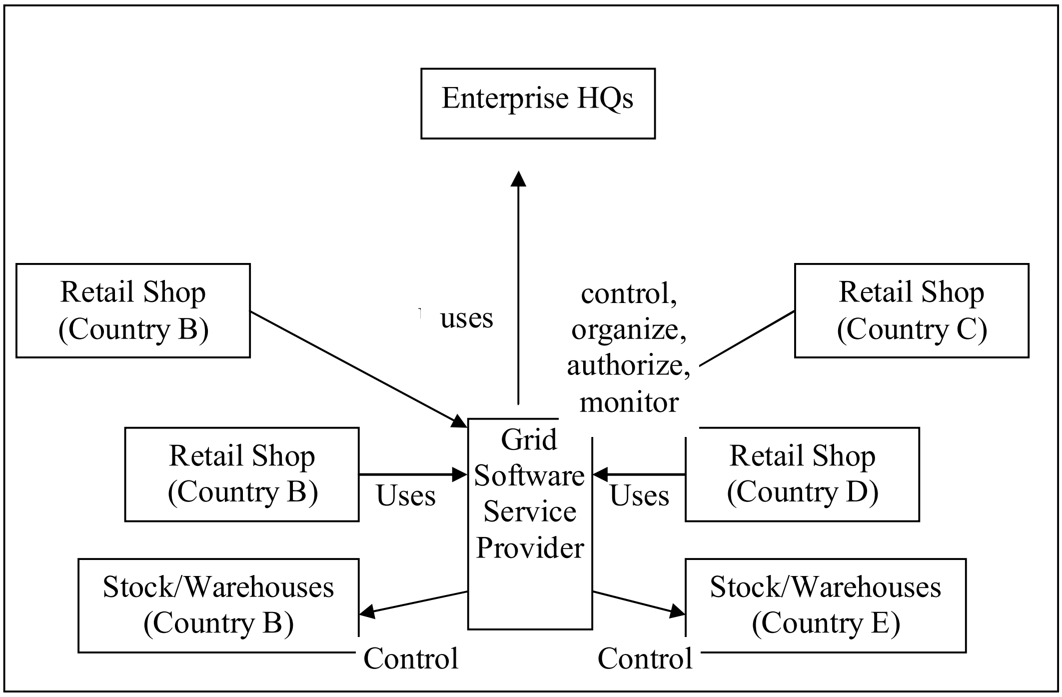
3.10.1 Scenario Meta data

Scenario name	Order Management System
Origin of scenario	Technical Annex
Typical users	Large (International) Enterprises
Kind of benefit	New Capabilities, originally not feasible for a non-IT enterprise
Alignment with EU goals	Empowering organisations to create and provide access to and use a variety of services, in a transparent and cost-effective way. Creating efficient components and lowering costs.
Classification	Software as a service

3.10.2 Scenario information

Synopsis	A large enterprise runs a large number of retail shops in many countries. It wants to control, organize and monitor the product orders and stock availability, to be able to process and supply thousands of retailers.
Strengths	Belongs to the inter-connecting grid problems, with examples on some very challenging economic questions regarding cost-effective hardware, dynamic scalability, QoS features on resource provisioning.
Weaknesses	More of a community grid rather than an open market. Not easy to design a sensible billing mechanism for the services and the authorities are rather centralized.
Opportunities	This can be a great advantage over competitive sales enterprises, since such a system can help in decreasing costs, increasing product availability and optimal scheduling of resource production.
Threats	None.
Complexity (technical)	Medium, related middleware has already been available and the challenge is in designing the economic components for optimizing transaction processing and resource usage.
Ambition	Medium. The scenario seems to be viable and realistic. The incentives of different entities in this organization are easy to understand.
Actors / entities in scenario	Enterprise HQs, retail shops, software service provider, hardware utility provider
Additional remarks	None.

Schematic overview



Long scenario description

A large enterprise runs a large number of retail shops in many countries. It wants to control, organize and monitor the product orders and stock availability, to be able to process and supply thousands of retailers. This can help in decreasing transportation costs and increasing total revenue. The demand of intensified number of transactions per day becomes critical: creating orders, fulfilling orders, defining sales and delivery routes, product availability etc. The technical challenge is to use a Grid for performance boost on cost-effective hardware that facilitates dynamic future growth. A possible solution may involve the provision of a Grid infrastructure for parallel processing engine as well as a distributed cache and a sophisticated billing and accounting mechanism to enable job completion in high performance.

3.10.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Stavros Routzounis
Date	2/11/06
Remarks	

3.11 *Consumer web services*

3.11.1 Scenario Meta data

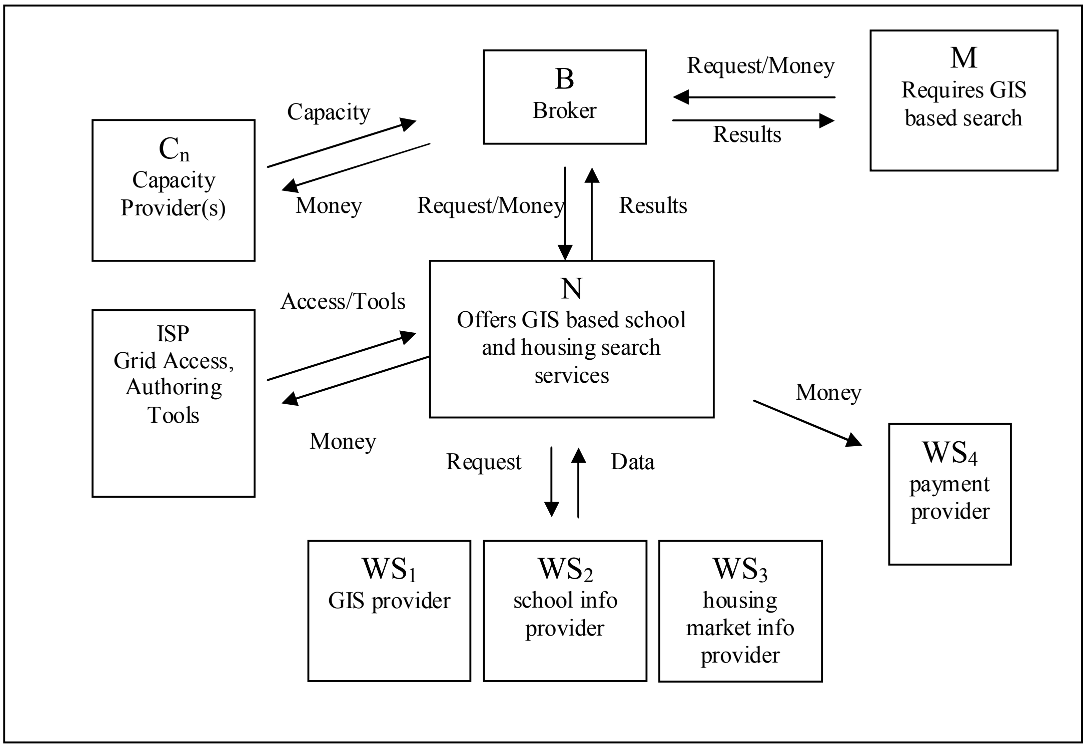
Scenario name	Consumer Web Services
Origin of scenario	Technical Annex
Typical users	Consumers, SME, ISV
Kind of benefit	New capabilities not originally available
Alignment with EU goals	Promotes consumer adoption of Grid technology based products, generates new business opportunities leveraging existing (non-computing) skills within the community
Classification	Software as a service

3.11.2 Scenario information

Synopsis	A citizen identifies a business opportunity to provide a service for fellow citizens and is able to implement this service by leveraging pre-existing Web services (GIS, payment and SLA), all through an easy to use authoring environment. The citizen sells access to the service through intermediaries who utilize utility computing to dynamically provision resources to meet demand.
Strengths	The developer can concentrate on the key commercial aspects of the application. The complexity of the underlying technology is hidden from both the developer and consumers. Grid technology ensures that the application dynamically scales to meet fluctuations in demand.
Weaknesses	Requires a level of maturity in web services – either open standards or dominant market leading services. Requires semantic Grid to fully achieve aims.
Opportunities	Easy to use authoring and deployment tools would open the Grid market place to many non/less technical EU citizens.
Threats	Domination by non-EU market leading companies.
Complexity (technical)	The complexity is considered to be medium since the process is broken down in identifiable pieces (Web service/Grid authoring and deployment tools, semantic standards for GIS, payment and SLA web services).
Ambition	The ambition level is considered to be high insofar as it depends on the existence of services, which are still under research/development and for which there are no clear standards.
Actors / entities in scenario	B – Broker – negotiation, SLA provision (could be a separate provider)

	C – Capacity providers – processing, storage, network ISP – Grid access, Grid authoring tools M – End-user N – Developer/application provider WS1 – Web service provider – GIS (Google) WS2 – Web service provider – school information (public) WS3 – Web service provider – housing market information WS4 – Web service provider – payment (Pay Pal/WorldCom)
Additional remarks	None.

Schematic overview



Long scenario description

The scenario outlined here demonstrates that the success of the Consumer Grid depends on easy and cheap access to a robust and well-defined Grid infrastructure. At a minimum this infrastructure must include capacity providers for processing, storage, and bandwidth as well as providers for brokering and payment. It does not necessarily require separate SLA providers as the application developer/provider could choose to provide his/her own simple SLA, for example a payment refund if not completely satisfied with the results.

In reality developers will often have a good business idea but will require help to convert the idea into something that can run on the Grid. For this to happen, the key elements required are:

- Widely accepted semantic descriptions for Web services – ideally standards lead (the EU could and should have a big input here), this would allow developers to discover Web services and choose between providers based on cost, performance, and features without having to develop to proprietary interface definitions.
- Web service discovery – even when Web services are semantically enabled, developers must still be able to find them on the Grid. The most efficient way for this to happen is to create a few widely known and trusted repositories to store the semantic descriptions and provide dynamic links to the physical Web services. Ideally a hierarchy of repositories would exist, organised by an international organisation and similar in nature to the DNS hierarchy.
- Web authoring tools which simplify the design and deployment of an application to the Grid. One can envisage extensions to existing/new easy to use Web authoring tools such as Microsoft Front Page or Macromedia Dreamweaver to allow these tools to discover and bind new applications to semantic enabled Web services for payment, brokering, GIS etc. In practice this requires Web service discovery via semantic service repositories.
- Provision of cheap and universal capacity. Adhoc pay-per-use consumer Grid applications require access to low cost and widely distributed processing and storage capacity. This is already happening with American companies such as Amazon, Sun, and Google taking the lead. However, it is not currently possible to for a developer to switch an application designed to use processing/storage from Amazon over to using Sun instead. Again success within the consumer Grid market will only come once developers are able to rely on standards based interfaces to these capacity providers.

3.11.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Derek McKeown
Date	30-Nov-2006
Remarks	Revised to include actor and long scenario

3.12 *Virtual company*

3.12.1 Scenario Meta data

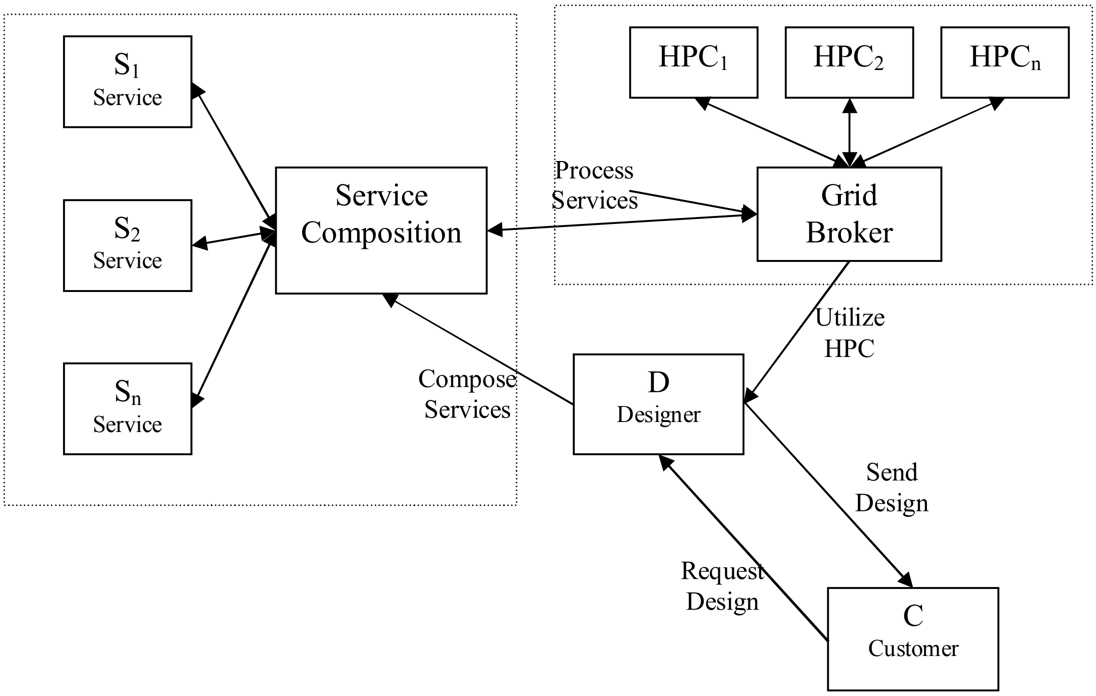
Scenario name	Virtual Company
Origin of scenario	Technical Annex
Typical users	Consumer
Kind of benefit	New business opportunities for end-users to create start-up companies
Alignment with EU goals	Improves the competitiveness of EU by allowing end-users to create start-up companies without the need of large investment. End-users are given the opportunity to capitalize on innovation and use existing services (in a pay-per-use model) in order to market their ideas to the general public.
Classification	Software as a service

3.12.2 Scenario information

Synopsis	A designer wants to start his own company, advising individuals how to furnish and design their house. While he is technically literate, he does not have neither the extensive technical knowledge nor the capital to develop the whole business chain required to provide his service. By utilizing existing services on the Web, he combines a range of modelling, 3-D visualisation, and rendering Web services in order to allow users to construct models of their house interiors, place furniture, and colour walls. Once he has set-up the appropriate service composition, he delegates the process intensive services to external HPCs in order to offer users services in almost real-time. He uses a pay-per-use model for all these transaction, allowing his business to grow along with customer demand. This model allows him to use external HPCs for widening his customer base without requiring a large investment in infrastructure.
Strengths	The scenario utilizes existing Web services and Grid platforms to minimize the capital required to create new businesses. The virtual company does not actually require any infrastructure as it acquires processing from external Grids based on a pay-per-use model.
Weaknesses	The number of end-users expected to create new businesses based on existing services is low. Complications of business conflicts when combining services is also expected to be a barrier to such a scenario.
Opportunities	This will allow new companies to be formed and revenue to be generated without the need for capital.

Threats	Legislation issues and IPR conflicts in combining existing services could prove a barrier to such a scenario.
Complexity (technical)	The complexity of the whole scenario is quite high, but from a Grid point of view, the components required are quite distinct and could be implemented. The high complexity resides mostly on the service composition side, which is out of the scope of the project.
Ambition	Quite high since a new business logic is introduced.
Actors / entities in scenario	Customer (C) = end-user of the new service Designer (D) = The small enterprise or entrepreneur that wants to provide new services based on existing ones Service (S) = Existing service Hyper-Computing Centre (HPC) = Publicly available HPC Grid Broker & Service Composition = Components made available through existing or future infrastructures, including additional economic components made available through GridEcon.
Additional Remarks	B2B scenario for small enterprises or entrepreneurs

Schematic Overview



Long scenario description

Designer, D, has aptitude for interior design. He would like to start his own company, advising individuals how to furnish and design their house. D is technically literate but not a computer specialist. He goes on the Web and discovers a range of modelling, 3-D visualisation, and rendering Web Services. Some of these services are commercial, available on a pay-per-use basis. The SLAs of these services specify the terms on which they can be used as components in other services. For using these services, a simple scripting language is provided as part of a Web Service authoring toolkit by an Internet broker. D constructs a service that interactively allows users to construct models of their house interiors, place furniture, and colour walls. Rendering and modelling services allow realistic 3-D walk-throughs of the prototype designs. The facilities provided by the Internet broker allow D to equip this service with the ability to negotiate a price for its use and, when used, to remit payment, through Pay Pal, to D's account.

D's service is launched on the Web and proves successful. For each use, a broker finds appropriate HPC resources on Utility-computing platforms to execute the computationally intensive parts of the service. For each run, appropriate payments are made automatically to D, to the component providers and to the execution service selected. These payments are not large, so use is encouraged. Over time, the volume of use ensures that significant revenues accrue to all players in the value chain. As usage increases, the broker utilises utility-computing platforms, which, in turn, incrementally acquire adequate resources so that no degradation of service is observed. At no time, does D directly purchase the component software or the execution services used. Indeed, over time, these services themselves are replaced by semantically equivalent but faster versions. These improvements are available to D transparently without any rebuilding of his application.

A house-building company, B, spots D's service and incorporates it as part of its service promoting its new houses, enabling prospective buyers to visualise their interiors, if desired. No contract or licence is agreed between B and D but every time the visualisation is used as part of B's service, appropriate payments are made to D, the component service providers and the execution provider. D's service proves popular. Over time, he develops and enhances the service.

3.12.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Dimitris Sotiriou
Date	1/11/06
Remarks	-

3.13 *Social network services*

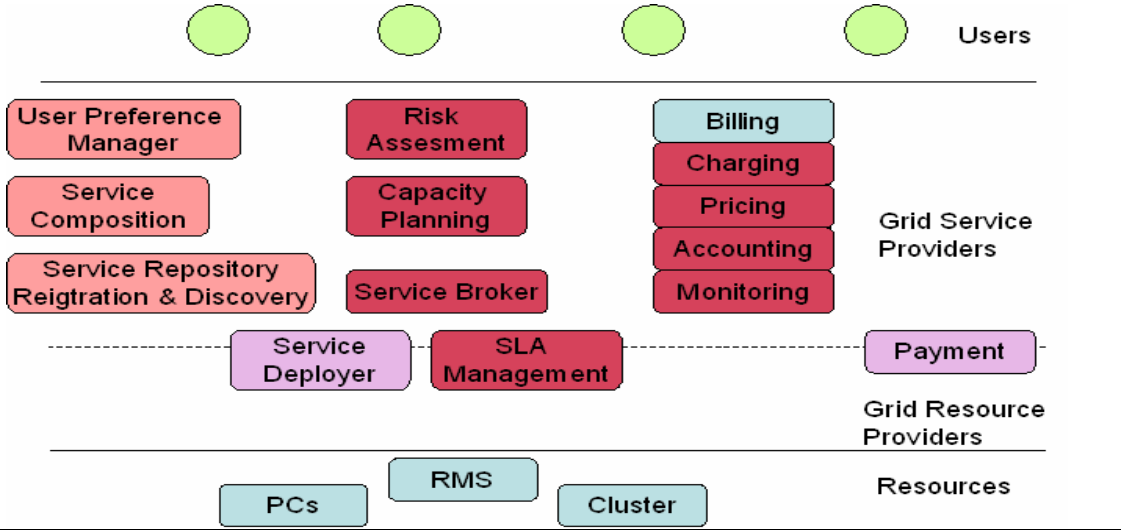
3.13.1 Scenario Meta data

Scenario name	Social Network Services
Origin of scenario	IU
Typical users	Internet end-users
Kind of benefit	Selling service to end user
Alignment with EU goals	Deploying/selling service on the Grid infrastructure. From the case of this work, many other services can also be built on the Grid.
Classification	Service-oriented architecture

3.13.2 Scenario information

Synopsis	The Grid service provider provides a social network services to end-users. He allocates some Grid resources from the Grid resource provider and deploys his service on the Grid resource. The end-users use service and pay the cost to the service provider. The service provider pays the cost for hiring resources on the Grid.
Strengths	This work can be seen as the connection between commercial Grid resources infrastructure and the end user. Basing on this work, many other new services can also be offered in the same manner.
Weaknesses	None
Opportunities	A very large number of end-users could use this work, almost all Internet users.
Threats	The problem happening inside the Grid resource provider can make the service unavailable.
Complexity (technical)	To realize this scenario, many issues, which are not considered in the present system, must be solved. Some of them are: – The communication protocol between service provider and resource provider to ensure the success of hiring Grid resource process and deploying the service. – The communication protocol between service provider and end user to ensure the success of using/paying the service. – A billing stacks system for accounting and charging.
Ambition	The ambition level is considered to be medium. The scenario seems to be viable and realistic (it does not involve any paradigm shift or discontinuities)
Actors / entities in scenario	end-users = consumers, Social network service provider Hardware utility service provider
Additional Remarks	None

Schematic overview



Long scenario description

Formally, a minimal social network is composed of a number of objects linked by relationships, the objects are also called actors or nodes and the relationships are called arcs or edges (Carrington et al., 2005). Many Internet Portals such as the Open Business Club (OpenBC, http://www.openbc.com) or LinkedIn (http://www.linkedin.com) are examples for applications that help users create a social network and hence are referred to as social network applications. OpenBC is widely in use in Europe while LinkedIn is more popular in the United States.

Now, a service provider wants to hire resources on the Grid to deploy his Social Network Services in order to save the cost. With the present working method of the Grid, the service provider will face following problems:

- The lacking of the communication protocol between service provider and resource provider to ensure the success of hiring Grid resource process and deploying the service.
- The lacking of the communication protocol between service provider and end user to ensure the success of using/paying the service.
- The lacking of a billing stack system for accounting and charging.

Besides that, many issues such as how to determine the risk of the hired resources or how to determine the capacity for future requirement still exist. All those problems should be solved to encourage providing business service on the Grid infrastructure.

3.13.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Dang Minh Quan
Date	2-11-2006
Remarks	-

3.14 *Squads*

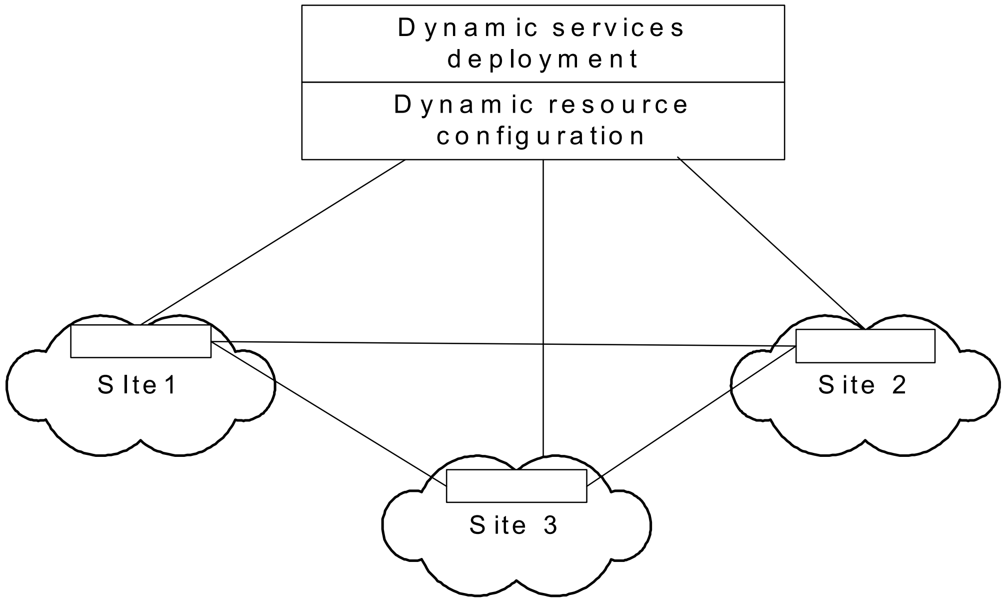
3.14.1 Scenario Meta data

Scenario name	Squads
Origin of scenario	Technical Annex
Typical users	People in organizations who need to use the service deployed on the Grid.
Kind of benefit	Selling Grid services to user
Alignment with EU goals	This work makes the Grid management/deployment easier. Through that, it promotes the deployment and usage of Grid as the infrastructure for the future.
Classification	Service-oriented architecture

3.14.2 Scenario information

Synopsis	Squads are mobile (dynamic) teams of people that maintain and develop inter-organizational infrastructures. This scenario assumes that Grids are already common and has a strong relationship with accounting and other economic issues since the business services involved use directly the Grids of these participating customers.
Strengths	This scenario support managing/using the Grid infrastructure in a dynamic way. A component of the Grid can come in and out of the Grid dynamically because of the changing in political, organisational and operational policies. Thus, the dynamic re-arrangement of the Grid structure both in architecture and services distribution can ensure the Grid to work properly and meet the requirement of the users
Weaknesses	Not all users need this service. The target of this service is a user within an organization, which contribute resource to the Grid.
Opportunities	This is a new way of maintaining the service.
Threats	Because of the dynamic changes in the structure of SQUAD organizations, the security can be a big problem, especially in a competitive business environment.
Complexity (technical)	To realize this scenario, many issues, which are not considered in the present system, must be solved. Some of them are: • The mechanism to dynamically configure the system coping with the change which can happen at any time • The mechanism to deploy the service on the infrastructure with many constraints of user requirement and complex set of incentives.
Ambition	The ambition level is considered to be medium. The scenario seems to be viable and realistic (it does not involve any

	paradigm shift or discontinuities)
Actors / entities in scenario	People in Squads, who maintain and develop inter-organizational infrastructures. Customer are SMEs as well as departments inside a big corporation. Organizations owning Grid resources, which provide resources for Squads.
Additional remarks	None
Schematic overview	



Long scenario description

Squads are mobile (dynamic) teams of people that maintain and develop inter-organizational infrastructures. This scenario assumes that Grids are already common and has a strong relationship with accounting and other economic issues since the business services involved use directly the Grids of these participating customers. It involves a variety of political, organisational and operational decisions. The components of the infrastructure might belong to various organisations and working on a component of one of the Grids may involve a complex set of incentives since not all participants may benefit by maintaining such a component.

The Squads support managing/using the Grid infrastructure in a dynamic way. A component of the Grid can come in and out of the Grid dynamically because of the changing in political, organisational and operational policies. Thus, the dynamic re-arrangement of the Grid structure both in architecture and services distribution can ensure the Grid can work properly and meet the requirement of the users. The working model of Squads can be presented in following figure.

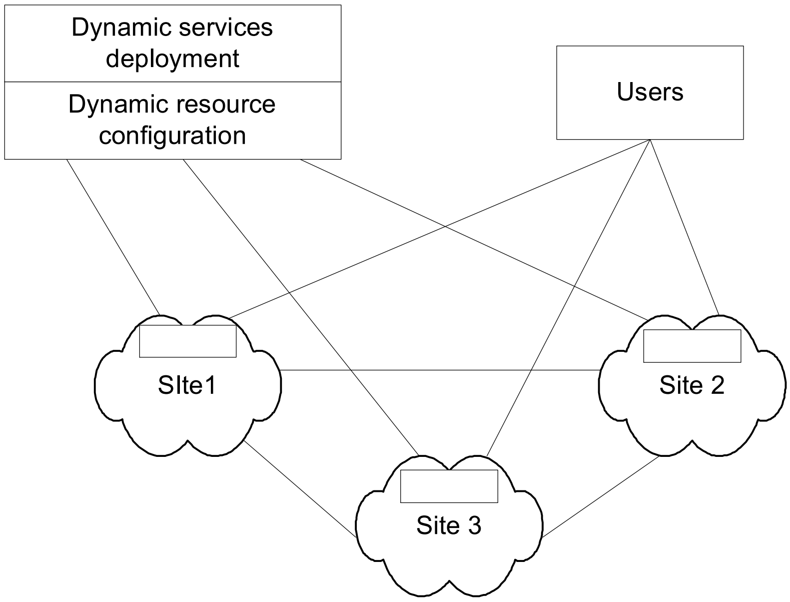


Figure 3.14.2-1 Model of Squads

To realize this scenario, many issues, which are not considered in the present system, must be solved. Some of them are:

- The mechanism to dynamically configure the system coping with the change, which can happen at any time. When one component releases the system, the system must be re-configured to ensure the integrity character. This is a complex task with various considerations inside the managing middleware.
- The mechanism to deploy the service on the infrastructure with many constraints of user requirement and complex set of incentives.

3.14.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Dang Minh Quan & Jörn Altmann
Date	5-11-2006
Remarks	-

3.15 *Insurance car repair*

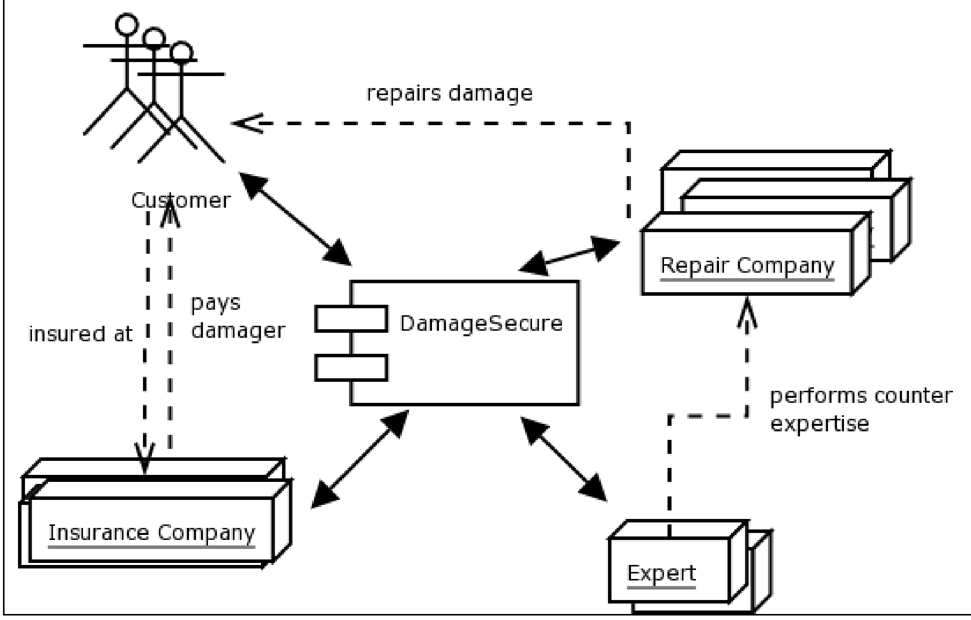
3.15.1 Scenario Meta data

Scenario name	Insurance Grid
Origin of scenario	ONTOGRID
Typical users	Players in the value chain of car repair
Kind of benefit	Reduction in the time to handle a repair claim
Alignment with EU goals	Cost reduction because of shorter time to handle dossier and reduction of cost because of negotiation between multiple possible repair shops
Classification	Software as a service

3.15.2 Scenario information

Synopsis	CarRepairGrid is build for an imaginary company called DamageSecure. DamageSecure looks after and controls all businesses involved in dealing with car damage claims for a number of insurance companies. .
Strengths	Car insurance is historically on the loss leading side. Cost reduction is always interesting for indemnity insurances to get out of the red.
Weaknesses	Model is flawed because calculations do not seem realistic. Grid is not needed to create a system like this. Car repair value chain is quite mature and less prone to act as case model.
Opportunities	Cost reduction in general is interesting.
Threats	Optimum system is already implemented; insurance companies will not buy into this.
Complexity (technical)	The system is not complex.
Ambition	Ambition level is low, since low complexity.
Actors / entities in scenario	End-users = insurance companies, repair companies Resource brokers = DamageSecure
Additional remarks	Grid does not add any value in this scenario.

Schematic overview



Long scenario description

CarRepairGrid is build for an imaginary company called DamageSecure. DamageSecure looks after and controls all businesses involved in dealing with car damage claims for a number of insurance companies. The goal of DamageSecure is to enhance the quality and efficiency of the total damage claims handling process between consumer, damage repair companies, and insurance companies. Every year around 100.000 damages are reported to DamageSecure, of which 40% are repairs and 60% replacements. If CarRepairGrid can work without human intervention it could potentially save 172M Euro.

3.15.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Rob Blaauboer
Date	2-11-2006
Remarks	-

3.16 *Grids in healthcare*

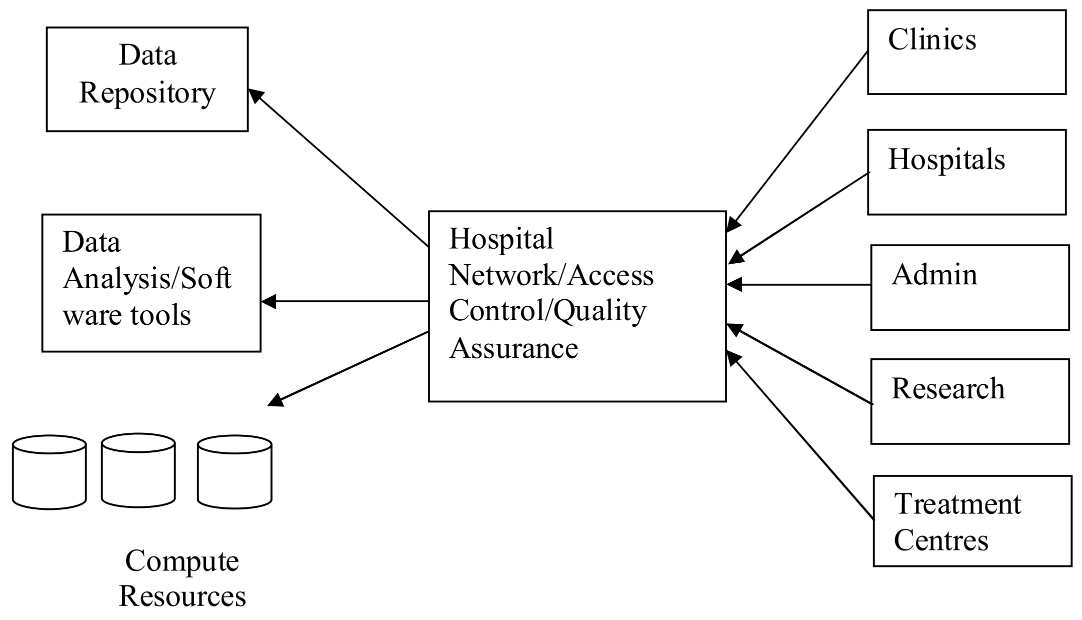
3.16.1 Scenario Meta data

Scenario name	Grids in Healthcare
Origin of scenario	Experience from existing London e-Science Centre projects at ICL
Typical users	Hospitals, clinics, healthcare networks
Kind of benefit	Local and remote access to high powered computational and data resources enabling hospitals to provide cutting-edge treatments and increasing the efficiency of hospital administrative tasks, including information sharing and monitoring.
Alignment with EU goals	Focus on improving the healthcare in Europe.
Classification	Interconnection of Grids

3.16.2 Scenario Information

Synopsis	This scenario describes the need for hospitals, clinics and other healthcare-related institutions to access compute and Data Grids and to enable them to provide cutting-edge treatments like image-guided neurosurgery etc to patients.
Strengths	This scenario addresses a very important application of Grid computing which will serve the wider community and improve healthcare.
Weaknesses	Issues such as patient confidentiality, data security and the sheer scale of many healthcare services make this a system that would be prone to regulatory as well as technical issues. This would be very expensive to implement.
Opportunities	Will assist the development of technologies that can be shared with other large-scale Grid deployments and projects.
Threats	The adoption of such a system may be hindered through regulatory and legal issues.
Complexity (technical)	Highly complex. From a technical view, this scenario requires very large, cross-organisational system deployment along with interfaces to advanced medical equipment.
Ambition	None.
Actors / entities in scenario	Hospitals, hardware utility provider.
Additional remarks	None.

Schematic overview



Long scenario description

This scenario describes the need of hospitals, clinics and other healthcare-related institutions to access compute and Data Grids and to enable them to provide cutting-edge treatments like image-guided neurosurgery etc to patients. Additionally, we aim to improve the efficiency of administrative tasks and information and data-sharing amongst doctors and medical staff. This includes the digitization of data like X-rays, MRI scans etc, which allows doctors to access vital information at the click of a button. Furthermore, complex analytical tools, e.g. image analysis, can be performed on patient data to monitor progress, detect anomalies, and improve diagnostic procedures.

3.16.3 Meta data about scenario record

Author	GridEcon technical annex
Filled in by	Ali Afzal / Jeremy Cohen
Date	9th November, 2006
Remarks	-

References

[1] Enterprise Grid Alliance (EGA), Reference Model Working Group, "Reference Model and Use Cases, Part 1 of 2," version 1.5, 10 March 2006.
[2] Global Grid Forum (GGF), "GGF Use Case Repository," http://forge.ggf.org/twiki/bin/view/UseCases .