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

Deliverable

D.2.3.3: Report on SIMDAT Integrated Grid Infrastructure, evaluation and validation

and

D2.3.4: Report on finalised production ready SIMDAT Integrated Grid Infrastructure

Start date of project: 1 September 2004 Duration: 48 months
Due date of deliverable: 01/07/2008
Actual submission date:
Lead contractor for this deliverable: IT Innovation Centre
Revision: 1.0 public

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

This document is the deliverable D.2.3.3: Report on SIMDAT Integrated Grid Infrastructure, evaluation and validation and D2.3.4: Report on finalised production ready SIMDAT Integrated Grid Infrastructure of the EU IST-2002-511438 SIMDAT project, as specified in the Annex 1- “Description of Work” [1].

The document is the final technology report from WP2 Integrated Grid infrastructure. The document presents a chronology of the Grid technology evolution within SIMDAT and how generic enabling technology was delivered to industrial application sectors to support collaborative product design. A description of the final software delivered to the application activities for integration in prototypes and demonstrators is provided. The document analyses each requirement targeted at WP2 from both technology and application activities. The analysis describes how the requirement was met in terms for technical innovation and how the technology was taken up within application prototypes. Finally, the document provides a concluding section that outlines the overall impact of WP2 within industry and research communities.

The Grid infrastructure software is based on GRIA and has been since the successful demonstration of the aerospace prototype in the Connectivity phase. We have seen a migration towards GRIA’s security and management approach within SIMDAT and in other projects. Regular releases of GRIA have been distributed to SIMDAT partners in accordance with the delivery schedule incrementally including features relevant to application prototypes. The final infrastructure is based on GRIA 5.3 and includes all features agreed with partners and outlined in the description of work. The significant innovations delivered by WP2 include:

- Innovation 1: New modular service-oriented architecture supporting a range of deployment scenarios required by industrial application prototypes incorporating a standardised message structure based on SIMDAT’s industrial WSRF profile

- Innovation 2: The first Grid middleware to include a SLA management service within a B2B procurement model supporting monitoring, constraining and billing for service usage from a wide variety of application services

- Innovation 3: Contextualised B2B registry incorporating security policies with resource publication and discovery

- Innovation 4: Grid licensing infrastructure supporting integration of FlexLM with Service Level Agreements

- Innovation 5: Standardised job service supporting integration of multiple resource managers

GRIA has been successfully delivered and deployed in seven SIMDAT application prototypes. The prototypes demonstrate how the Grid infrastructure can support inter-Enterprise collaboration for the design of complex products. Partners involved in each prototype have evaluated and validated the solution for their scenario. A summary of the prototypes is given below (detailed descriptions are available in the relevant deliverables).

- Aerospace collaborative product design (BAE, EADS, CEDG, MSC)

---

[1] SIMDAT Annex 1 Description of Work

- Auto CAE/CAD/CAT integration (Audi)
- Auto crash compatibility testing (Renault, Audi)
- Auto grid licensing (ESI)
- Auto knowledge (Audi)
- Pharma target identification B2B outsourcing (GSK, Galapagos)
- Pharma data distribution B2A outsourcing (GSK, EMBL, ULB)

The objective of the SIMDAT project was to develop generic-enabling Grid technology that can support improvements in how industry collaborative for the design of complex products. We can see from the evaluation and validation by companies across a range of sectors such as BAE, EADS, and GSK that SIMDAT has achieved this goal. As we enter the demonstration phase we can reflect on a project that has delivered considerable value to the industrial players and software vendors participating in the programme. We have now moved from research and development into a world of opportunities. Industrial partners have the opportunity to use SIMDAT demonstrators to influence IT decision makers so that technologies selected support scenarios that make a real difference to collaborative design. Software vendors have the opportunity to incorporate advanced service-oriented capabilities within products that customers want. Grid technologists have an opportunity to exploit the knowledge gained and to start meeting the challenge of future service-oriented systems. SIMDAT represents the end of an era, as after more than 10 years of RTD industry is finally beginning to exploit the flexibility and diversity of the Internet for business benefit now that service-oriented technologies can be shown to mitigate many of the risks.

1 Introduction

This document is the deliverable D.2.3.3: Report on SIMDAT Integrated Grid Infrastructure, evaluation and validation and D2.3.4: Report on finalised production ready SIMDAT Integrated Grid Infrastructure of the EU IST-2002-511438 SIMDAT project, as specified in the Annex 1- “Description of Work”.

The document is the final technology report from WP2 Integrated Grid infrastructure. The document presents a chronology of the Grid technology evolution within SIMDAT and how generic enabling technology was delivered to industrial application sectors to support collaborative product design. A description of the final software delivered to the application activities for integration in prototypes and demonstrators is provided. The document analyses each requirement targeted at WP2 from both technology and application activities. The analysis describes how the requirement was met in terms for technical innovation and how the technology was taken up within application prototypes. Finally, the document provides a concluding section that outlines the overall impact of WP2 within industry and research communities.

The intended audience for this document are the application and technology partners within the SIMDAT consortium, as listed in Section 3 of [1], as well as the European Commission Services. SIMDAT partners (stakeholders) have a diverse range of expertise representing both application sectors and horizontal technology activities.

2 SIMDAT Grid Infrastructure Chronology (2004 – 2008)

Grid and Web Services infrastructures have evolved significantly throughout the SIMDAT project as the technological and political landscape have shifted. We have tracked this landscape and regularly updated project partners on the state-of-the-art technologies and how emerging ideas impact the development plan for the SIMDAT Grid infrastructure. The SIMDAT project is structured in four phases each building on the lessons learnt from previous phases to deliver advanced Grid infrastructure solution to application prototypes and demonstrators.

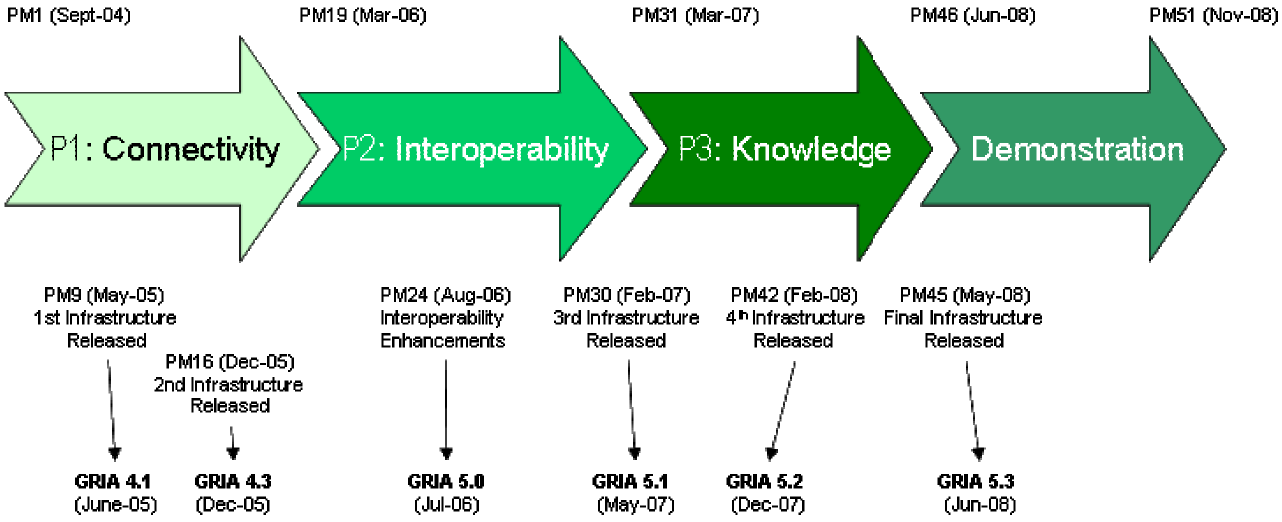

Figure 1: Grid infrastructure releases

Figure 1 shows the SIMDAT project phases and the Grid infrastructure releases delivered to the project. GRIA was the primary infrastructure for SIMDAT from the interoperability phase and all future Grid infrastructure releases where based on GRIA distributions. Each release included innovations from SIMDAT and other RTD projects developing advanced Grid solutions. The following sections provide a summary of each phase and the innovations delivered from WP2.

2.1 Connectivity Phase

The objective of the Connectivity phase 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 initial application prototypes. The key success measure was to understand the needs of industrial users and deploying Grid technology within an industrial context allowing businesses to explore for the first time how the Grid technology can support the collaborative design of complex products.

Back in 2004, the state-of-the-art in Grid systems was characterized by Web Service infrastructures aiming to support inter-Enterprise collaboration scenarios alongside traditional academic and scientific applications. Grid technologies moved from second-generation services (Globus Toolkit 2.X, UNICORE) based on low-level bespoke protocols to Web Service based services (GT4, GRIA, OMII, gLite) using open standards, although there were significant differences between security and management approaches.

The selection of Grid infrastructure components was made by application sectors based on evaluating and selecting publicly available middleware components. 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 PM12 prototypes. The aerospace and automotive sectors chose GRIA for the PM12 prototypes due to the GRIA's explicit focus on industrial Grid deployments and immediate availability, although MSC and Audi from the automotive sector reviewed their decision after PM12 following the release and stabilisation of gLite and GT4. 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 after PM12. The pharma sector decided to develop a web services Grid leveraging E2E security components developed during the GEMSS project.

By the end of the connectivity phase the technologies deployed within application prototypes reflected the heterogeneity of Grid infrastructures available on the market. SIMDAT application users and tools vendors were still facing significant uncertainty regarding the applicability of technologies for inter-domain collaboration and longevity of solutions and standards. GRIA 4.3 was delivered to SIMDAT as the 2nd Grid infrastructure release with major robustness and usability improvements, however, it was clear from evaluation feedback that significant architectural changes were needed if SIMDAT Grid technology were to meet the industrial challenge. The Connectivity phase bridged the gap between the commercial users' perception of the Grid and the technical realities. The hype was broken and we then had a better understanding of how SIMDAT research and development needed to be structured to deliver an infrastructure that could support the collaborative design of complex products.

2.2 Interoperability Phase

The overall goal of the Interoperability phase was to develop generic Grid infrastructure services by improving interoperability of Grid services (e.g. Web Service and Grid specifications and standards) 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.

After almost 10 years of RTD, the exploitation of Grid technology by industry and commerce was a key target for researchers and Grid systems vendors in 2006. Many ongoing initiatives were working to understand how Grid technology can support a new market-based economy in addition to the conventional e-Science and business co-operative sectors. Projects began addressing these challenges by major advancements in Grid management services to support operational security and business objective requirements. A new-style Grid operation model emerged, as supported by GRIA and NextGRID experiments that reflected business models found in the real world and enabled more flexible Peer-to-Peer (P2P) collaboration architectures based on distributed management principles.

In 2004, we saw the launch of a new collection of standards called the "Web Services Resource Framework" (WSRF) by The Globus Alliance with IBM and others with the expectation that these standards would provide convergence between Web Service and Grid specifications. These proposals were made to OASIS, built on existing and emerging Web Service standards, and were seen as a key step that allows convergence between Web Services and the Grid by supporting applications that require services to support stateful long-running activities. However, although WSRF was ratified by OASIS in spring 2006 and is compatible with wider Web Services standards (and their likely future development), it was somewhat controversial. Key vendors, Sun and Microsoft, did not back the proposals and went another direction. The major vendors (HP, Intel, Microsoft, IBM) realised that the dispute needed to be resolved and they published a joint WS-Convergence white paper detailing a roadmap for converging these specifications between 2006 and 2009, however, progress has been difficult and we are still some way from middleware products that support open standards derived from the WS-Convergence dialog.

In SIMDAT, we monitored the situation and published a report D2.2.2[2] outlining standards and specifications that could be safely adopted within SIMDAT prototypes. The decisions made by the technical management board have proven to be good and have reduced the SIMDAT effort spent on aligning infrastructure with standards/specifications that are no longer beneficial for interoperability.

The Interoperability Phase Grid infrastructure was based on GRIA 5, which included key innovations resulting from SIMDAT RTD based on industrial requirements captured from the application sectors during the connectivity phase evaluation and architectural analysis:

- ***INNOVATION 1: New modular service-oriented architecture supporting a range of deployment scenarios required by industrial application prototypes incorporating a standardised message structure based on SIMDAT's industrial WSRF profile.*** *This innovation transitioned SIMDAT from reliance on a monolithic Grid infrastructure to a toolkit of services that can be configured and deployed to support a range of application needs. Application users could choose the capabilities they needed and integrate these services with other services developed in platforms such as .NET or even Globus.*

- ***INNOVATION 2: The first Grid middleware to include a SLA management service within a B2B procurement model supporting monitoring, constraining and billing for service usage from a wide variety of application services.*** *This innovation allowed service providers could now offer services relevant to their sector and specify the terms and*

[2] D2.2.2 Report on Grid infrastructure interoperability challenges

conditions for using. Customers could procure those services and clearly understand the commitments their suppliers were making and price they would have to pay.

The success of the Integrated Grid Infrastructure software for inter-domain industrial partnerships during the connectivity phase within the aerospace and automotive prototypes resulted in a greater uptake of GRIA by other application sectors during the Interoperability phase. All automotive prototypes and pharma prototypes decided to adopt GRIA, whilst the meteo activity re-evaluated GRIA's security and service management packages for use within VGISC and has incorporated GRIA to provide a single point of access to elaborated products with the BRIDGE project.

2.3 Knowledge Phase (PM30-PM45)

The overall goal of the Knowledge phase was to further develop generic Grid infrastructure services by improving alignment with operational security requirements and support for cost-effective working practices required for industrial uptake of the technology at project end.

By 2007, inter-enterprise service-oriented collaborations had been proved in many major sectors and companies were adopting the technologies in real-world pilot applications within their businesses. Many of the architectural principles for creating and maintaining business relationships were understood and frameworks implementing the underlying policy management capabilities were available. Core security federation standards were now largely approved or at least in a recommendation phase. There were still differences of opinion about the future of some core infrastructure standards between IBM and Microsoft. Service level agreements emerged as a core building block for governing the provision of a wide range of IT assets such as applications, information, computation and storage resources. Market-based service economies began to be discussed and organisations began to adjust their business models and operating policies to the changing market conditions that incorporate more flexible and adaptive working beyond their traditional organisation boundaries. The Grid and Web Services brands had largely converged under the banner of service-oriented architectures (SOA) and service-oriented infrastructures (SOI). Further convergence occurred within enterprise IT data centres as virtualisation, cluster management and multi-core processing solutions all jostled for position.

The knowledge phase prototypes were based on GRIA 5.2 and 5.3, which included further significant innovations resulting from SIMDAT RTD:

- ***Innovation 3: Contextualised B2B registry incorporating security policies with resource publication and discovery.*** *This innovation allowed consumers to publish distributed resources (datasets, SLAs, etc) for use by only members of certain project teams ensuring they automatically have access to the capabilities and data they need to perform their work*

- ***Innovation 4: Grid licensing infrastructure supporting integration of FlexLM with Service Level Agreements.*** *This innovation allowed the three key stakeholders (consumer, Grid service provider, and independent software vendor) to collaboration in Software-as-a-Service models in an accountable and secure way*

- ***Innovation 5: Standardised job service supporting integration of multiple resource managers.*** *This innovation helped improve interoperability between GRIA and other middleware supporting the JSDL specification and allowed service providers to connect multiple clusters managed by different resource management technologies to a single service provider.*

During the knowledge phase (PM31 – PM45) pharma and aero prototypes exploited the new innovations and WP2 focused on supporting the automotive activity in the adoption of GRIA's security and management models. The final months of this phase focused on the testing and release of GRIA 5.3 which incorporates all of the innovations developed within SIMDAT. Please see section 3 for a detailed description of the final infrastructure.

2.4 Demonstration Phase and Future Outlook

Following a series of prototypes SIMDAT is now entering a demonstration phase where the objective is to demonstrate technologies from the SIMDAT Grid Solution Portfolio to business units within industrial partner organisations and the wider community. Application partners have defined a set of proof-of-concepts that are relevant to various actors within their organisations such as scientists, engineers, operational security, and project management representatives. The desired outcome for SIMDAT partners is to significantly influence IT architecture and technology choices and in some cases for SIMDAT technologies to be directly adopted as components in the B2B toolkit.

Throughout the lifetime of SIMDAT we have seen an increased adoption of service-oriented systems within the Enterprise to improve the agility of internal business processes. SIMDAT has challenged the enterprise to explore scenarios that go beyond these traditional enterprise boundaries and understand the policy impact of emerging security and management technologies. After more than 10 years research and development of service-oriented systems economically viable global service networks have yet to materialise. We will see that the reliance on service-oriented systems dramatically increase the enterprise endeavour to meet the economic challenges of the next decade. However, the complexity and scale of applications will expand by orders of magnitude due to increasing need for flexibility, dynamicity, mobility, and distributed policy and regulatory compliance in a network of highly connected diverse physical and IT assets under different stakeholder control. Even though SIMDAT has achieved a great deal many challenges exist that will require real innovation, such as the tools, services, architectures and theories to support the application of emerging service ecosystems, supporting new business models and new forms of collaborative working, with self-governance to address business, social and legal challenges presented by these new paradigms.

Current software engineering theories and service specifications assume software lifecycle models that significantly restrict the potential of service-oriented systems and the ability of such systems to support meaningful and dynamic economic relationships required by future collaborative design scenarios. Today, service-oriented systems are developed using software engineering approaches that assume system boundaries and behaviour (functional and non-functional properties) can be defined, analysed, verified and validated in advance of service deployment and operation. The assumption that an entire system can be known a priori no longer holds because consumers need to recruit, monitor and govern resources they need for applications from a diverse set of devices, content, services and network connectivity options based on dynamically changing business objectives. Non-functional characteristics of these applications cannot be known at design time and availability of such applications will lie in different domains of control. Such service-oriented systems will be extremely flexible, but this flexibility stems from inherently volatile and new design theories and modelling approaches will be required to assure security, reliability, availability and commercially viability by design in this environment.

Increasingly, complex distributed systems and information will be built by non-IT experts who will potentially vastly outnumber those who are formally trained in software and systems engineering. Service domain abstractions and interaction patterns will be needed to build application specific solutions and transition to a world of user generated content and services developed, provisioned and programmed by an increasingly diverse set of non-IT experts. The complexity of such systems and lifecycle interactions will require new goal oriented programming abstractions that are sufficiently expressive to support consumers with a wide range of technical and domain expertise to intelligently compose services, to engineer applications and systems at all infrastructure levels (service, network, energy). These abstractions should allow service operators and consumers to assess risk, and manage the security, availability and trustworthiness of applications considering the scale, complexity and dynamic nature of opportunities and threats in a global service network. Self-governance will be essential to support cost-effective operation of such systems and new service platforms will need to be developed combining monitoring and predictive techniques that allow services be managed to achieve stakeholder objectives by behaving intuitively and recognising business insights in operational environments.

The Grid brand has now returned to its origin with most players considering the technology as providing access to computational resources rather than wide aspects of distributed computing. The Future Internet has now taken centre stage and will be as much more about converging technologies support consumers and suppliers. The creation and governance of applications of service-oriented infrastructures must become much easier for all types of users as the diversity and scale of assets dramatically increases, especially for applications that span multiple administrative domains. Socio-economic issues will grow in importance, and an interdisciplinary approach will be essential to ensure success.

3 Final Grid Infrastructure Software Description

The final Grid infrastructure software is based on GRIA 5.3. GRIA is a service-oriented infrastructure designed to support B2B collaborations through service provision across organisational boundaries in a secure, interoperable and flexible manner. GRIA uses Web Service protocols based on key interoperability specifications. GRIA makes use of business models, processes and semantics to allow service providers and users to discover each other and negotiate terms for access to high-value IT assets. By focusing on business processes and the associated semantics, GRIA enables users to provision for their computational needs more cost effectively, and develop new business models for their services.

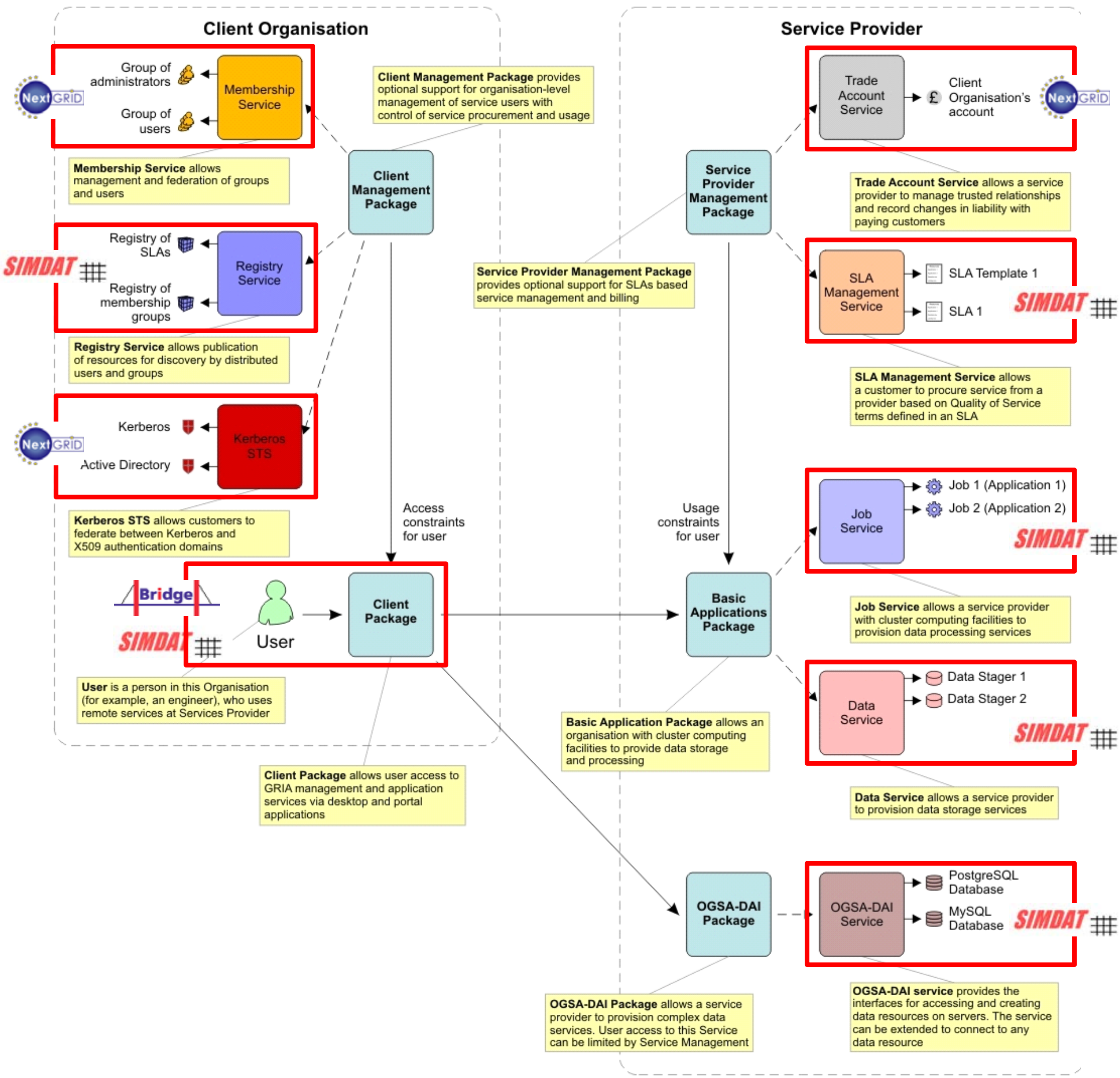

Figure 2: Final Grid infrastructure based on GRIA 5.3

The high-level architecture for GRIA is shown in Figure 3 showing software packages, services and resources. The diagram also shows the primary RTD project that has contributed innovation to each service. We see that SIMDAT's applied nature has resulted in a focus on the management of applications (i.e. SLA, data, job and OGSA-DAI) whilst NextGRID has focused on economic markets where accounting and dynamic federations between administration domains is critical (account, membership and Kerberos). Each project has significantly benefited through concertation at conceptual and technical levels.

GRIA has provided the foundation infrastructure technology for developments within the aerospace, automotive and pharmaceutical sectors and provides an important technology within the SIMDAT Grid Solution Portfolio. Figure 3 shows the overall SIMDAT architecture and how GRIA provides many of the capabilities within the offering.

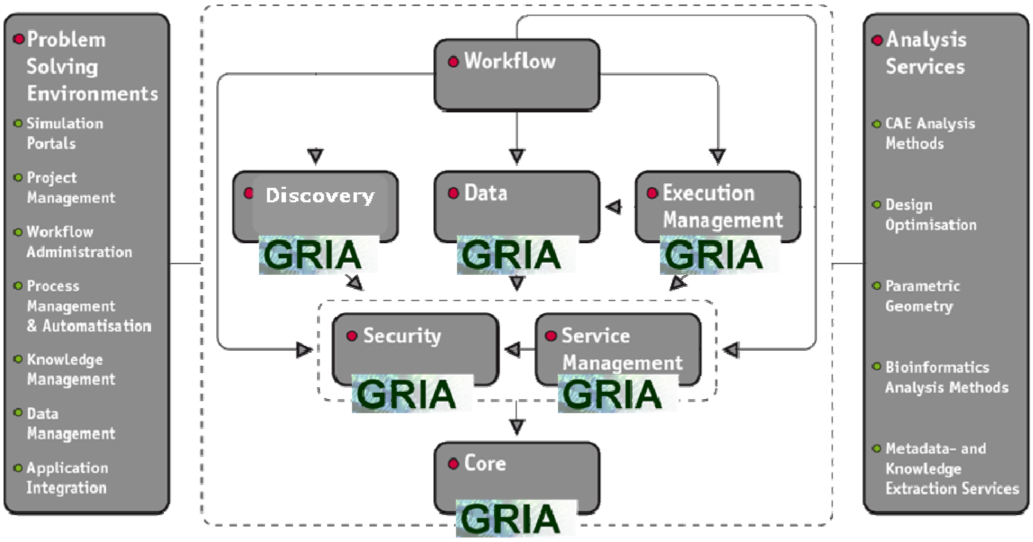

Figure 3: SIMDAT Architecture

An outline how GRIA supports each of component in the architecture is described below:

- **Core:** underlies all Grid service developments and provides web service containers that support contextualised (stateful) Web Services compliant with interoperability specifications WS-I and WSRF. GRIA provides a developers toolkit for Web Services that extends Apache AXIS for alignment with standards and specifications in accordance with the SIMDAT industrial Grid profile.

- **Security:** 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. GRIA provides a security toolkit that includes a WS-Security implementation based on extensions to Apache WSS4J, Process Based Access Control for service authorisation policies, dynamic service federation policies described using WS-Policy and security token services for group management and integration with MS Active Directory.

- **Service Management**: 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. GRIA provides an SLA management service allowing consumers and providers to govern provision of application services.

- **Data:** provides data transfer, storage, federation, replication, synchronisation as well as semantic mediation between different data repositories. GRIA provides a file storage service that is used in conjunction with the job service.

- **Execution Management:** responsible for managing the lifecycle of data processing tasks including staging data transferred from the Data subsystem and reporting usage metrics to Service Management.

- **Information:** supports the publication and discovery of services, data and workflows providing data processing, storage, and transfer capabilities. GRIA provides a standardised job service supporting the integration of heterogeneous resource management systems.

4 Evaluation and Validation

4.1 Requirements Reprise

SIMDAT Grid infrastructure was developed using an iterative development cycle as shown in Figure 4. The project is currently completing the final development phase "knowledge" at PM45 (extended from PM42) prior to industrial demonstrators from PM46-PM50.

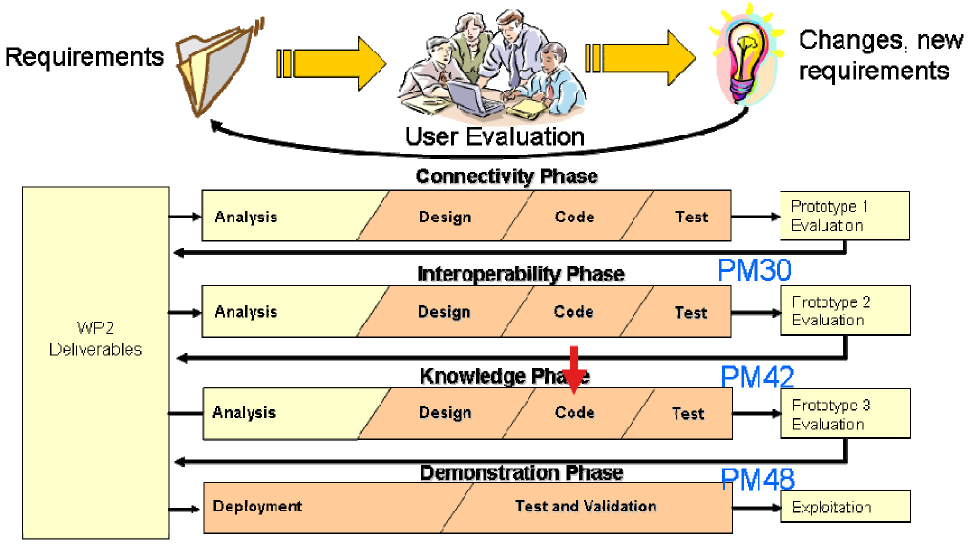

Figure 4: WP2 software engineering process

Three requirements capture and evaluation phases of SIMDAT identified 40 requirements for WP2. To help to identify the importance of the requirements classification criteria was identified.

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

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

A summary of the key WP2 requirements is shown in Figure 5. The requirements are structured

The requirements were incrementally delivered throughout the project with the final infrastructure meeting 85% of all requirements and 94% of all priority 1 requirements. Further requirements were identified following Infrastructure accreditation completed by WP4 Virtual Organisations during the knowledge phase. Although early insights were factored into the final work plan not all of these requirements (audit, improved usability of management services) have been addressed in the final infrastructure due to the late inclusion.

ID	Description	Applications	Delivered
CI1	Provide a generic Job submission portal	Aero, Auto	
CI2	Refactor GRIA 4.3.0 Java API	Auto	4.3
CI3	Client-side evaluation of service provider performance and heart-beat checking.	Pharma	
EM1	Support interactive real-time visualise of engineering simulations	Auto	
EM10	Clients shall be able to copy data from local stagers to the job working directory during job execution	All	5.2
EM11	Clients shall be able read intermediate output data generated by applications that have been terminated	All	5.2
EM12	Resource Managers shall be able to report platform usage metrics for SLAs	All	5.2
EM13	Applications shall be able to report application usage metrics SLAs	All	5.2
EM14	Application Metadata shall support command line arguments, stager arrays, modifiable stagers and optional stagers	All	5.2
EM2	Support integration of multiple backend queuing systems with a single analysis service	Auto, Aero	5.2
EM3	Allow users or service providers to pause and resume jobs	Auto	
EM4	Provide a simple way of giving parameters to an analysis service	Auto	5.2
EM5	Support a standard analysis service interface	Auto	5.2
EM6	A client shall be able to specify their exact requirement for the number of input and output stagers	All	5.2
EM7	A client shall be able to describe a job's resource requirements at job creation time.	All	5.2
EM8	A client shall be able to specify command line arguments at job creation time	All	5.2
EM9	A client shall be able to view application status information, in addition to the overall status of an activity	All	5.2
IN2	Deployment of services within JBoss	Auto	5.2
IN3	Integration with Visual Composer	Auto	5.3
IS1	Support analysis services discovery	Auto, Aero	5.1
IS2	Support information discovery	Pharma, meteo	5.1
IS3	Support for building registry services for data and analysis service in the automotive SOA -Architecture	Auto	5.1
IS6	Extended registry for service selection	Aero	5.1
IS7	Semantic broker integration with discovery infrastructure	Pharma	5.2
R1	Allow QoS commitments with softer commitments on resources and time periods	Auto, Aero	5.0
R2	Support a very generic QoS model handling a wide class of application services	Auto	5.0
R3	Support a resource reservation model	Auto	
R4	Support management of application and data resources	Auto	5.0
R5	Support Grid license distribution	Auto, Aero	5.2
S1	Compliance with organisation security policies and standards such as ISO/IEC 17799/27001	All	
S2	Support trust relationships that are grounded in out-of-band policies rather than the ability to pay for a service	Meteo	5.1
S3	Support auditable security logging	Aero, auto	
S4	Allow client organisations to locally manage access control lists for remote resources provided by other organisations	Meteo	5.0
S5	Decentralised relationship management	Meteo	5.0
S6	Assign access rights to services (per application suite)	Auto	5.0
S7	Convergence between GRIA and NEC DAC security architectures	Pharma	Architecture discussion completed
S8	Integration with home authentication schemes	Aero, meteo	5.3
S9	Portal security	Auto, Meteo, Pharma	5.3
U1	Improve GRIA 4.3.0 documentation	All	4.3
U2	Provide a technical FAQ	All	5.1
U3	Provide additional metadata about stateful resource ids	Auto	5.0
U4	Improve exception handling and error messages	All	5.x
U5	Improvement availability of service usage information	Auto	5.0
U6	Support other databases that PostgreSQL	Auto, Aero	4.3
U7	Refactor GRIA 4.3.0 services so that they can be used independantly	Pharma, Auto, Aero	4.3
U8	Use a web-based installer for the keystore generation	Auto	5.0
WSC1	Standard mechanism to access stateful resources using web services (WSRF)	Auto	5.0

Figure 5: Key WP2 requirements

4.2 System Validation

The Integrated Grid infrastructure was developed using rigorous software engineering best practices. The methodologies selected were based on the need to ensure that the software delivered to application partners was to sufficient quality that proof-of-concepts could be deployed and insights into inter-domain service-oriented collaboration could be understood. SIMDAT is a near-to-market RTD project with close industrial partner engagement. The expectations of industrial partners in terms of software quality was high and there was a need to deliver robust and reliable software releases that addressed the core requirements outlined in Section 4.1.

When developing software in RTD projects we must balance the need for innovation against the need for software quality and this often the results in compromises that are not seen in commercial product development. It is well known that there is a factor of 10x resources required to take a proof-of-concept from the research lab into a production environment. If SIMDAT attempted to do this with the integrated Grid infrastructure the level of innovation generated within the project would have been significantly reduced. Therefore, the software engineering approach was balanced but grounded in best practice.

The software lifecycle was iterative and user centred allowing early feedback at the end of each development cycle for the identification of new and changed requirements. Throughout the WP2 deliverables we can see how the evaluation of the infrastructure had identified new challenges and these were factored into the description of work and software development plans. For example, software licensing issues were identified during the Connectivity phase and a proof-of-concept was finally developed in the automotive sector with ESI in the Knowledge phase prototype. Within each development cycle a range of approaches were used to design, develop and test the software releases. A summary of the techniques adopted is given below:

- **Use case and Scenario Analysis**: the final aerospace prototype moved from a supply chain model to a VO model based on bi-lateral relationships governed by service level agreements. This scenario required the development team to understand how the infrastructure could support different collaboration scenarios where security and federation contexts could be discovered using different business processes. The interactions between a client, token services and registry services was extremely complex and only by working through use cases and aerospace scenarios could a suitable solution be proposed.

- **Simulation and Mock Ups**: developing and testing a distributed system based on SOAP messaging requires each component developer to reach an advanced implementation stage to allow integration to occur. Often it is during integration that architectural weakness is identified requiring significant rework of existing code. To address this issue WP2 adopted a simulation approach where developers built an entire mock up system in process using Python in a few days. A mock up of each component was produced very quickly allowing interactions between components and underlying algorithms to be tested based on the scenario analysis. The simulation code was then translated to Java as a starting point for Web Service development. The result was that architectural issues were identified early in the software development lifecycle and reducing costly integration time later.

- **Unit, Continuous Integration, Compatibility, Upgrade, Stress and Release Testing**: a structured software testing framework is essential to deliver a target level of quality in software distributions. The infrastructure adopted testing throughout the development lifecycle with continuous integration providing developers with ongoing monitoring of the state of any build. Unit and integration tests were developed during the implementation phase and these were run continuously using the Hudson continuous integration tool. In addition, stress testing is executed regularly to gauge a measure of performance and scalability of key services. During the release phase of any distribution release testing is performed by the testing team who operate the system in a range of deployment scenarios driven by SIMDAT requirements. The flexible nature of SOA means that services can be deployed to operate in many different configurations. The adaptability and customisation of services and the fact that in some situations services are used in ways that are not imagined at design time presents a significant testing challenge. Test cases grow exponentially and backwards compatibility between different infrastructure versions only adds to the number of tests required. In SIMDAT, we adopted a pragmatic approach based on the need to deliver software on time for prototypes and demonstrators, including key features for the project and ensuring that the software worked. Test coverage at the unit test level was monitored to ensure that at least 75% of code paths were executed and for integration we focused on partner needs across a range of prototypes and core capabilities rather than working through all scenarios.

We can see from the end-user evaluation given in Section 4.3 that the overall software engineering approach adopted by WP2 has been successful in delivering robust solutions on time for application prototypes and demonstrators. GRIA support played a critical role in the success of the SIMDAT project providing best efforts response times but usually an answer within 24 hours of receiving a request. There have been bugs and suggested improvements from partners but the support process adopted by WP2 has ensured that when users face problems they have immediate access to expertise. This has ensured that other application and technical partners can get on with developments either by the adoption of workarounds or the deployment of a patch. New features were factored into the development roadmap and delivered to partners later in project phases.

4.3 Industrial Evaluation and Validation

The Integrated Grid Infrastructure software has been successfully delivered and deployed in seven SIMDAT application prototypes. The prototypes demonstrate how the Grid infrastructure can support inter-Enterprise collaboration for the design of complex products. Initial prototypes were successfully demonstrated at the 3rd SIMDAT Review in November 2007 and these have now been refined further during the Knowledge Phase. A summary of the prototypes is given below (detailed descriptions are available in the relevant deliverables).

- Aerospace collaborative product design (BAE, EADS, CEDG, MSC)
- Auto CAE/CAD/CAT integration (Audi)
- Auto crash compatibility testing (Renault, Audi)
- Auto grid licensing (ESI)
- Auto knowledge (Audi)
- Pharma target identification B2B outsourcing (GSK, Galapagos)
- Pharma data distribution B2A outsourcing (GSK, EMBL, ULB)
- Meteo VGISC (ECMWF)

Evaluation feedback was captured throughout the project through a series of structured face-to-face and telephone interviews. Key partners involved in the prototypes from both technology and application work packages were contacted such as MSC, BAE, LMS, ESI, EADS, UoS, Inforsense and GSK. The following sections describe each prototype and the evaluation feedback captured from each sector.

4.3.1 Aerospace collaborative product design (BAE, EADS, CEDG, MSC)

This prototype focused on distributed process-centric workflows for the design optimisation of an aerospace component. The scenario was driven by the need to integrate engineering capabilities delivered by different partners in distributed locations. The early prototypes provided the partners with some insights into how this could be achieved using Web Service security and management technologies and how different collaboration patterns could be implemented. The collaboration model adopted by the aerospace prototype changed from the initial design supply chain to a business cooperative but always building business relationships based on bi-lateral agreements between companies.

The evaluation of the infrastructure focused on understanding how SLAs and dynamic trust and security technologies can govern business relationships. The partners successfully demonstrated that a virtual organisation can be established and operated using the infrastructure and that the infrastructure provided an acceptable approach to managing users. A key requirement was to support federated management of teams within the virtual organisation with no one partner responsible for VO membership of individuals (Requirement: S5). This was achieved using GRIA's membership security token service building on the WS-Trust and SAML specifications. Another key requirement was the ability to govern access to resources committed by each participant of the virtual organisation (Requirement R2). This was achieved by allowing each partner to publish an SLA to the virtual organisation outlining the services, applications, resources and quality of service they are willing provide. Each participant providing service remained in control of their resources and the prime contractor and other consumers could monitor that a participant is delivering to their commitments. Dealing with heterogeneous resource managers was important as each aerospace partner used a different cluster management technology (Requirement EM2). This was delivered in the Interoperability phase and within prototypes since we have connectivity between Condor, Sun GridEngine, Torque and Windows Scheduler.

The aerospace prototype has validated that the architecture of the SIMDAT infrastructure is appropriate for collaborative design teams. This was really established during the Interoperability Phase with GRIA 5 and enhanced for usability in the Knowledge phase with GRIA 5.3. The aerospace partners have sent a clear message to other companies that the SIMDAT approach is valid and can provide significant benefit. We see in the demonstration phase engagement with both Airbus and Rolls Royce, aerospace partners not directly involved in the project but engaged through the enthusiasm for SIMDAT technologies from BAE and EADS.

In addition to architectural and feature validation BAE systems have focused on understanding how GRIA can be taken from a research prototype into a production. BAE performed a security assessment as part of the VO work package and identified a series of technology gaps that would need to be addressed before GRIA could be adopted including areas such as home authentication, auditability and manageability of large scale distributed systems. BAE also examined how service-oriented systems could be accredited for production use for a complex engineering scenario. They concluded (not surprisingly) that reliance on service-oriented systems for inter-enterprise collaboration would require a framework for certification and accreditation of suppliers and the technology/working practices used. This certification framework for service providers does not exist today and although the technology research is fundamentally completed and we are in the realms of product development rather than RTD, production adoption of SIMDAT for commercial aerospace design remains some distance off.

4.3.2 Auto CAE/CAD/CAT integration (Audi)

This prototype focused on integrating distributed design teams using heterogeneous data management teams with the VW Audi Group. The initial prototypes examined how semantic technologies could support integration between data schemas and only during the Interoperability and Knowledge phase did the partners begin to analyse the governance of data transfers between the teams.

A security model for 3rd party semantic mediation between consumer and provider was developed that integrated the MSC SimManager and LMS Tec Manager data management systems via OntoBroker. The security model built on GRIA's dynamic security and SLA management services for propagating authentication/authorisation checks and mapping PDM user roles to SLA subscriptions. The initial evaluation identified that GRIA's client API needed to be updated to support multiple X509 identity tokens when used with a portal (the user interface technologies for most PDMs) (Requirement S9). The portal would need to act as a certificate authority issuing X509 credentials from the basic authentication provided by the PDM and that a replying service (OntoBroker) would need to trust the PDM CA certification policy. This capability was delivered in GRIA 5.3 ready for the demonstrators.

The validation of the security model has shown that security federation between problem solving environments can be achieved using GRIA. The validation of security has been weaker in this prototype than aerospace because although security and management is seen as an important requirement the focus has been to understand data integration using ontologies and rules. Therefore engagement has been with design engineers rather than operational security experts within Audi. However, the design engineers who have evaluated the solution have said that federated access to engineering data can be controlled in a satisfactorily way.

4.3.3 Auto crash compatibility testing (Renault, Audi)

This prototype aimed to utilise Grid technology to protect intellectual property rights (IPR) when two OEMs collaborated to perform a crash compatibility test. The scenario built on earlier work achieved for the integration of suppliers with OEM design processes.

The prototype relied heavily on Grid security models for delegation of rights to fine-grained access control policies (Requirements S4, S5, S6). A trusted-third party was required to execute the simulation to ensure neither party had direct access to cluster which they could use to access competitor models through deviant techniques. The key technical innovation in this prototype was to automate security policy side effects (access control decisions) within the engineering process. By coupling dynamic security with workflow, OEM input models and the resultant behaviour of each vehicle could be controlled to ensure that each OEM only had access to their data. Execution management was a key objective of this prototype and various requirements where identified by ESI during the initial project phases that resulted in enhancements to GRIA's job service (Requirements EM2-EM14). These enhancements were delivered at the beginning of the knowledge phase in GRIA 5.2 and included in later prototypes.

The validation of the security model was performed by ESI who demonstrated a crash compatibility test between representative Audi and Renault vehicles. A test plan was created that executed the scenario within different stakeholders. At each point access control privileges where checked to ensure appropriate policies where in place to protect the data. This was demonstrated successfully at the 3rd SIMDAT review.

4.3.4 Auto Grid licensing (ESI)

This prototype aimed to explore Grid licensing scenarios where consumers want to run vendor applications at a 3rd party service providers. Grid licensing models have been a known issue for many years and the objective was to explore how SIMDAT technologies can support management of application usage. Two scenarios were explored:

1) Service provider sub-licenses ESI PAMCrash application bundles to consumers as part of their service provision offering with ESI.

2) Service provider has agreement with ESI to run PAMCrash providing the consumer has a valid license with ESI

For each case ESI needed to be notified of PAMCrash usage by the consumer. For the prototype we focused on scenario 1.

This scenario exploited the generic capabilities of GRIA's SLA management service. ESI where able to define application metrics for PAMCrash bundles (standard crash, safety, etc) and have the SLA service constrain usage of the PAMCrash application (e.g. number of crash hours) and apply subscription or pay-per-use charging terms as appropriate. The architecture linked Marcrovision FlexLM to GRIA's SLA management service allowing usage metrics generated by PAMCrash as a result of FlexLM licensing checks to be reported to the SLA service. Given this information GRIA was able to record usage, charge and take management decisions based on the consumer usage.

The validation of the architecture did identify some limitations with the current management protocol between the SLA service and managed services. ESI wanted to have SLA's to represent the relationships between different parties. An SLA was required between the consumer and service provider to govern the commitment to provide PAMCrash and computational resources. An SLA was also required between the service provider and software vendor to govern the usage of PAMCrash at the service provider and to monitor customer usage. Maintaining a link to the consumer was important for ESI and this would have been lost.

The problem identified was that a GRIA service (e.g. Job Service) can only be managed by one SLA management service. All usage reports generated by the Job Service would only be reported to the service provider's SLA service. In the scenario ESI needed to report usage to a SLA management service at both the service provider and the software vendor and this was not possible because the Job Service only support WS-Notification pull points. Further implementation of the WS-Notification specification would be required to support subscription of SLA services to metric topics but this could not be achieved within the timescales of the SIMDAT project.

4.3.5 Auto meshing (Audi)

This prototype focused on outsourcing a CAD meshing tasks to 3rd party engineering consultancy companies. The benefits of outsourcing specialist capabilities are well known (improved innovation, reduced costs, increased flexibility) and automotive companies increasingly looking to exploit the Internet and Web Services when outsourcing in collaborative design processes. The prototype utilised a reverse auction negotiation protocol to allow suppliers to compete on quality of service and price for meshing tasks.

The implementation used GRIA's SLA management service as the basis for governing business relationships between the OEM and their approved meshing suppliers. Project managers created SLAs with suppliers to provide the context for meshing task negotiation, which was carried out at a technical level between CAE engineers and meshing experts. The CAE engineer created a meshing proposal request which was submitted to suppliers who had existing SLAs previously agreed by the project manager. A receipt of a request a meshing expert created a meshing proposal that outlined metrics such as number of meshing hours based on the task information received. These metrics were governed by the SLA between the OEM and supplier and could be monitored by project management within each organisation.

The prototype was demonstrated at a workshop hosted by Audi in Ingostadt, Germany. IT Innovation demonstrated to major automotive companies and technology suppliers how GRIA can help support collaborative product design. Audi challenged different IT vendors to develop and demonstrate solutions to a complex engineering design scenario including CAE (Computer Aided Engineering), CAD (Computer Aided Design) and external consultant engineering teams. Significant business and technical challenges had to be tackled. A key aspect of the scenario was outsourcing CAD meshing tasks to 3rd party engineering consultancy companies. The SIMDAT prototype was the only solution that tacked all of the technical challenges and was highly commended by Audi and other OEM's such as Ford. The prototype demonstrated the SIMDAT has delivered generic enabling technology that can support a range of collaboration scenarios.

The validation of this prototype will be completed during the demonstration phase where the meshing service will be integrated with ESI's VisualDSS and deployed at Renault with a Korean distributor. The seamless vertical integration of Grid technology with ISV products, such as ESI's VisualDSS, is key during evaluation and validation.

4.3.6 Pharma B2B and B2A outsourcing (GSK, Galapagos, EMBL, ULB)

The focus for this prototype was to allow GSK to explore how external service providers from business and academic can deliver advanced and commodity capabilities to scientist working within their drug discovery pipeline.

The scenario was based an integration between Inforsense and GRIA allowing GSK scientists to develop workflows that seamlessly include tasks from within the enterprise and from suppliers in distributed locations. Security and management of pharmaceutical data was a key concern along with the accountability of business relationships. The implementation built on GRIA's security and management capabilities and GSK gained significant value from the lessons learnt by the aerospace prototypes which were leading the way in terms of complexity of business relationship management.

The validation of the prototype has been performed by all members of the pharmaceutical activity. The value that GRIA provides has been indentified by all participants and a future workshop at EMBL will examine how research costs in pharma can be reduced by outsourcing parts of their analysis pipelines using virtual organizations (VOs). GRIA is a foundation for this event and GSK will show how significant cost savings can be achieved by using GRIA to outsource part of their sequence analysis pipeline. In fact, GSK see GRIA as an important part of their future B2B toolkit and are championing its use in production environments at GSK today.

4.3.7 Meteo VGISC (ECMWF)

In the EU IST BRIDGE[3], the meteo activity is aiming to develop elaborated meteorological products from model outputs that are produced by geographical distributed applications, initially focusing on rainfall prediction. The meteo activity use GRIA to distribute input and output data on a Grid infrastructure. In this context, a scriptable interface that implements the VMC Simple Data Repository has been developed and deployed at the ECMWF node. The new BRIDGE data repository uses the VMC Grid infrastructure to expose BRIDGE resources as Grid services and handles any request to these services. Thus, any user with the right access credentials is able to submit a formula to a BRIDGE service provider via the Portal. The formula is parsed by the

[3] http://www.bridge-grid.eu

BRIDGE client, executed and the results are generated on the fly. Results are delivered in the format choose by the user, and can be instantly retrieved from the Portal.

VMC infrastructure has provided a single point of entry to the BRIDGE applications where any user can submit, monitor and retrieve results from running distributed models part of a GRIA infrastructure. By leveraging the VMC infrastructure with GRIA, BRIDGE applications have became accessible to any user worldwide.

5 Conclusions

The objective of the SIMDAT project was to develop generic-enabling Grid technology that can support improvements in how industry collaborative for the design of complex products. We can see from the evaluation and validation by companies across a range of sectors such as BAE, EADS, and GSK that SIMDAT has achieved this goal. By adopting an iterative user-centred approach structure around themed phases SIMDAT has been able to incrementally deliver capability to industry and help them understand the impact of Grid and now service-oriented infrastructures within their business. Industrial users have moved from initial prototypes, where just getting connected was a challenge, to proof-of-concept demonstrators that show new business operation models that can only be achieved securely and cost-effectively using generic service-oriented technologies, such as those developed in SIMDAT.

The Grid infrastructure software is based on GRIA and has been since the successful demonstration of the aerospace prototype in the Connectivity phase. We have seen a migration towards GRIA's security and management approach within SIMDAT and in other projects. Even in the meteorology application activity, where GRIA was not used for governance of the virtual organisation, many of the principles around distributed management and bi-lateral relationships were adopted following a period of consultancy with WP2 and wider state-of-the art analysis. Regular releases of GRIA have been distributed to SIMDAT partners in accordance with the delivery schedule incrementally including features relevant to application prototypes. The final infrastructure is based on GRIA 5.3 and includes all features agreed with partners and outlined in the description of work. The significant innovations delivered by WP2 include:

- Innovation 1: New modular service-oriented architecture supporting a range of deployment scenarios required by industrial application prototypes incorporating a standardised message structure based on SIMDAT's industrial WSRF profile
- Innovation 2: The first Grid middleware to include a SLA management service within a B2B procurement model supporting monitoring, constraining and billing for service usage from a wide variety of application services
- Innovation 3: Contextualised B2B registry incorporating security policies with resource publication and discovery
- Innovation 4: Grid licensing infrastructure supporting integration of FlexLM with Service Level Agreements
- Innovation 5: Standardised job service supporting integration of multiple resource managers

We have shown that robust software engineering and significant RTD innovation can be incorporated into a collaborative research project ensuring that the maximum impact of the research is achieved and delivered to industry and research communities.

As we enter the demonstration phase we can reflect on a project that has delivered considerable value to the industrial players and software vendors participating in the programme. In 2004, Grid hype was at its peak yet today the brand has little value in business and has returned back to its origins as a technology supporting access to computational resources. SIMDAT's focus on the development of a service-oriented architecture consisting of interoperating technologies rather than just Grid has paid dividends and proved to be a good decision. We have now moved from RTD into a world of opportunities. Industrial partners have the opportunity to use SIMDAT demonstrators to influence IT decision makers so that technologies selected support scenarios that make a real difference to collaborative design. Software vendors have the opportunity to incorporate advanced service-oriented capabilities within products that customers want. Grid technologists have an opportunity to exploit the knowledge gained and to start meeting the challenge of future service-oriented systems. SIMDAT represents the end of an era, as after more than 10 years of RTD industry is finally beginning to exploit the flexibility and diversity of the Internet for business benefit now that service-oriented technologies can be shown to mitigate many of the risks.