SIMDAT

Data Grids for Process and Product Development using Numerical Simulation and Knowledge Discovery
Project no.: 511438

Grid-based Systems for solving complex problems – IST Call 2
Integrated project

SIMDAT

Deliverable

D2.1.3 Report on SIMDAT Integrated Grid Infrastructure and D2.1.4 Report on SIMDAT Integrated Grid Infrastructure Evaluation and Validation

Start date of project: 1 September 2004 Duration: 48 months

Due date of deliverable: 01/04/2006
Actual submission date: 09/05/2006

Lead contractor for this deliverable: IT Innovation Centre
Revision: 1.0

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination level		
PU	Public	X
PP	Restricted to other programme participant (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Copyright

Table of contents

Executive Summary

The document presents a description of the technology outcome for WP2 Integrated Grid Infrastructure for the Connectivity phase PM0-PM18 and future plans for the Interoperability phase PM19-PM30.

The Integrated Grid Infrastructure software has been successfully delivered and deployed in four SIMDAT application prototypes. The prototypes demonstrate how the Grid infrastructure can support inter-Enterprise collaboration for the design of complex products. The prototypes were successfully demonstrated at the 1st SIMDAT Review in November 2005. A summary of the prototypes is given below (detailed descriptions are available in the relevant deliverables: D17.1.3 [9], D11.1.5 [8], and D8.1.3 [7]). Evaluation feedback was captured from the Connectivity phase prototypes through a series of structured telephone interviews. Key partners from both technology and application work packages were contacted including MSC, BAE, LMS, ESI, EADS, UoS, ECMWF, Inforsense, Lion and FhG AIS. Although, WP2 did not deliver technology directly to the pharma and meteo activities leading partners from each sector where interviewed to establish initial experiences with Grid technology and future plans.

The evaluation feedback shows that the Integrated Grid Infrastructure software, based on GRIA 4.3, is well suited for SIMDAT's industrial application sectors where partners are providing value added service, at a cost, rather than direct access to computational resource. In addition, the security in-depth approach adopted by GRIA is largely compatible with the operational security requirements of industrial companies. The functional, usability and robustness enhancements were well received and essential for management and maintenance of Grid services. The majority of challenges experienced by users during the Connectivity phase were related to the Resource Allocation service.

A SIMDAT architecture has been defined that provides a vehicle for communicating SIMDAT ideas, structure, component technologies and their relationships to application communities within and beyond SIMDAT, whilst defining the scope of the technical developments to be undertaken by technology developers within the project. The architecture provides an abstract representation of Grid technology being developed by SIMDAT and partitions the overall Grid solution into a set of cleanly bounded subsystems that interact through well-defined interfaces based on important Grid and web service specifications. These specifications are being investigated during the Interoperability phase and will only be adopted after considering aspects of maturity and applicability to operational requirements of SIMDAT's industrial application sectors. WP2 (Integrated Grid Infrastructure) is responsible for the *Core, Security, Resource Management, Execution Management and Information* subsystems.

During the Interoperability phase, the Integrated Grid infrastructure software will be redesigned to provide a set of loosely coupled services based on the architectural subsystems. The services will allow application users to share assets in accordance with business objectives. Application users will be able to monitor and bill for usage of applications, information, clusters or storage depending upon their role within the value chain. The loose coupling of components will enable technology vendors easily integrate generic Grid capabilities into existing Problem Solving Environments. We will see a consolidation in the developments through conformance to the SIMDAT architecture and larger uptake of Grid technology across all application sectors transforming their ability to solve complex problems.

1 Introduction

1.1 Purpose

This document is the public version of deliverables D2.1.3 “Report on SIMDAT Integrated Grid Infrastructure” and D2.1.4 “Report on SIMDAT Integrated Grid Infrastructure evaluation and validation” of the EU IST-2002-511438 SIMDAT project.

The document presents a description of the technology outcome for WP2 Integrated Grid Infrastructure for the Connectivity phase PM0-PM18 and future plans for the Interoperability phase PM19-PM30. The technology delivered to the application sectors for the 1st prototypes is described and the contribution made to the state-of-the-art in respect to new developments analysed. The document presents the evaluation feedback captured from the application sectors following a series of structured interviews with key application and technology partners. The evaluation feedback has been consolidated into a set of specific new requirements identified to satisfy the needs of SIMDAT application prototypes. Finally, the document outlines the development plans for Integrated Grid Infrastructure to address these new requirements during the Interoperability phase

1.2 Scope

The objective of the first 12 months for Integrated Grid Infrastructure was to analyse the Grid infrastructure requirements of the four SIMDAT application activities in the context of the state-of-the-art of Grid technology, develop enhancements to existing infrastructure components that satisfy the application needs and to support deployment within application prototypes. The first integrated Grid infrastructure software (D2.1.2) was released and has been successfully deployed within four SIMDAT prototypes demonstrating how Grid technology can support the inter-Enterprise collaborative development of complex products. These prototypes were successfully demonstrated at the 1st SIMDAT review in November 2005. Between PM13 to PM18, the objective was to consolidate the first prototype developments allowing the application sectors to elaborate their application scenarios whilst capturing evaluation feedback regarding lessons learnt in the connectivity phase. The second Integrated Grid Infrastructure software was released (PM16) and the application prototypes successfully upgraded their prototypes. Sections 2, 3 and 4 provide a review the results from the Connectivity phase. Section 5 provides consolidated list of new requirements resulting from the evaluation feedback.

WP2 was instrumental in defining the overall architecture for the SIMDAT Grid solution portfolio (See Section 6.1). The portfolio provides a mechanism for SIMDAT to communicate project ideas and technologies to interested parties outside of SIMDAT and provides a framework for scoping the technical developments associated with the technology work packages within the project. WP2 is responsible for the core, security, resource management, execution management and information services.

1.3 References

1. SIMDAT Annex 1 Description of Work

2. D2.1.1 Consolidated requirements report, roadmap, and SIMDAT infrastructure design

3. Boniface M., Surridge M., "Industrial Grid Profile Analysis", IT Innovation SIMDAT Report, R01 V1.0

4. D2.2.2 First integrated Grid infrastructure

5. D14.1.2 Production of first working prototype

6. D 20.1.2 Software Design of SIMDAT Meteorological Scenario

7. D8.1.3 First Prototype of Knowledge Services Tools

8. D11.1.5 Description of Technology, Implementation and Evaluation of Automotive Prototypes

9. D17.1.3 Enhanced version of the aerospace scenario

10. I. Foster, C. Kesselman, S. Tuecke., The Anatomy of the Grid: Enabling Scalable Virtual Organizations, http://www.globus.org/research/papers/anatomy.pdf

2 Technology Improvements in the Connectivity Phase

The selection of Grid infrastructure components was made by application sectors following a detailed requirements capture phase and review of the state-of-the-art in Grid technologies provided by WP2 [2]. The selection was based on evaluating and selecting publicly available middleware components to ensure that the development of the Connectivity phase prototypes could begin quickly. The state-of-the-art presented various Grid infrastructure technologies some of which were available immediately (GRIA, OMII, UNICORE) and some with imminent release schedules (gLite, GT4). The emphasis during the analysis was to allow the application activities to come to their own conclusions and select Grid technology that meets the needs of their application domain rather than prescribing a single Grid infrastructure.

In the end, the application sectors selected a variety of technologies to support their first prototypes. The aerospace and automotive sectors chose GRIA due to the GRIA's explicit focus on industrial Grid deployments and immediate availability. The meteo sector concluded that Grid technology was still emerging and that the state-of-the-art in Feb-2005 would not meet their data grid requirements. However, OGSA-DAI and gLite had potential and would be re-evaluated for the Interoperability phase. The pharma sector decided to develop a web services Grid leveraging E2E security components developed during the GEMSS project.

The Integrated Grid Infrastructure software delivered by WP2 during the Connectivity phase was based on GRIA 4.3, providing various functional and usability enhancements to satisfy SIMDAT requirements. GRIA was initially developed in the EC IST GRIA project which ended Oct-2004. The final software delivered by the project was functionally good but robustness, usability and maintainability needed improvement to support commercial Grid deployment in SIMDAT. Also, additional functional enhancements were identified following an analysis of both the application and higher-level technology work package requirements.

GRIA 4.3 integrated OGSA-DAI WS-I 2.1 into a business process for secure and managed access to database resources. The GRIA resource model was extended to monitor and bill for usage of database resources. In addition, a security model was defined based on well-known database usage roles to support necessary access rights for a distributed design database controlled by a client actor but populated by workflow services operated by suppliers.

GRIA 4.3 also provided extensive usability enhancements to help users install and administer the software. Previous versions of GRIA typically took a few days to install and required in-depth knowledge about target platforms and the GRIA software itself. The new release makes the installation much simpler by utilizing Tomcat's web application archive (WAR) deployment facility. Users can now upload a GRIA WAR file directly to a Tomcat server and provide configuration information using the GRIA service provider administration portal. The administration portal is very intuitive and leads the user through all aspects of the installation. If problems exist with the configuration the user is notified directly. The administration portal also supports upgrade from previous GRIA installations whilst maintaining previous configuration options.

The application requirements showed that SIMDAT users need Grid infrastructure software to be portable between many operating systems and hardware platforms. GRIA is written in Java and is therefore inherently portable to most operating systems. However, providing deployment documentation and testing on every operating system/Linux distribution would be too costly. Through discussions with the application sectors, target operating systems were identified that would be supported for deployment in SIMDAT. GRIA was originally documented and tested for the SuSE Linux distribution. This has now been extended to support Fedora Core 3 and Windows XP. Fedora was selected due to its close relationship with the Linux Ret Hat distribution required by the automotive sector. Fedora is more upstream than Red Hat with more frequent release intervals and newer features where more inline with a Grid RTD project.

The application activities use various queuing systems to manage access to computational clusters. GRIA integrates queuing systems using Platform Connector scripts. In the aerospace and automotive sector Condor, PBS and SunGrid were already deployed. Additional Platform Connector scripts where developed to support both systems.

A major objective of the Connectivity phase in the aerospace and automotive sector was integration between Grid infrastructure and workflow technologies Taverna/Freefluo, Inforsense and Optimus. All of the technologies provide a composition tool that allows users to compose workflows from a set of available services/tasks. GRIA was extended to provide a basic service registry of applications deployed at a service provider. The registry allows service providers to publish application metadata and clients such as workflow technologies to discover application metadata for use during workflow composition.

Future releases of Grid infrastructure software will begin to address interoperability challenges between middleware components deployed in each sector. We will begin to see a move away from monolithic Grid infrastructures to a set of services that can work together in a coherent way to delivering the needs of the application activities. See Section 6.1 for a detailed description on these issues

3 SIMDAT Contribution to the State-of-the-Art

The state-of-the-art in Grid systems is characterized by web service-based infrastructures aiming to support inter-Enterprise collaboration scenarios alongside traditional academic and scientific applications. Grid technologies have moved from second-generation services based on low-level bespoke protocols (Globus Toolkit 2.X, Unicore) to web service based services using open standards (GT4, GRIA, OMII, gLite). Grid technologies today are typically compliant with basic web service standards (WS-I) providing interoperability at the core protocol and security layers but standards for higher-level Grid functionality remains a matter for debate. In 2004, The Globus Alliance with IBM and others launched a new collection of standards called the “Web Services Resource Framework” (WSRF), part of which (concerned with notification) was later decoupled to become “WS-N”. These proposals were made directly to OASIS (not GGF), built on existing and emerging Web Service standards, and are seen as a key step that allows convergence between Web Services and the Grid. WSRF certainly is more compatible with wider Web Services standards (and their likely future development), but remains somewhat controversial. This is partly because it retained many of the original OGSA “object-oriented” concepts, and partly because some of the Web Service standards it uses are not yet widely agreed or accepted.

WSRF on its own does not provide interoperability although limited integration between Globus and Unicore has been demonstrated. More complex scenarios required by dynamic federation of information and computational resources currently require all participants to use the same Grid infrastructure software. This is obviously not feasible unless collaboration is restricted to a closed community whose members are known a priori. Compliance with Grid standards and OGSA/WSRF is not sufficient to overcome this problem: OGSA addresses architectural issues through the development of standards profiles and OGSA WSRF Profile specifies how to integrate stateful resources and notification into a Web Service framework, but neither define modelling or operation of Grid resources. Thus the Grid is falling well short of its potential and the challenge of standardising the Grid programming model and associated management services is therefore still unfulfilled, but we do now understand much better what is needed.

The deluge of different, complex and sometimes competing specifications has led to various “profiling” initiatives. A “profile” aims to improve interoperability by identifying a group of related specifications that can be used together for a specific purpose and adding further constraints to how the specifications are implemented. This idea originated from a group of leading vendors in the web service community called WS-Interoperability (WS-I). For example, WS-I Basic Profile 1.0 specifies that only certain transport protocols should be used (even though WSDL can accommodate others), so that vendors don’t have to implement all possible protocols in their frameworks. Other profiles are now emerging from the Grid community including the OGSA WSRF profile and NextGRID profiles, however, both are work in progress and complaint systems do not exist today.

SIMDAT application sectors are driving the evolution of Grid technology through the specification of scenarios that require uniform and location-independent access to persistent and supported heterogeneous IT assets that are under different ownership or control. These scenarios present interesting challenges for the existing Grid specifications and require major advancements in Grid management services to support operational security and business objective requirements. WP2 will analysis of Grid-related interoperability specifications, with a view to defining an adoption policy for SIMDAT’s industrial Grid developments. The purpose is to understand how these specifications can be used in industrial, B2B scenarios where infrastructure that can support strict but flexible export policies is critical.

Most existing Grid infrastructures (gLite, GT4, CNGrid, etc) provide a persistent VO model based on the model described by Foster, Kesselman and Tuecke [10]. This traditional “resource sharing” VO involves trust between providers and the VO, allowing the VO to exercise control and to optimise for the community. The VO is becomes a persistent manifestation of the collaboration operating its own services. The characteristics of this VO model mean relationships take time to set-up and that it is only suitable for “big” collaborations. In SIMDAT industrial use cases, service providers are potentially in competition with no shared goals and participants cannot be expected to optimise for a specific community. Service providers don’t reveal the status of their resources because if they announced they had some spare capacity they would certainly find it harder to negotiate a good price. Typically, service providers won’t even tell you exactly how many CPU they have, let alone whether they are busy or idle, even if you are a customer.

The SIMDAT “VO” model doesn’t try to unite resource providers because we don’t think they would ever reveal this information to each other. Service providers continue to manage their own resources, but in accordance with service level agreements with consumers. The SLA provides a high-level abstraction for the resources provided, and the control consumers have over them.

However, SIMDAT can unite resource consumers once they have procured services, they can share them with each other.

In WP2, we are developing an SLA management service that incorporates resource allocation, usage monitoring and capacity management. The service aims to support service providers in the negotiation of service level agreements (SLAs) by deciding how to map business-oriented requirements to local resource management policies, checking and monitoring the status of commitments against available capacity and predicting future capacity to steer the negotiation strategy for new SLAs. The service will allow service providers to share assets in accordance with business objectives. Application users will be able to monitor and bill for usage of applications, information, clusters or storage depending upon their role within the value chain. This approach is novel compared other Grid technologies, which typically provide a lower level management infrastructure based on direct access to remote computation resources using O/S level accounts.

4 Connectivity Phase Application Outcome and Evaluation

The Integrated Grid Infrastructure software has been successfully delivered and deployed in four SIMDAT application prototypes. The prototypes demonstrate how the Grid infrastructure can support inter-Enterprise collaboration for the design of complex products. The prototypes were successfully demonstrated at the 1st SIMDAT Review in November 2005. A summary of the prototypes is given below (detailed descriptions are available in the relevant deliverables: D17.1.3, D11.1.5, and D8.1.3).

- Aerospace (BAE, EADS, CEDG): pan-European inter-enterprise multidisciplinary (optimisation, aeroacoustics, aerodynamics) collaborative configuration design of a low-noise, high-lift landing system (D17.1.3).
- Automotive (Renault/ESI/IDEStyle): Virtual organisation for collaborative car design supporting confidentiality constraints of components between organisations (D11.1.5).
- Automotive (LMS, MSC, AUDI): Analysis service provision (D11.1.5).
- Automotive Knowledge (Audi, SCAI, Inforsense): Distributed data mining using WEKA toolkit (D7.1.3).

The objective for the PM13 – PM18 period was to support the application activities in the elaboration of the connectivity phase prototypes, capture initial elaboration feedback to inform the architecture of the Integrated Grid Infrastructure for the Interoperability phase, contribute to the architectural vision that underpins the SIMDAT Grid solution portfolio. Evaluation feedback was captured from the Connectivity phase prototypes through a series of structured telephone interviews. Key partners from both technology and application work packages were contacted including MSC, BAE, LMS, ESI, EADS, UoS, ECMWF, Inforsense, Lion and FhG AIS. Although, WP2 did not deliver technology directly to the pharma and meteo activities leading partners from each sector where interviewed to establish initial experiences with Grid technology and future plans. The following sections describe each prototype and the evaluation feedback captured from each sector. The evaluation comments have been consolidated into a set of specific requirements specified in Section 5.

4.1 Aerospace

GRIA has been deployed as the Grid infrastructure to support the aerospace application scenario. GRIA's capability to support dynamic virtual organisations allows a project manager to create a multidisciplinary design team where distributed engineers working in different organisations can participate in the design and development of a low-noise/high-lift landing system. Figure 1 shows the overall architecture for the aerospace scenario. Each organisation within the aerospace virtual organisation operates as a GRIA service provider offering engineering services that form part of the overall aerospace application including an initial optimisation function (UoS), aerodynamics (BAE) and aero-acoustics (EADS).

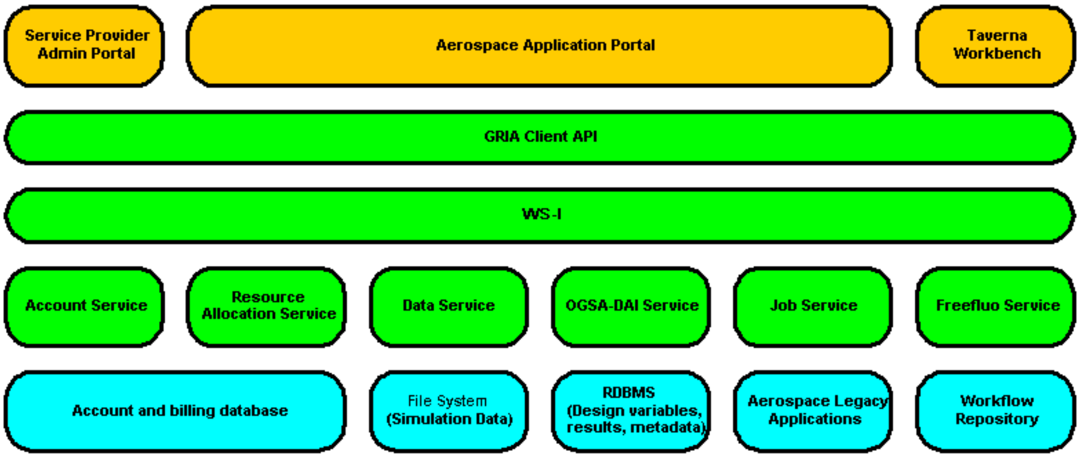

Figure 1 Aerospace architecture

Each service provider hosts six core services:

- Account service records resource usage for billing purposes
- Resource allocation service manages and assigns computation and data resources hosted by the service provider
- Data service stores simulation data files
- OGSA-DAI service provides access to database resources for storing metadata
- Job service executes legacy applications
- Freefluo service executes application workflows containing legacy jobs and post processing tasks for writing metadata to OGSA-DAI

Three user interfaces will be provided including:

- A service provider administration portal that allows service providers to manage resources such as accounts, data and applications
- Taverna workbench allows expert engineers to develop application workflows and deploy them to a Freefluo service
- Aerospace application portal provides a business-focused interface to the aerospace design workflow.

4.2 Automotive

The automotive sector has developed three prototypes each demonstrating progress towards different SIMDAT objectives. The first prototype from Audi is based on the simulation data and process management system (SDM system) MSC.SimManager. The prototype will focus on the implementation and intra-enterprise deployment of distributed database access, grid infrastructure and analysis services components. During the interoperability phase, the prototype will be extended to support virtual organisations addressing inter-enterprise issues such as firewalls, security, encryption, etc.

The second prototype from LMS/Noesis concentrates on the coupling between workflows and analysis services in the automotive scenario. The main purpose of the LMS SAMD is to develop a prototype demonstrating the current state-of-the-art in workflow definition and execution based on analysis services. GRIA was selected as the Grid infrastructure after examination of several Grid and web services frameworks. OPTIMUS has been integrated with GRIA demonstrating workflow composition and execution on the Grid. The prototype will be extended to compose and execute workflows over multiple service providers accessing a multi-CPU computer cluster.

The third prototype from RENAULT/IDESTYLE demonstrator uses Grid-enabled PSE environments to between automotive OEMs and partners. RENAULT involves numerous partners or subsidiary companies in the design process of vehicles, IDESTYLE in particular. The Grid-enabled PSE will enhance collaborative work between RENAULT and its partners by sharing analysis services and particularly by sharing or exchanging data in a more convenient and secure way. GRIA has been deployed to support this prototype specifically to provide distributed analysis services, secure data transfers and delegation to a trusted partner. The GRIA services have been integrated with Inforsense workflow technology to support the overall design process.

4.3 Meteorology

The meteo sector is developing a common infrastructure for the collection and sharing of distributed meteorological data: the Virtual Global Information Systems Centre (V-GISC). The prototype is based on the V-GISC architecture Data Communication Infrastructure (DCI) [6]. The aim of the prototype is to validate that V-GISC can be built on a distributed and loosely coupled Grid architecture. Grid technologies will be used to offer external interfaces to the V-GISC and to federate the partner’s data repositories. The main objectives of the demonstrator are the following:

- Build the virtual database foundations
- Implement the authentication service of the VO
- Implement part of the Access services

A combination of web service and Grid technologies such as OGSA-DAI has been used to integrate the legacy databases within a virtual database. The authentication service (AuthN) has been integrated to deliver single sign-on credentials to users. AuthN is based on the Public Key Infrastructure (PKI) allowing X509 user authentication. For the PM12 prototype a trivial authorization service (AuthZ) has been implemented that grants access to data for any authenticated user.

A portal providing a basic access to the virtual database functionalities has been built. This portal provides interactive access using web technologies and batch access using Grid technologies. This demonstrates that programs and humans can access the virtual database service. It also demonstrates that the Virtual database service can be integrated in a wider Grid.

4.4 Pharmaceuticals

The Pharma prototype coordinates pharmaceutical services, data and resources in order to support industrial and scientific pharmaceutical research. The prototype extends SRS software to provide client federation of multiple SRS services deployed across different domains. The system consists of three main components:

- Directory Service (Semantic Broker)
- Node Broker (SRS Nodes)
- Federated Portal (User Interface)

The Pharma prototype is deployed on an untrustworthy network, the internet, for communication of highly sensitive data and to give access to services and resources. The pharma prototype uses deployed a web services Grid with end-to-end security components from the GEMSS Grid developed by NEC. Pharma users require that access is granted only those who are authorized, that the data transmitted does not get eavesdropped, manipulated and is authentic. NEC's E2E security toolkit currently provides a security mechanism that can ensure a high level of security of the data transmitted from one Grid endpoint to another, even across intermediaries. In future, the E2E toolkit will be extended to provide dynamic access control mechanism that supports a variety of access models typical of Pharma collaboration scenarios.

4.5 Evaluation Summary

The Integrated Grid Infrastructure software has been deployed in four SIMDAT prototypes and successfully demonstrated at the 1st SIMDAT review. The early release of the first software infrastructure software enabled higher technology layers and application activities to begin work prototype developments early in the Connectivity phase. The evaluation feedback shows that the Integrated Grid Infrastructure software, based on GRIA 4.3, is well suited for SIMDAT's industrial application sectors where partners are providing value added service, at a cost, rather than direct access to computational resource. In addition, the security in-depth approach adopted by GRIA is largely compatible with the operational security requirements of industrial companies. The functional, usability and robustness enhancements were well received and essential for management and maintenance of Grid services.

The majority of challenges experienced by users during the Connectivity phase were related to the Resource Allocation service. GRIA 4.3 provides a simple QoS model based on the total usage across all jobs and data transfers over a specified interval. This type of model has been used with a so-called Resource Allocation service, separate from the job execution and data storage services, allowing user actions to be managed in three ways:

- Service providers respond to consumer QoS requests with offers based on predicted resource availability, allowing the user to detect which providers can meet their QoS needs, and accept offers only where capacity is sufficient for their needs.
- Users have to ask the QoS service to set up a job or data store, specifying the resource requirements, allowing the QoS service to check if these are within the agreed total before giving the user access to the job and data services themselves.
- The job and data services keep track of the resources used per job and data store, and terminate them when the declared usage limit is reached.

The evaluation feedback has shown that GRIA 4.3 has some weaknesses that can be traced to the specific QoS model and management mechanisms:

- Users find it hard to predict the QoS requirements for each job and data store. They often overstate the true requirements to avoid premature termination by the service, but this in turn means they have to overstate their needs when asking for QoS offers.

- Service providers have to assume that QoS requests are realistic, and often decide that they cannot meet (user-inflated) QoS requests, so that no services are provided.

- Users often want to reuse data for new calculations in ways that could not easily be predicted in advance. Thus they frequently find they want to use resources beyond the time period originally envisaged, but are prevented by the QoS model from doing so.

These drawbacks sometimes cause service providers to refuse user requests, even when the resources available are enough to meet the user's true requirements, which in a commercial situation would produce lower resource utilisation and higher prices. The refusals often come when users don't expect it, making GRIA services less usable to consumers. GRIA 4.3 has been very successful, and the QoS model has proven reasonably well matched to end user requirements, but the model is imperfect, and the compromises made in its implementation are unacceptable in the longer term.

5 Specific Requirements

This section provides a list of "new" requirements captured from the evaluation feedback following the SIMDAT connectivity phase prototypes. The requirements are organised according to the functional decomposition of SIMDAT infrastructure components defined by the SIMDAT architecture (See Section 6.1). These requirements extend those defined in during the initial requirements catyD2.1.1

Please note that these requirements are considered to be a "wish list" and that we cannot guarantee that these requirements will be satisfied within the SIMDAT project. The requirements will be prioritised through discussion with other technical and application work packages taking into account relevance to project objectives and resources available.

5.1 Requirements Classifications

To help to identify the importance of the requirements the classification criteria has been identified.

5.1.1 Priority

Allows the user to place some meaningful priority on the specific requirement

1) Should be delivered in the next phase

2) Should be delivered within the lifetime of SIMDAT

3) Would be required for production deployment but not critical for development within the lifetime of SIMDAT

4) Interesting new feature but beyond the scope of SIMDAT

5.1.2 Estimated Complexity

Allows the developers to place some estimate of how much effort the specific requirement will take to implement.

1) Significant development aligned with SIMDAT's technical objectives

2) Major enhancement to existing software

3) Minor enhancement to existing software

5.2 Requirements "Wish List"

The following sections provide a list of the specific requirements captured during evaluation feedback from the connectivity phase

5.2.1 Web Service Core

ID	Requirement	Source	User Priority	Estd Comp
WSC1	*Standard mechanism to access stateful resources using web services (WSRF)* Resource state should be accessible from the service provider rather than managed on the client.	MSC, Fhg AIS, Neosis	1	1

Table 1: Web service core requirements

5.2.2 Security

ID	Requirement	Source	Priority	Estd Comp
S1	*Compliance with organisation security policies and standards such as ISO/IEC 17799/27001* Many organisations have operational security policies that identifies levels of risk associated with types of data exchange and measures required by both infrastructure systems and operational working practices to mitigate those risks	All	3	1
S2	*Support trust relationships that are grounded in out-of-band policies rather than the ability to pay for a service* The trust relationship in GRIA is grounded in a limit of liability that a service provided is willing to risk against a specific customer. In many circumstances, the trust relationship is grounded in an out-of-band agreement such as the WMO-40	Meteo	1	2
S3	*Support auditable logging*	BAE, Renault	3	2/3

ID	Requirement	Source	Priority	Estd Comp
S4	*Allow client organisations to locally manage access control lists for remote resources provided by other organisations* Most Grids today require access control lists to be maintained by service providers resulting in a significant management overhead for both the client and service provider. Standards such as WS-Trust/WS-Federation provide a framework whereby a client can management locally access to remote resources within the scope of a specific trust relationship.	Meteo	1	1
S5	*Decentralised relationship management* In the meteo activity there is not centralised authority that can operate VO services; each party needs to be seen as equal within the VGISC. This means that concepts such as GT4's Community Authorisation Service (CAS) and gLite's Virtual Organisation Management Service (VOMS) are not well suited for supporting meteo's relationship management requirements. Infrastructure based on WS-Trust/WS-Federation is better suited for this scenario.	Meteo	1	1
S6	*Assign access rights to services (per application suite)* Today, anyone that has an approved account can run any of the services installed on the GRIA server. In some secured collaboration scenarios, some services are reserved for a certain type of users only.	ESI	1	

Table 2: Security requirements

5.2.3 Resource Management

ID	Requirement	Source	Priority	Estd Comp
R1	*Allow QoS commitments with softer commitments on resources and time periods* Users find it hard to predict the QoS requirements, users often overstate the true requirements to avoid premature termination, service providers assume QoS requests and often no service is provided. Users often want to reuse data in ways that could not easily be predicted in advance, users find they want to use resources beyond the time period originally envisaged	ESI, MSC, BAE, EADS	1	1

ID	Requirement	Source	Priority	Estd Comp
R2	*Support a very generic QoS model handling a wide class of application services* The resource model is too specific to job execution and data services and its difficult to extend to other types of services such as OGSA-DAI and workflow	MSC	1	1
R3	*Support a resource reservation model* Users want to be able to reserve a machine for a specific period of time in the future. The current resource model is based on a best effort strategy	ESI		2
R4	*Support management of application and data resources* The value to a service provider may be in the application or data rather than the physical compute and storage capacity. Users want to be able to record and possible bill for usage of a wider range of resources	ESI	2	1
R5	*Support Grid license distribution* Some ISV software cannot be accessed as a service by external organisations because of running an application in this way contravenes the conditions of most licenses. ISV's do not currently support these type of license distribution within their business models	ESI, EADS, UoS, BAE	2	1
R6	*Support for long term relationship management* The meteo partners exchange data based on a policy defined by the WMO-40 resolution. The resolution states that certain data must made available freely to participating organisations	Meteo	2	1

Table 3: Resource management requirements

5.2.4 Information Services

ID	Requirement	Source	Priority	Estd Comp
IS1	*Support analysis services discovery* Consumers need to be able to discover analysis services as part of the publish, discover, bind paradigm for loosely coupled SOA architecture. The metadata descriptions of GRIA 4.3.0 applications is too basic and needs to be extended to include to include addition metadata and the possibility of domain knowledge For example, it should be possible to put constraints on the input and output files such	MSC, Inforsense, BAE, EADS, NEOSIS, FhG AIS	1	1

	as the requirement for 1 .dat file, n .inc files and optionally n config files.			
IS2	*Support information discovery* Consumers need to be able to discover data that is published on the Grid. This is especially relevant to data mining, meteorology and bioinformatics data.	FhG AIS, ULB, Meteo	1	1
IS3	*Support for building registry services for data and analysis service in the automotive SOA -Architecture* The data and analysis services which will be built up out of PSE's need a reliable infrastructure for registering their services. The gathered information about input, output and location of services with possible semantic annotations should be published to the registry service and managed there to support the client to find the appropriate service.	MSC	1	1

Table 4: Information services requirements

5.2.5 Execution Management

ID	Requirement	Source	Priority	Estd Comp
EM1	*Support interactive real-time visualise of engineering simulations* Engineers want to be able to steer design simulations through interactive graphical interfaces.	ESI	2	4
EM2	*Support integration of multiple backend queuing systems with a single analysis service* GRIA 4.3.0 interface to queuing systems is currently designed for a single backend system. Integration of multiple queuing systems is possible through the development of a special Platform Connector script but a more elegant solution would be desirable	ESI, UoS		2
EM3	*Allow users or service providers to pause and resume jobs* GRIA 4.3.0 only supports a direct kill of applications running within the job service. In some situations, it may be desirable to pause and restart and executing job for the perspective of a user analysing the progress or a service provider who is trying to comply with a set of SLA's. The ability to support this pause/resume depends upon the functionality provided by applications.	ESI	3	2
EM4	*Provide a simple way of giving parameters to an analysis*	NEOSIS		2

ID	Requirement	Source	Priority	Estd Comp
	service GRIA 4.3.0 job service supports input from a data stager or a command line argument. It should be possible to give typed parameters (ints, double, string) to an Analysis Service that can be queried using service registry.			
E5	*Support a standard analysis service interface* Users want to have a standard web service interface for job submission allowing applications such as PSE's and workflow tools to be developed without maintaining complex adaptor code for each Grid middleware	MSC, Inforsense, EADS, NEOSIS	1	1

Table 5: Execution management requirements

5.2.6 Data

ID	Requirement	Source	User Priority	Estd Comp
D1	*Support data transfer between data stagers* In GRIA 4.3.0, it is not possible to directly transfer data between data stagers. This can only be achieved through the client federation or a specific data transfer job service	Inforsense FhG AIS	1	1
D2	*Optionally compress input and output files* Possibility to compress input/output files from within GRIA (make this an option for every file).	NEOSIS	3	3

Table 6: Specific GRIA data requirements

5.2.7 Client Interfaces

ID	Requirement	Source	User Priority	Estd Comp
CI1	*Provide a generic Job submission portal* Users want a job submission portal for use when deploying and testing installations and new applications. This is generally seen as a better option than a CLI or GUI because there is no need for a client footprint. Normally, engineers and scientists would use existing PSE or workflow tools.	BAE, ESI, UoS	3	2
CI2	*Refactor GRIA 4.3.0 Java API*	NEOSIS	3	3

ID	Requirement	Source	User Priority	Estd Comp
	ApplicationOutputType and ApplicationInputType are completely different types for the moment, however they share a significant amount of functionality, it would be possible for these types to implement a mutual interface Provide an easier way to create objects in plain java from the classes ResourceAllocationType, JobSpecType. Now this can be done from an XML file, but to do this in any other way is rather difficult.			

Table 7: Client interface requirements

5.2.8 Usability

ID	Requirement	Source	User Priority	Estd Comp
U1	*Improve GRIA 4.3.0 documentation* The documentation could be improved by adding more detailed examples to the Java client API.	ALL	1	3
U2	*Provide a technical FAQ* Detailed GRIA 4.3.0 configuration options were sometimes required from GRIA support. FAQs would help especially related to common problems associated with Firewalls and Apache	ESI	1	3
U3	*Provide additional metadata about stateful resource ids* Resource ids are currently URIs, which are difficult for a user to remember and understand	NEOSIS	2	3
U4	*Improve exception handling and error messages* Errors resulting from resource allocation and job failures were especially difficult to debug and often required help from GRIA support. Providing errors that are actionable by the user would be a desirable	ALL	2	2
U5	*Improvement availability of service usage information* Users want more information about account and resource usage	ESI, NEOSIS	2	3
U6	*Support other databases that PostgreSQL* Database independence is important for organisations who have a preferred choice of either supported/or favourite	EADS, NEOSIS	1	2

ID	Requirement	Source	User Priority	Estd Comp
	database			
U7	*Refactor GRIA 4.3.0 services so that they can be used independently* GRIA 4.3.0 services are tightly coupled and cannot be used independently very easily. Users want to use functionality of some components such as security, account, SLA without having to deploy the entire GRIA release.	MSC, ECMWF, GSK, NEC	1	1
U8	*Use a web-based installer for the keystore generation* Users currently have to use the KeyToolGUI to generate a keystore. It would be better if the keystore was generated within the web browser rather than an external program.	NEOSIS	3	3

Table 8: Usability requirements

6 Integrated Grid Infrastructure Developments

6.1 SIMDAT Grid Solution Portfolio

SIMDAT is developing generic-enabling Grid technology aimed at enhancing applications in key industrial sectors. The technical developments and outcomes are structured according to an overall SIMDAT architecture defined by the technology champions during the Connectivity Phase (See Figure 2). The architecture provides a vehicle for communicating SIMDAT ideas, structure, component technologies and their relationships to application communities within and beyond SIMDAT, whilst defining the scope of the technical developments to be undertaken by technology developers within the project. The architecture provides an abstract representation of Grid technology being developed by SIMDAT and partitions the overall Grid solution into a set of cleanly bounded subsystems that interact through well-defined interfaces based on important Grid and web service specifications. These specifications are being investigated during the Interoperability phase and will only be adopted after considering aspects of maturity and applicability to operational requirements of SIMDAT's industrial application sectors.

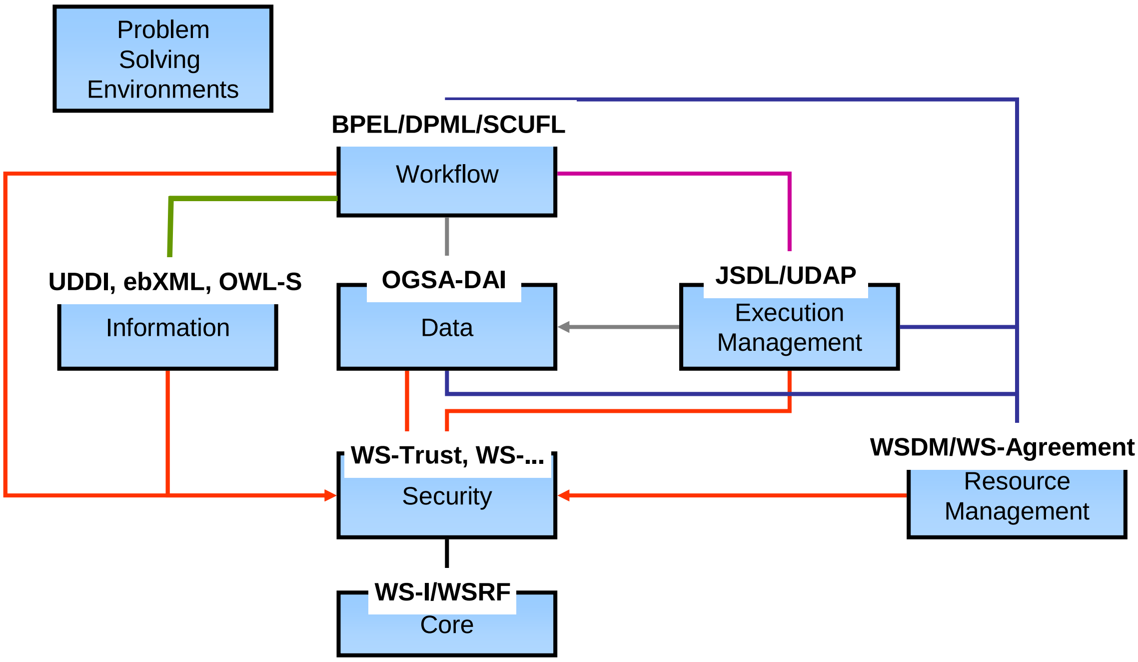

Figure 2: SIMDAT Architecture

Each subsystem in the architecture has specific responsibilities within the overall solution and dependencies to other subsystems indicated by directional arrows in Figure 2. The *Core* subsystem underlies all Grid service developments and provides web service containers that support contextualised (stateful) web services compliant with interoperability specifications WS-I and WSRF. The *Security* subsystem provides capabilities based on industrial web service specifications (WS-Security, WS-Trust, WS-SecureConversation, WS-Policy and WS-Federation) to maintain security policies, enable access rights and enforcing restrictions for other actors consistent with the terms of industrial security policies for protection of IPR. The *Resource Management* subsystem is responsible for supporting service providers in deciding how to translate business-oriented requests for service to a local resource management policy using service level agreements. The *Data* subsystem provides data transfer, storage, federation as well as semantic mediation between different data repositories. The *Execution Management* subsystem is responsible for managing the lifecycle of data processing tasks including staging data transferred from the *Data* subsystem and reporting usage metrics to the *Resource Management* subsystem. The *Information* subsystem supports the publication and discovery of application services providing specific data processing, storage, and transfer capabilities. The *Workflow* subsystem provides a programming environment that allows users to develop application scientific processes and organisational business processes. *Workflow* is decomposed into composition, enactment and dynamic semantic workflow support based on semantic service discovery and workflow mining. The *Problem Solving Environment* subsystem consists of various tools targeted at application sectors that are enhanced through integrated with subsystems from the SIMDAT Grid Solution Portfolio.

The architectural subsystems do not map directly onto SIMDAT workpackages, some workpackages are responsible for multiple subsystems and some are shared between workpackages. WP2 (Integrated Grid Infrastructure) is responsible for the *Core*, *Security*, *Resource Management*, *Execution Management and Information* subsystems. WP2 works closely with WP4 (Administration of Virtual Organisations) to understand the developments required for the *Security* subsystem to satisfy the needs of industrial security policies. WP2 also works closely with WP7 (Analysis Services) and WP8 (Knowledge Services) to understand the requirements for discovery and execution of services, as part of the *Information* and *Execution Management* subsystems. WP3 (Data Access and Integration) is responsible for leading the *Data* subsystem and works with WP2 to understand integration with *Execution Management*, WP6 (Ontologies) for semantic mediation/integration and WP5 (Workflow) to understand data federation requirements. WP5 is responsible for the *Workflow* subsystem. All technical workpackages consult closely with application workpackages through the technology champion role to understand requirements and to identify application specific developments that can be considered generic and part of SIMDAT's Grid Solution.

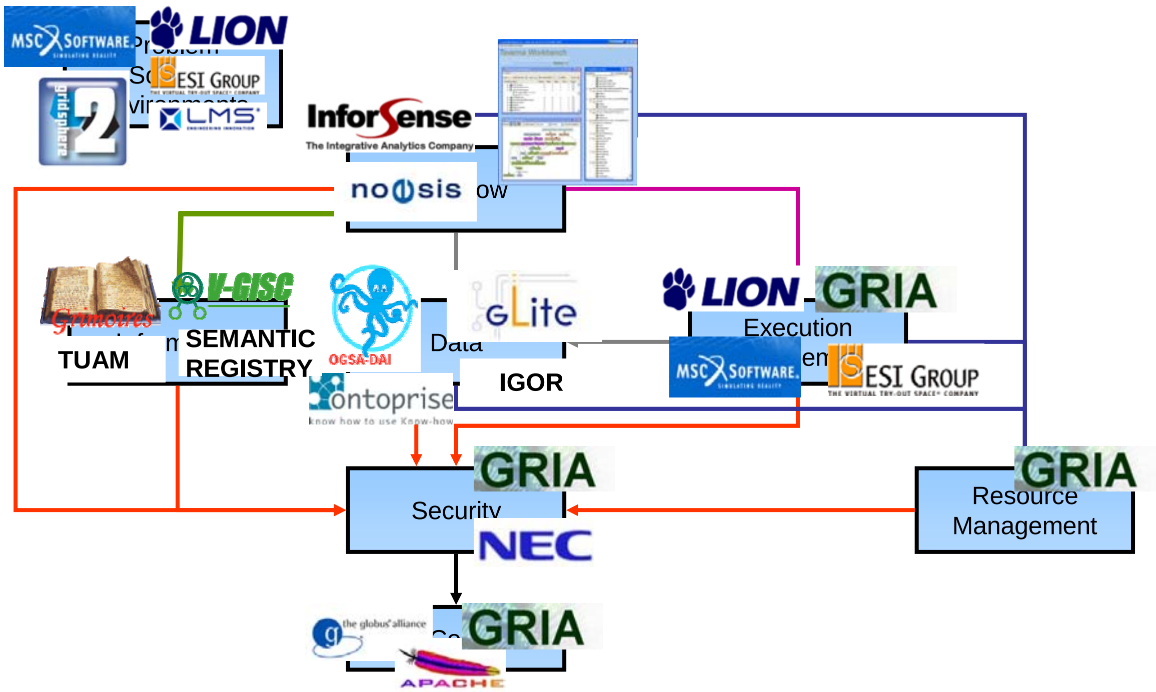

Figure 3: SIMDAT Technologies

Figure 3 shows the specific technologies that are being deployed within SIMDAT application sectors along with the relationship between the technologies and the SIMDAT architecture. The objective is to development a set of loosely coupled Grid components that can used together to satisfy the needs of individual application sectors. For each component in the architecture, SIMDAT offers various technology alternatives that application sectors select from when building a Grid solution in a specific domain.

6.2 *Subsystem Developments*

The overall goals of this Integrated Grid Infrastructure workpackage during the Interoperability phase is to further develop generic Grid infrastructure services by improving interoperability through architectural analysis and adoption of relevant standards, integrating additional services by exploiting ongoing developments both in SIMDAT and the Grid community at large, and to respond to specific needs of the application sectors as their scenarios mature. The work includes an update to the state-of-the-art survey and gap analysis of existing technology offerings that can be used to inform the generation of a second "roadmap" for supporting emerging requirements throughout the project. The following sections describe in detail the tasks for Integrated Grid Infrastructure related to each subsystem within the Grid solution portfolio.

6.2.1 Web Service Core

The objective for the web service core subsystem is to provide an analysis of important Grid-related interoperability specifications, with a view to defining an adoption policy for SIMDAT's industrial Grid developments. The purpose is to understand how these specifications can be used in industrial, B2B scenarios where infrastructure that can support strict but flexible export policies is critical. A "profile" aims to improve interoperability by identifying a group of related specifications that can be used together for a specific purpose and adding further constraints to how the specifications are implemented.

6.2.2 Security

The objective for the security subsystem is to provide a set of security components based on web service standards that can be used to enforce the policy requirements of SIMDAT's virtual organisations. The security components will be based on emerging industrial specifications (WS-Security, WS-Policy, WS-SecureConversation, WS-Trust and WS-Federation) allowing SIMDAT to be aligned with enterprise web service technologies from IBM and Microsoft. The specifications and their use within industry is currently being analysed by the EU IST NextGRID project. SIMDAT will build on the results of NextGRID experiments through concertation to realise the security subsystem.

During the Connectivity phase, GRIA (IT Innovation) and E2E Toolkit (NEC) both provided web service security solutions to the application sectors [4, 5]. Both GRIA and E2E implement complimentary aspects of the web service security stack that when integrated will provide a powerful security solution for SIMDAT applications. GRIA, through collaboration with NextGRID, will be enhanced to provide dynamic trust and access control through extensions of PBAC and accounting to improve the manageability of federations within the Grid, by permitting client organisations to manage authorisations locally rather than distributed access control lists at each service provider (WS-Trust and WS-Federation). E2E toolkit will add mutual authentication of endpoints (WS-Trust), shared-key exchange between endpoints (WS-SecureConversation), message-level encryption (WS-Security) and security policy.

6.2.3 Resource Management

The objective for the resource management subsystem is to provide generic management services for resource allocation, monitoring and capacity management based on decentralised and actor centric principles. These principles mean that services are loosely couple and that the management services the needs of the actors running the service. The services aim to support service providers in the negotiation of service level agreements (SLAs) with customers, specifically:

- deciding how to map business-oriented requirements to local resource management policies;
- checking and monitoring the status of commitments against available capacity and usage allowing the detection of potential SLA violations;
- predicting future capacity to steer the negotiation strategy for new SLAs.

Based on the evaluation feedback for the Resource Allocation service in GRIA 4 (See Section 4.5), SIMDAT will provide an SLA management service.

6.2.4 Execution Management

The objective for the execution management subsystem is to provide standards-based analysis and knowledge service containers. Various application specific (MSC.SimManager, Lion SRS) and generic analysis service containers (GRIA) were deployed during the Connectivity phase. Each analysis service container implements a bespoke interface for starting and controlling data processing activities. In this task, we will analysis available interoperability specifications related to data processing such as JSDL, UDAP, OGSA-BES and even OGSA-DAI, which provides a pipelined activity processing framework. We will work closely with WP3, who is leading the Data subsystem, to understand standard data delivery interfaces for staging tasks. The results should show how a standards based analysis service containers can be exploited directly by workflow tools through a common API.

6.2.5 Information

The objective for the Information subsystem is to provide service registries that can support the publication and discovery of analysis, knowledge and data services. SIMDAT applications have a variety of registry requirements depending on the use of the Information subsystem. In aero and pharma, dynamic workflow is a critical target where extensive semantic descriptions are required. In auto, a business registry support engineering services is needed. In meteo, where the structure of metadata is largely agreed and adaptive workflow is not a priority an XML based solution is sufficient. The discovery requirements drive the needs of the information model and expressiveness of the query language. As with all subsystems, the compliance with standards, availability and stability of solutions are also important when selecting technologies. We will evaluate SIMDAT developments (VGISC data registry, NEC's semantic broker) along with and external registry technologies such as UDDI and ebXML, Grimoires Semantic Registry making recommendations to the application activities as appropriate.

The following list summarises the technologies that have been evaluated:

- ***UDDI*** XML registry for business information (White, Yellow, Green Pages), standardised through OASIS and part of the WS-I Basic Profile. UDDI has a limited information model, as it was not designed for WSDL although best practice binding has been published. Queries are very limited and based on attribute value pairs.

- ***ebXML*** is an XML-based B2B infrastructure (business processes, collaboration protocol agreements, registries, etc) standardised through OASIS in collaboration with CEFACT. The Information model is quite good with core high-level business entities alongside service entities and an OO based classification scheme. Queries based on ebXML FilterQuery language and SQL (optional) although XPATH is likely to be support soon

- ***Semantic Registries*** are ontology based metadata registries based on standards such as RDF and OWL. The Information model is excellent allowing complex descriptions of domain knowledge to be represented. Upper ontologies for web services exist such as OWL-S. Queries and reasoning can be support with specifications such as SPARQL approaching standardisation through W3C.

- ***V-GISC registry*** developed by the meteo activity provides generic access to XML metadata fragments and can be used to discover and query any kind of datasets. The architecture and information model is flexible (indexing, schema) and queries are based on XPATH.

7 Conclusions

The Integrated Grid Infrastructure workpackage has successfully supported all application sectors in the development and elaboration of prototype developments during the Connectivity Phase. The Grid infrastructure software has been deployed in four SIMDAT prototypes and successfully demonstrated at the 1st SIMDAT review. The prototypes have shown that SIMDAT will change the way that scientists and engineer design artefacts from drugs to aeroplanes. Global industrial companies are now "Connected" to the process of development and deployment of Grid technology having attained an understanding of the potential benefits to their businesses. The application activities are now driving the evolution of Grid technology by supplying complex scenarios that require uniform and location-independent access to persistent and supported heterogeneous IT assets that are under different ownership or control.

The Integrated Grid Infrastructure software has been extensively enhanced to meet the needs of industry. A wide range of operating systems and queuing systems are now supported whilst robustness, usability and maintainability issues have all been addressed. SIMDAT application users can now participate in the collaborative design of complex products through the development of distributed workflows that cross domain boundaries.

Key partners in each application activity engaged with the Integrated Grid Infrastructure have provided evaluation feedback through a series of structured telephone interviews. These interviews have informed the specific requirements for the Interoperability phase and contributed to the overall architectural vision that underpins the SIMDAT Grid solution portfolio. SIMDAT application partners want to collaborate, but within the constraints of their business interests. To support these requirements, the Grid needs to provide management infrastructure that enforces organisation level security policies whilst sharing assets at an appropriate level of abstraction for individual relationships. The Grid needs to provide management capabilities that to support provision of service rather than direct access to a remote computational or data resources typically supported by most infrastructures today.

A SIMDAT architecture has been defined that provides a vehicle for communicating SIMDAT ideas, structure, component technologies and their relationships to application communities within and beyond SIMDAT, whilst defining the scope of the technical developments to be undertaken by technology developers within the project. During the interoperability phase, the Integrated Grid infrastructure software will be redesigned to provide a set of loosely coupled services based on the architectural subsystems. The services will allow application users to share assets in accordance with business objectives. Application users will be able to monitor and bill for usage of applications, information, clusters or storage depending upon their role within the value chain. The loose coupling of components will enable technology vendors easily integrate generic Grid capabilities into existing Problem Solving Environments. We will see a consolidation in the developments through conformance to the SIMDAT architecture and larger uptake of Grid technology across all application sectors transforming their ability to solve complex problems.