SIMDAT

Mike Boniface and
Colin Upstill
University of Southampton IT Innovation Centre

Introduction

In the context of this project, a Grid is defined to be a software system that provides uniform and location independent access to geographically and organizationally dispersed, heterogeneous resources that are persistent and supported. Typically, these shared assets are under different ownership or control. The SIMDAT project[1] is developing generic Grid technology for the solution of complex application problems and demonstrating this technology in several representative industry sectors. Special attention is being paid to security, e.g. where third-party suppliers have need-to-know access to data, and correlation and inference may provide insight into confidential processes. The objective is to accelerate the uptake of Grid technologies in industry and services, provide standardised solutions for capability currently missing, and validate the effectiveness of a Grid in simplifying processes used for the solution of complex, data-centric problems.

The SIMDAT consortium is comprised of leading software and process system developers–IBM, IDESTYLE Technologies, InforSense, Intel, Lion Bioscience, LMS International, MSC Software, NEC, Ontoprise and Oracle; Grid technology specialists–Fraunhofer Institute AIS, Frauenhofer Institute SCAI, IT Innovation, Universitat Karlsruhe, Universite libre de Bruxelles and the University of Southampton; and representatives from strategic industry and service sectors–Audi, BAESystems, DWD, EADS, ESI, EUMETSAT, ECMWF, GlaxoSmithKline, MeteoFrance, Renault, and the UK Met Office. IT Innovation is leading the basic Grid infrastructure level architecture work in SIMDAT and this article will therefore be focused on this aspect of the project rather than the applications.

Grids for complex problem solving in industry

Development of industrial and large-scale products and services poses complex problems. The processes used to develop these products and services typically involve a large number of independent organisational entities at different locations grouped in partnerships and supply chains. Grid is connectivity plus interoperability and is a major contributor to improved collaboration and an enabler of virtual organisations. It has the potential to substantially reduce the complexity of the development process, thereby improving the ability to deal with product complexity.

The heart of the issue is data. Applications and their associated computing power are central to the product development process. Grid technology is needed to connect diverse data sources, to enable flexible, secure and sophisticated levels of collaboration and to make possible the use of powerful knowledge discovery techniques.

Key to seamless data access is the federation of problem-solving environments using grid technology. The federated problem solving-environments will be the major result of SIMDAT. Seven key technology layers have been identified as important to achieving the SIMDAT objectives:

- an integrated grid infrastructure, offering basic services to applications and higher-level layers

[1] http://www.simdat.org/

- transparent access to data repositories on remote Grid sites
- management of Virtual Organizations
- workflow
- ontologies
- integration of analysis services
- knowledge services.

The strategic objectives of SIMDAT are to:

- test and enhance data grid technology for product development and production process design
- develop federated versions of problem-solving environments by leveraging enhanced grid services
- exploit data grids as a basis for distributed knowledge discovery
- promote de facto standards for these enhanced grid technologies across a range of disciplines and sectors
- raise awareness of the advantages of data grids in important industry sectors.

SIMDAT focuses on four exemplar application areas: product design in the automotive, aerospace and pharma industries; and service provision in meteorology. For each of these application areas a challenging problem has been identified that will be solved using Grid technology, e.g. distributed knowledge discovery to enable better understanding of the different Noise, Vibration and Harshness (NVH) behaviour of different designs of cars based on the same platform; Grid technology will allow seamless access to all relevant data for all engineers of the development centers of large multinational car manufacturers.

SIMDAT Architecture

All application sectors deploy existing problem-solving environments for product and process design. Each application activity in SIMDAT is integrating Grid middleware into existing applications to provide a demonstration of distributed, collaborative work in complex problem solving. The vendors of these environments require an acceptable level of stability in middleware technology before it will be adopted with their products and delivered to customers. Well-designed and accepted standards are essential for technology uptake.

Examining Grid infrastructure state-of-the-art, it is clear that even the core technology, which underpins higher-level services such as resource and execution management, is still evolving. In the future, core features should be part of a standards-compliant architecture, so application developers can use them more easily and so they can choose between different interoperable Grid implementations.

The Open Grid Services Architecture[2] (OGSA) represents an evolution towards a Grid system architecture based on Web services concepts and technologies. OGSA Profiles for various higher-level functions are beginning to be developed but there are certainly no OGSA compliant grid implementations. Even the underlying proposed standard WS-Resource Framework[3] (WS-RF) is still somewhat controversial and has yet to prove its value. Therefore, the challenge of standardising the Grid programming model and associated management services is still unfulfilled.

[2] http://www.globus.org/ogsa/

[3] http://www.globus.org/wsrf/

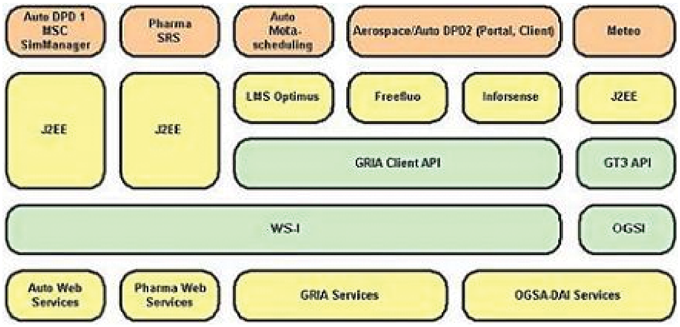

Figure 1: Initial SIMDAT Architecture

SIMDAT has adopted a pragmatic approach for using existing Grid infrastructure and Web Service technologies. The application sectors faced the challenge of selecting Grid technologies that best fitted their scenarios, even if they did not provide all of the necessary functionality. Key among these are GRIA[4] (open source Grid middleware developed by IT Innovation to enable commercial use of the Grid in a secure, interoperable and flexible manner), Globus,[5] and J2EE[6] portals. The application activities concluded that in the short-term, unless and until a standardised Grid programming model is agreed upon, new developments should be based on Web Service standards such as WS-I.[7] GRIA emerged as a core technology to support collaborative work in the SIMDAT aerospace and automotive activities because of its availability, its adherence to WS-I, and its explicit support for B2B collaborations. The SIMDAT meteorology activity is developing a Data Grid based on Open Grid Services Architecture Data Access and Integration[8] (OGSA-DAI) and Web Service technology, while the SIMDAT pharmaceutical activity is deploying a Web Services Grid leveraging E2E (end-to-end) security component developed during the GEMSS project.[9] The initial architecture (Figure 1) shows how WS-I can provide a common API for distributed services, but it does not currently meet Grid infrastructure requirements for providing a standardised approach for managing stateful resources, as proposed by WS-RF.

[4] http://www.gria.org/
[5] http://www.globus.org/
[6] http://java.sun.com/j2ee/index.jsp/
[7] http://www.ws-i.org/
[8] http://www.ogsadai.org.uk/
[9] http://www.gemss.de/

SIMDAT provides application sectors with a Grid infrastructure roadmap that tracks the rapidly changing Grid landscape. During 2005, the situation has evolved significantly as GRIA continues to be developed and as technologies such as GT4[10] and gLite[11] emerge. Following the recent delivery of the first SIMDAT prototypes, the application activities are feeding back lessons learned, which will be factored into the roadmap and into the Grid technologies. In the longer term SIMDAT aims to achieve interoperability between different Grid infrastructures such as GRIA and GT4, with a Grid service API based on WS-RF, although the level of compliance may differ between implementations.

[10] http://www.globus.org/toolkit/
[11] http://glite.web.cern.ch/glite/

Aerospace Case Study

The aerospace industry deals with highly complex products that have data creation, management and curation requirements that span hundreds of collaborating organisations over a 50-year lifecycle. Partners in a product team need to collectively manage thousands of interrelated processes and this leads the industry to expend considerable time and effort in the access, transmission, control, translation and sharing of data.

The primary focus of the aerospace activity in SIMDAT is the development and deployment of existing and emerging Grid technologies and concepts to enhance the collaborative engineering of sophisticated products. The improvement in the ability to handle complex problems is not delivered simply through the connectivity that Grid offers, but through the deployment of industry-strength middleware and advanced ontology-based techniques to radically improve the efficiency of the data exchange, both between applications and between organisations

The initial aerospace deployment simulates the multi-disciplinary collaborative configuration design of a low-noise, high-lift landing system. The scenario is typical of sub-system design problems in the context of future-concept, unmanned cargo vehicles that need to use airfields in noise-sensitive locations. The scenario has been designed to show how Grid technologies can support the aggregation of distributed capabilities operating across organisational boundaries.

Figure 2: Pan-European deployment of Grid technologies in Aerospace

The deployment of the aerospace prototype demonstrates how Grid technologies can support pan-European inter-Enterprise collaborative development of complex products (Figure 2). Each organisation within the aerospace deployment operates as a GRIA service provider offering specialised engineering services such as optimisation (University of Southampton), parameterised CAD generation (University of Southampton), aerodynamics (BAE SYSTEMS) and aero-acoustics (EADS). GRIA's explicit business process support for dynamic, bi-lateral QoS agreements allows project managers at aerospace companies to create inter-Enterprise multidisciplinary design teams in a secure, managed, auditable and accountable environment.

GRIA has been significantly enhanced to support the aerospace application scenario through integration with other key Grid technologies. OGSA-DAI WS-I has been integrated with GRIA to provide distributed data access for relational data and simulation files. The GRIA OGSA-DAI service provides security and enforces a business model for managing distributed data resources. A GRIA workflow service has been developed, based on IT Innovation's open source workflow enactor, Freefluo.[12] This allows aerospace engineers to publish workflows as services that can be executed by distributed clients.

[12] http://freefluo.sourceforge.net/

The Grid programming environment is provided by the Taverna[13] workbench, which has been enhanced to support GRIA's business process and WS-I Basic Security Profile. Aerospace engineers integrate legacy applications as Grid services using GRIA and compose these applications into workflows using Taverna. The workflows can then be published to GRIA's workflow enactment service, allowing clients to compose hierarchical distributed workflows that cross organisation boundaries. At the lowest level, the aerospace workflows consist of a simple computational sequence of meshing, solving and post-processing. At higher-levels, the workflows are much more complex. For example, the design workflow (Figure 3) explores the design space by creating a design of experiments (DoE) from a given input specification and iteratively calculating the results for each design point in the DoE. The workflow implementation invokes optimisation, CAD generation and compute workflows, along with staging input and output data in a design database accessed through OGSA-DAI.

[13] http://taverna.sourceforge.net/

Programming Grid-based engineering workflows is a complex problem that can only be currently achieved by expert aerospace engineers. Workflow management and accessibility by less experienced users is a key requirement for aerospace companies. In the first SIMDAT prototype, an aerospace design portal has been developed and deployed using the GridSphere[14] portal framework to provide a simple interface to the design workflow. However, future work will focus on workflow management and integration with existing aerospace problem solving environments.

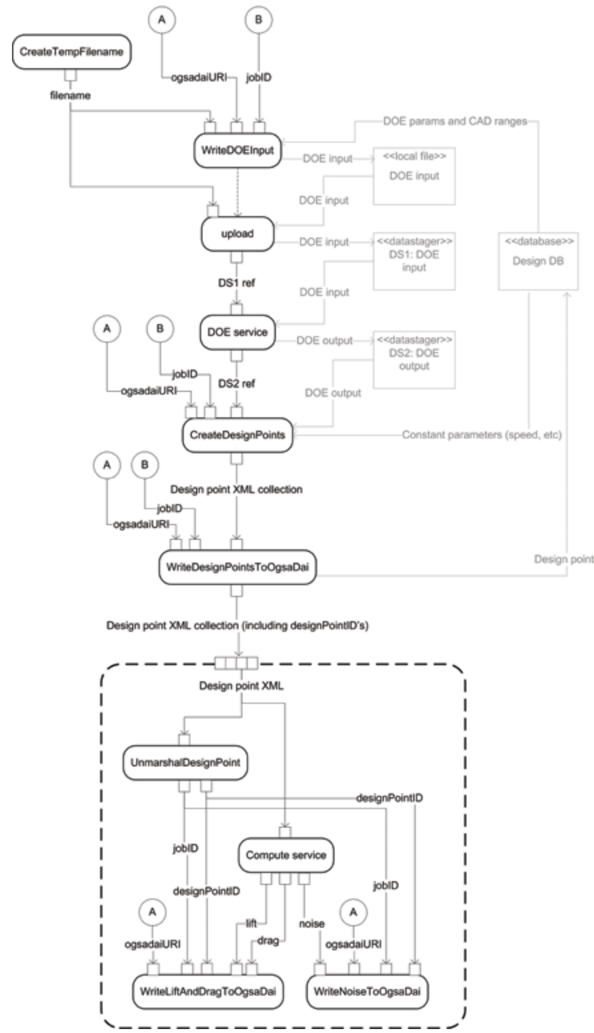

Figure 3: Aerospace design workflow

Conclusion

SIMDAT will change the way we design artefacts from drugs to aeroplanes. It will enable chemists and engineers to access geometry information and simulation results in a transparent way. Actors in supply chains will have immediate access to modifications, which can be securely disclosed to them for their own tasks. For the first time, not only the shape of a product but also its functional characteristics will be part of an integrated software environment. The underlying set of distributed federated databases will be viewable as a one logical whole, enabling the use of powerful knowledge discovery techniques and transforming our ability to solve complex problems.

Acknowledgement
SIMDAT has received research funding from the European Commission under the Information Society Technologies Programme(IST), contract number IST-2004-511438.

[14] http://www.gridsphere.org/gridsphere/gridsphere/