Business Experiments in Grid

www.beingrid.eu

With Grid middleware stacks reaching maturity, industrial uptake of Grid solutions is paramount for the European economy. The 74-partner BEinGRID project includes top European Grid, IT and business experts, supporting two waves of business experiments as they pilot Grid solutions in diverse market sectors. Instrumental in the design and management of the project, EPCC also leads Data Management support to the business experiments.

Of EPCC in Europe

As part of its mission, EPCC is committed to transferring skills and knowledge to UK and European industry. Having in the past been a leading partner in the successful HPC-TTN and other related European projects, EPCC was perfectly positioned to design and contribute to the objectives of the Business Experiments in Grid (BEinGRID) project. Taking the lead from our past involvement in MPI, our data management expertise appears to be key in modern-day Grids.

Of Grids in Europe

Grid research has enjoyed good uptake in the academic community. The Globus Toolkit, OGSA-DAI, gLite, Unicore and other such Grid stacks have reached a level of maturity through development effort, but more importantly through the engagement of an enthusiastic user community in Europe and further abroad.

Having funded Grid research during the past 5 years, the European Commission now sees the Grid as a key building-block towards sustainable infrastructures. The key component missing at this stage is business penetration. In order to develop a market for a new technology, such as the Grid, it is essential to make the transition from an early market, dominated by a few visionary customers, to a mainstream market, dominated by a large block of customers who are predominantly pragmatists in orientation.

Objectives

The strategic objectives of BEinGRID are:

- To understand the requirements for Grid use in the commercial environment, involving software vendors, IT integrators, service providers and end-users.
- To enable and validate the adoption of Grid technologies by business.
- To design and build a Grid toolset repository with components and solutions based on the main Grid software distributions and basic Web Service specifications.
- To develop and deploy a critical mass of Grid-enabled pilots, embracing a broad spectrum of economic sectors with different needs and requirements.

BEinGRID Bridge

BEinGRID aims to bridge the chasm between early adopters and early majority through a series of business experiments that employ mature Grid technologies to add value to European businesses. In doing so, BEinGRID will establish effective routes to foster the adoption of Grid technologies across the EU and to stimulate research into innovative business models.

A wave of eighteen business experiments, across a broad spectrum of European business sectors (entertainment, financial, industrial, chemistry, gaming, retail, textile, etc), will set the scene for the uptake of the Grid in industry. A wave of further business experiments (selected from a 2M€ competitive call) will validate and consolidating the results of the project.

The overall result of the project will be a collection of key middleware components, Grid solutions and successful case studies resulting from the real-world pilots and the best-practice guidelines derived from the Grid pilot experiences. The creation of a toolset repository that gathers high-level services, new tools and innovative Grid application solutions will result in a "Grid market place" enabling individuals and organisations to create, provide and use Grid technologies to meet their business challenges.

Taking the Lead

One of the five managing partners in this 74-partner project, EPCC leads the Data Management support to the Business Experiments. This follows quite naturally given our expertise in OGSA-DAI, our involvement in the Globus Alliance and our long-standing commitment to transfer knowledge to Small and Medium-sized Enterprises. Early indications suggest that at least two thirds of the Business Experiments have requirements for Data Management components to:

- Automate and rationalise their workflows
- Simplify access to heterogeneous and/or distributed data
- Enhance the performance of their distributed applications

Our involvement in BEinGRID demonstrates the transfer of our OGSA-DAI research work into business. The BEinGRID project results will validate the claim that Data Management is to the Grid what MPI was to parallel computing.

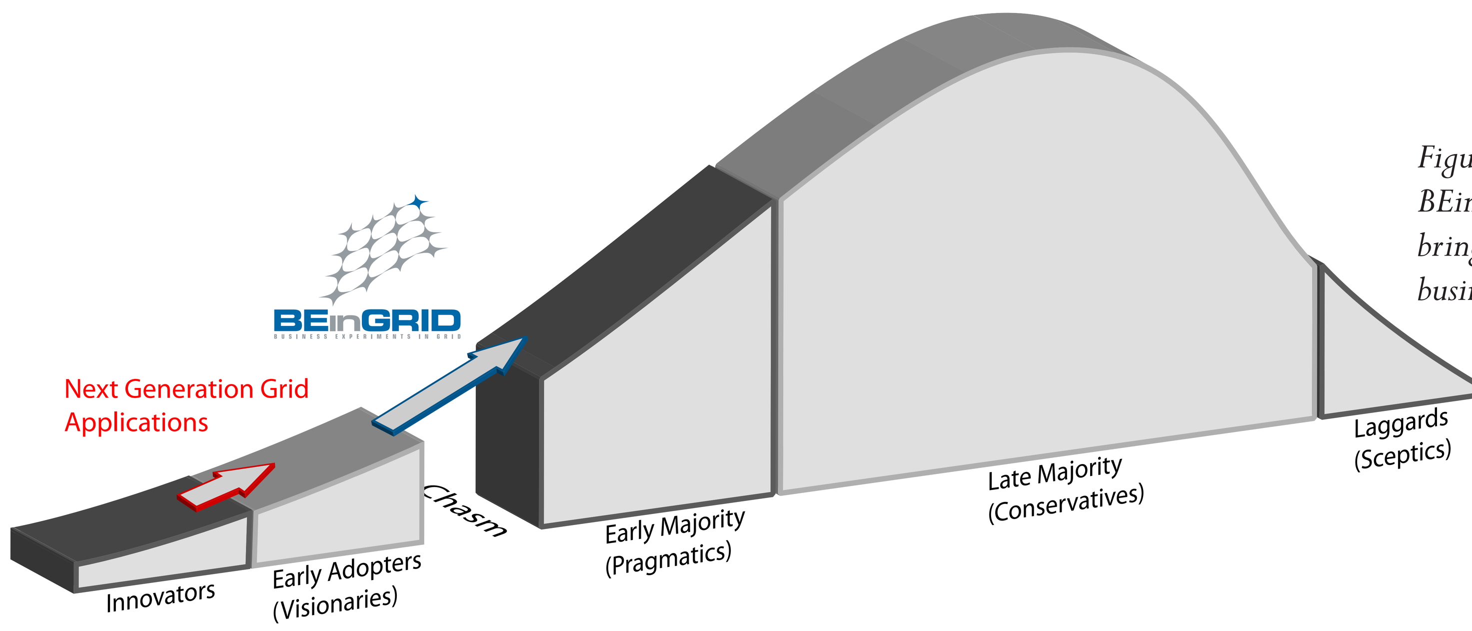

Figure 1: How BEinGRID will bring the Grid to business.

NextGRID: Towards a Next Generation Grid Architecture

www.nextgrid.org

In the NextGRID project, the University of Edinburgh is leading a €15million, 21 partner pan-European project to develop the next generation of Grid architectures.

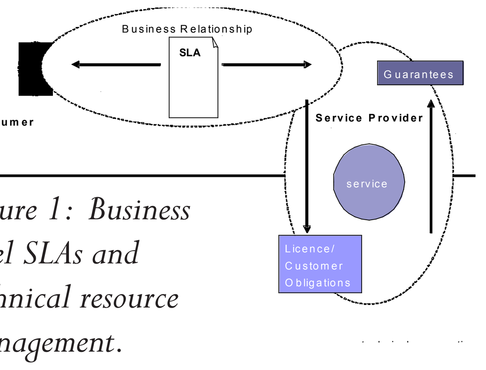

Figure 1: Business level SLAs and technical resource management.

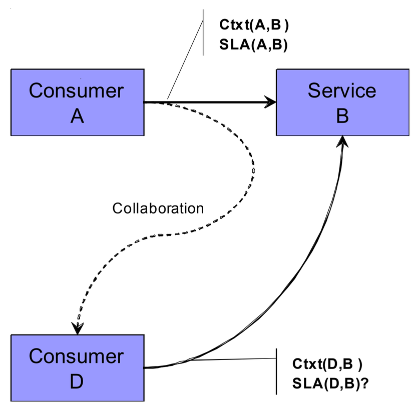

Figure 2: Resource Sharing (Consumer Federation).

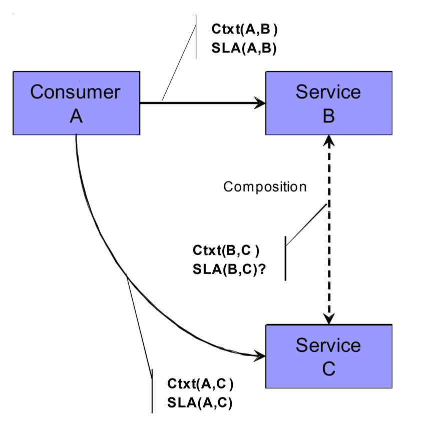

Figure 3: Resource Orchestration.

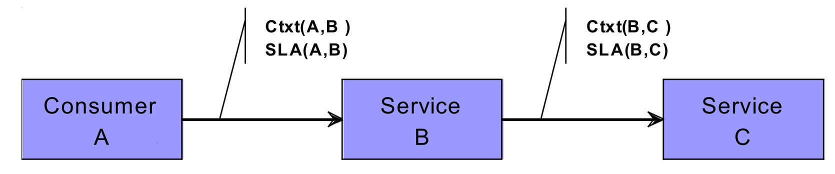

Figure 4: Resource Encapsulation.

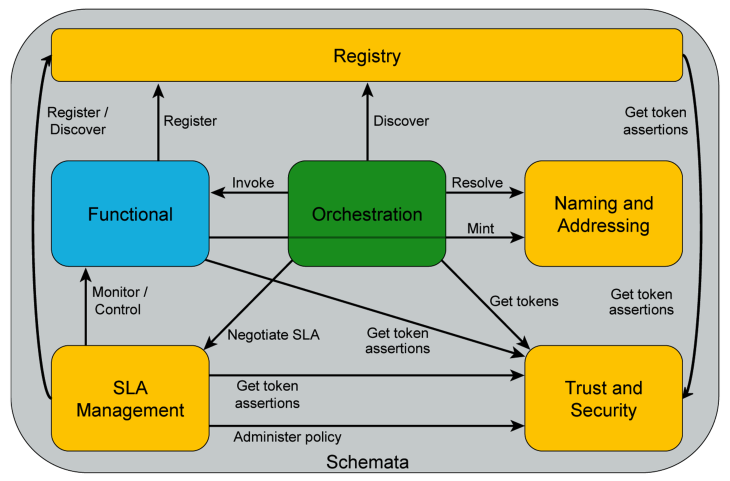

Figure 5: Overview of NextGRID Component Model and basic interactions.

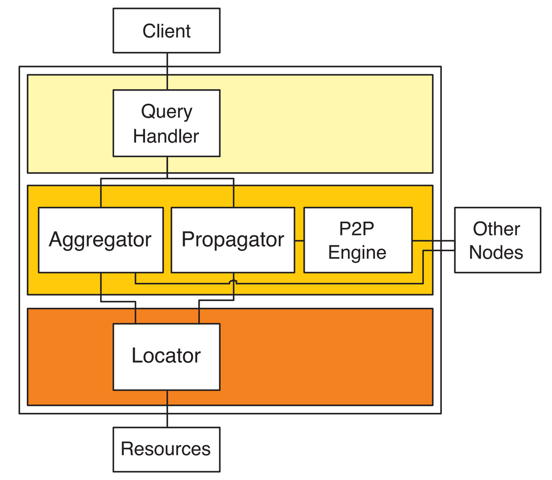

Figure 6: Architecture of a NextGRID Information Discovery Framework node.

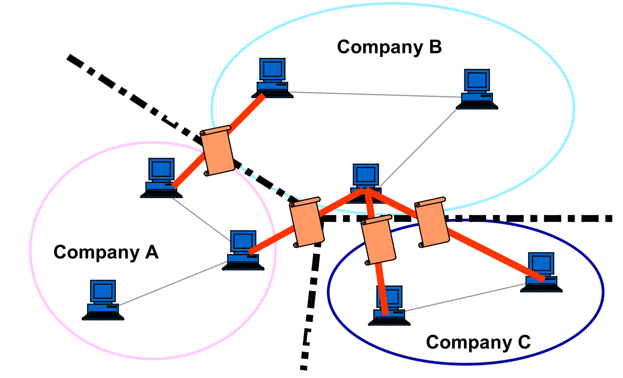

Figure 7: The Information Discovery Framework can bridge queries across organisational boundaries.

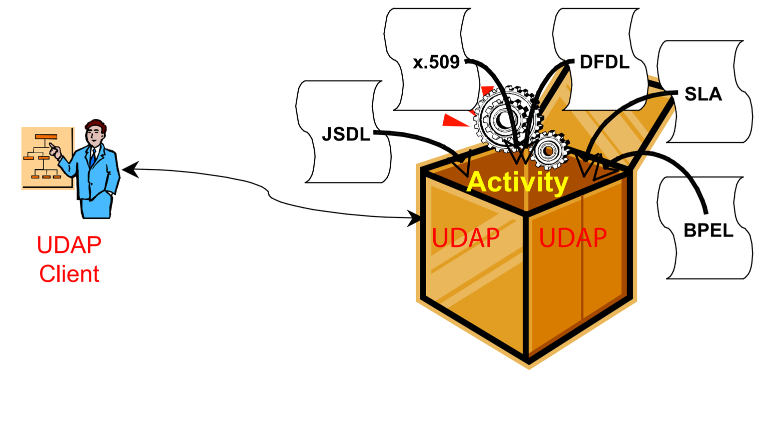

Figure 8: UDAP provides contextual information on activities.

EPCC coordinates the 21 partner NextGRID project, which seeks to ensure that Europe is a world leader in the next generation of Grid technology. The three-year project envisions the development of an architecture for Next Generation Grids which will enable their widespread use by research, industry and the ordinary citizen thus creating a dynamic marketplace for new services and products. The NextGRID vision is of future grids, which are economically viable; in which new and existing business models are possible; in which development, deployment and maintenance are easy; and in which the provisions for security and privacy give confidence to businesses, consumers and the public.

The full realisation of the NextGRID vision will lead to the emergence of Next Generation Grids by the end of the decade.

Requirements of Next Generation Grids

There are eight fundamental requirements driving the design of NextGRID. These grew out of the experience of both technology and application experiments undertaken by the project; they have been used to draw together the architectural design process and provide threads of continuity through the project as a whole. The requirements are:

R1 – Flexible Business Models
R2 – Specific Quality of Service Terms
R3 – Dynamic Security
R4 – Dynamic Composition
R5 – Economic Sustainability
R6 – Privacy
R7 – Facilitated Management
R8 – Interactive Support

From Requirements to Principles

The primary architectural principles of NextGRID fall into three categories.

Service Level Agreements

First, the Service Level Agreement is central to the conceptual model of NextGRID and therefore forms a key aspect of the underlying NextGRID infrastructure (Figure 1). All interactions in NextGRID are predicated by an SLA, dynamically created, and aimed at ensuring that the relationship between provider and consumer is well defined and understood. Follow up capabilities allow for monitoring, violation management, and reputation creation. The SLA-based approach to all non-functional (as well as functional) aspects of NextGRID provides a uniform framework for the management and operation of all QoS aspects of NextGRID, eg performance, security, provenance management, adherence to privacy regulations, etc.

Dynamic Grid infrastructure

Second, as a dynamic Grid infrastructure, NextGRID provides extensive capabilities for service construction and composition (Figures 2, 3 and 4), including traditional interface composition, various forms of workflow-enabled orchestration, and support for dynamic extension of services capabilities. NextGRID also includes requirements and guidance on how an algebra for service composition might be constructed.

Minimal set of capabilities

Third, that all services operating in a NextGRID environment can expect to find, but are not required to exploit, a minimal level of capabilities either available in the environment or exhibited by peer services. These capabilities are further refined as communication protocols and languages, behavioural interfaces available on all services, support services from the environments, and a common infrastructure of underlying schemas.

Architectural decomposition

NextGRID tackled the problem of decomposing its architecture through the regular analysis of the various iterations of the architecture. Here we have used the results of this process to present a conceptual view of NextGRID as it has evolved. This is a more complex task than it appears at first. Frequently systems can be decomposed into a "layered" architecture, where each layer communicates only with the adjacent layers. However, the increasing complexity of Grid systems has resulted in the erosion of this simple approach, with some aspects of the system (e.g. security and provenance) spanning all layers of the architecture. Figure 5 depicts this decomposition and some of the interactions expected between components, which can be decomposed into four concepts: Schemas, Support Systems, Functional Systems and Orchestrators.

An Information Discovery Framework

In the NextGRID project, The University of Edinburgh has been developing a more advanced implementation of the communication algorithm known as the Echo pattern. This comprises a collaborative network of autonomous and decentralised discovery nodes that act both as entry points and resolution points for data and information queries. Data and information queries for the discovery mechanism can range from resource discovery to data resolution, and include metadata discovery about both resources and data within a Grid.

The discovery nodes interface with local Grid resources using custom components, called locators, which encapsulate search algorithms and access methods for specific resources, allowing any type of information of data to be accessed through this system, if somebody implements a locator object for that information or resource. Queries can be submitted to any discovery node within the system, as discovery nodes forward (propagate) queries to other nodes to provide adequate search coverage for any given query.

Propagators handle the forwarding of queries to other discovery nodes, using both locally defined and user supplied propagation rules. The aggregator receives results from both the local locator component, and from other discovery nodes the query has been propagated to, and is responsible for aggregating those results in some way. This combination of propagators and aggregators allows complex queries to be executed across organisational boundaries (Figure 7).

A Unified Approach to Resource Description

The University of Edinburgh has also contributed to a new area of research which is studying how to build unified models of compute and data resources. This has led to the proposal of a Universal Dynamic Activity Packet Model. Here an activity is defined as a unit of work, e.g. a job, a data access operation, an application execution or a Web Service invocation. It is important to consider all that there is to know about an activity: its properties (i.e. its behaviour), attributes and their values (i.e. what it has), dependencies (on data and other activities) and contextual information (e.g. topical information, SLAs).

The heterogeneity of information schema cannot be avoided but the different schemes for handling various information fragments do inhibit interoperability between Grid components and different framework implementations. Therefore the Universal Dynamic Activity Package (UDAP) is a package that contains all of the information associated with an activity on the Grid: it is a framework for the management of all activity information.

The UDAP package can contain any information about an activity, regardless of the schema used to present that information. The values of the activity information in the UDAP package can be changed, and can be updated or appended to reflect the past, present, and future state of the activity.

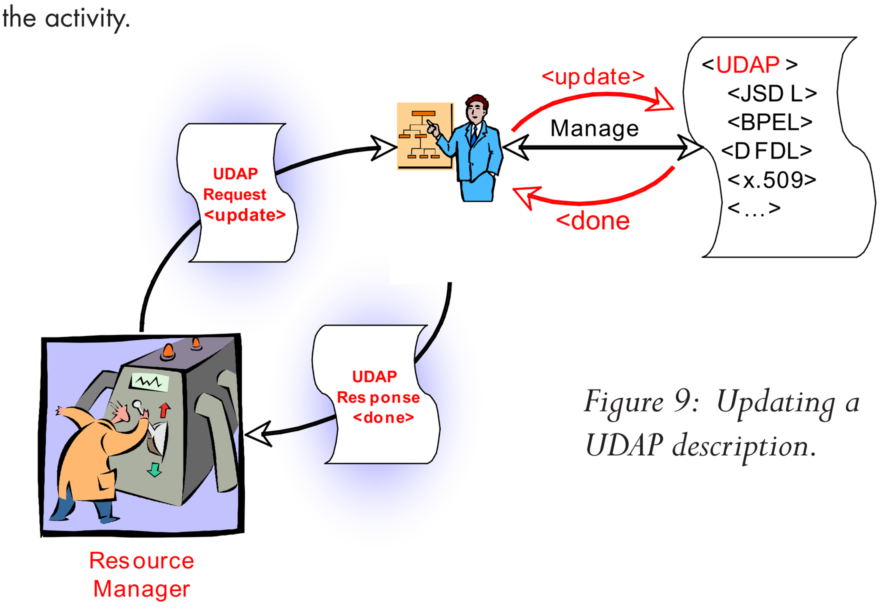

Figure 9: Updating a UDAP description.