Reasoning Services for the Semantic Grid

Oscar Corcho
The University of Manchester.
Kilburn Building, Oxford Road, Manchester, M13 9PL, United Kingdom.
Oscar.Corcho@manchester.ac.uk

1 Background: The Semantic Grid and S-OGSA

The Grid aims to support secure, flexible and coordinated resource sharing through providing a middleware platform for advanced distributing computing. Consequently, the Grid's infrastructural machinery aims to allow collections of any kind of resources—computing, storage, data sets, digital libraries, scientific instruments, people, etc—to easily form Virtual Organisations (VOs) that cross organisational boundaries in order to work together to solve a problem. A Grid depends on understanding the available resources, their capabilities, how to assemble them and how to best exploit them. Thus Grid middleware and the Grid applications they support thrive on the metadata that describes resources in all their forms, the VOs, the policies that drive then and so on, together with the knowledge to apply that metadata intelligently.

The Semantic Grid is a recent initiative to systematically expose semantically rich information associated with Grid resources to build more intelligent Grid services [1]. The idea is to make structured semantic descriptions real and visible first class citizens with an associated identity and behaviour. We can then define mechanisms for their creation and management and protocols for their processing, exchange and customisation. We can separate these issues from both the languages used to encode the descriptions (from natural language text right through to logical-based assertions) and the structure and content of the descriptions themselves, which may vary from application to application.

In practice, work on Semantic Grids has primarily meant introducing technologies from the Semantic Web [2] to the Grid. The background knowledge and vocabulary of a domain can be captured in *ontologies* – machine processable models of concepts, their interrelationships and their constraints. Metadata labels Grid resources and entities with concepts, for example describing a job submission in terms of memory requirements and quality of service or a data file in terms of its logical contents. Rules and classification-based automatic inference mechanisms generate new metadata based on logical reasoning, for example describing the rules for membership of a VO and reasoning that a potential member's credentials are satisfactory.

1.1 S-OGSA

The Grid has a reference architecture, called OGSA (Open Grid Service Architecture) [4], which defines a core set of capabilities and behaviours for Grid systems. However, currently the Semantic Grid lacks a Reference Architecture or any kind of systematic framework for designing Semantic Grid components or applications. In the context of the project OntoGrid (http://www.ontogrid.net/), which the project RSSGRID is associated to, an architecture called S-OGSA has been recently proposed [6]. S-OGSA extends OGSA by explicitly defining a lightweight mechanism that allows for the explicit use of semantics and defining the associated knowledge services to support a spectrum of service capabilities.

S-OGSA has three main aspects: the *model* (the elements that it is composed of and its interrelationships), the *capabilities* (the services needed to deal with such components) and the *mechanisms* (the elements that will enable communication when deploying the architecture in an application).

S-OGSA Model. Although there is no standardized overall model of the Grid and its basic concepts, there are project specific models [3,5], capability focused models emerging from the Global Grid Froum such as CIM (Common Information Model) and JSDL (Job Submission Description Language), and a vocabulary associated with OGSA. S-OGSA introduces the notion of Semantics into the model of the Grid defining *Grid Entities*, *Knowledge Entities* (e.g. ontologies, rules, text), *Semantic Bindings* between these two for a Grid Entity to become *Semantic Grid Entities*. Semantic Bindings are (possibly temporary) metadata assertions on Grid entities and are Grid resources with their own identity, manageability features and metadata.

S-OGSA Capabilities. S-OGSA is a mixed economy of these semantically enabled and disabled services. We add to the set of capabilities that Grid middleware should provide (as defined by OGSA) the Semantic Provisioning Services and Semantically Aware Grid Services (Figure 1).

Semantic Provisioning Services dynamically provision an application with semantic grid entities in the same way a data grid provisions an application with data. The services support the creation, storage, update, removal and access of different forms of Knowledge Entities and Semantic Bindings. *Ontology services* store and provide access to the conceptual models representing knowledge; *reasoning services* support computational reasoning with those conceptual models; *metadata services* store and provide access to semantic bindings and the *annotation services* generate metadata from different types of information sources, like databases, services and provenance data.

Semantically Aware Grid Services exploit knowledge technologies to deliver their functionality, for example metadata aware authentication of a *given identity* by a VO Manager service or execution of a *search request* over entries in a semantically enhanced resource catalogue. Sharing this knowledge brings flexibility to components and increases interoperability.

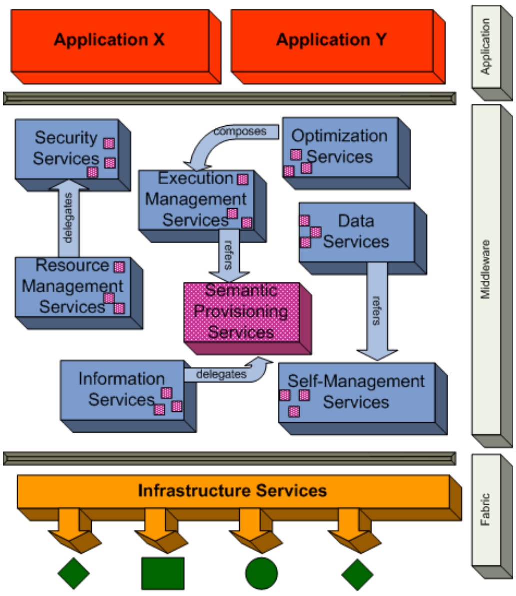

Figure 1: The S-OGSA semantic provisioning services positioned in the OGSA services

S-OGSA Mechanisms. All this conceptual work grounds out into concrete Grid modeling elements. Specifically, S-OGSA OntoGrid's implementation builds on Globus Toolkit 4, grounding to WS-RF (Web Services Resource Framework), incorporating S-OGSA entities into the CIM Resource Model. These groundings demonstrate how semantic bindings can be incarnated in the form of stateful Grid resources, including bindings on inter-service messages that are not strictly Grid entities but nonetheless useful.

2 Reasoning Services in S-OGSA

The project RSSGRID is specifically focused on the development of reasoning services, which are part of the semantic provisioning services described in the S-OGSA architecture. RSSGRID considers the following two situations in the adaptation and refinement process of reasoning infrastructure:

a) **Knowledge in the Semantic Grid is distributed among Grid resources**. Reasoning services will have to gather knowledge distributed in many sources and process it to fulfil a task.

b) **Reasoning in the Semantic Grid can be performed collaboratively by distributed Grid resources**. In order to allow a good scaling up of reasoning services, they will surely have to distribute their reasoning workload among distributed resources so that they can operate more efficiently.

Taking into account these situations (knowledge and reasoning distribution), common problems of Grid computing will also appear in the Semantic Grid. The most relevant problems that we will deal with are heterogeneity, scalability and adaptability:

a) Regarding **heterogeneity**, we will deal with knowledge distributed among Semantic Grid resources, which will be heterogeneous in different aspects: it will be represented in different languages and formats (such as the Semantic Web languages RDF – Resource Description Framework –, RDF Schema, or OWL – Web Ontology Language –), it will consider different vocabularies or ontologies, etc. Some work has been already done to deal with distributed knowledge in the context of the Semantic Web, regarding knowledge localization and format heterogeneity. However, the current solutions are not mature enough yet.

b) Regarding **scalability**, we will improve the efficiency of complex reasoning processed by distributing their execution in several resources. We must note that the efficiency gaining will not be linear with regard to the number of resources used, since we will have to consider the overhead inherent to the dispatching of reasoning subprocesses, latencies, etc. Work has been also done in the past in this issue (from Problem Solving Methods to Internet reasoning services, and from domain oriented solutions to generic ones), which will be improved and adapted to the Grid architecture.

c) Regarding **adaptability**, we will consider the possibility of failures of resources and of their reasoning processes. This will deal to problems when trying to ensure that the results delivered by a reasoning process are complete and/or correct, whether we are considering all the relevant content, etc. RSSGRID's innovation will go in the direction of facing failures in reasoning resources and processes with guarantees. Finally, we will provide information about the quality of service by a continuous process of evaluation (which will be also innovative itself).

Acknowledgements

I would like to give thanks to the OntoGrid team at the University of Manchester, with whom I have worked on the proposal of the S-OGSA reference architecture. Besides, this work is supported by the Marie Curie Intra-European Fellowship "RSSGRID" (contract number 006668) and is being deployed in the context of the EU FP6 OntoGrid project (STREP 511513).

References

1. C.A. Goble, D. De Roure, N. R. Shadbolt and A.A. Fernandes *Enhancing Services and Applications with Knowledge and Semantics* in The Grid 2 Blueprint for a New Computing Infrastructure Second Edition eds. Ian Foster and Carl Kesselman, 2003, Morgan Kaufman, November 2003
2. James Hendler *Science and the Semantic We*b Science 299: 520-521, 2003
3. L. Pouchard, L. Cinquini, B. Drach, et al., *An Ontology for Scientific Information in a Grid Environment: the Earth System Grid*, CCGrid 2003 (Symposium on Cluster Computing and the Grid), Tokyo, Japan, May 12-15, 2003.
4. Foster, H. Kishimoto, A. Savva, D. Berry, A. Djaoui, A. Grimshaw, B. Horn, F. Maciel, F. Siebenlist, R. Subramaniam, J. Treadwell, J. Von Reich, *The Open Grid Services Architecture,* http://www.ggf.org/documents/GFD.30, 2005.
5. N. Sharman, N. Alpdemir, J. Ferris, M. Greenwood, P. Li and C. Wroe, *The ^{my}Grid Information Model*, Proceedings of UK e-science All Hands Meeting, 2004, available from http://www.mygrid.org.uk
6. Kotsiopoulos I, Alper P, Bechhofer S, Corcho O, Goble C, Kuo D, Missier P, Pérez-Hernández MS. *Towards a Semantic Grid architecture.* PPAM'05 Workshop on Knowledge and Data Management in Grids (1st CoreGrid workshop). Poznan, Poland. September 2005