## S-OGSA as a Reference Architecture for OntoGrid and for the Semantic Grid

Pinar Alper, Oscar Corcho, Ioannis Kotsiopoulos, Paolo Missier, Sean Bechhofer, Dean Kuo, Carole Goble.
University of Manchester, UK

### What is a Semantic Grid?

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; for example a model of a VO [3]. 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.

Currently the Semantic Grid lacks a Reference Architecture or any kind of systematic framework for designing Semantic Grid components or applications.. In this paper we describe Semantic-OGSA (a.k.a. S-OGSA), which is one of the early results of the EU-IST project OntoGrid (*http://www.ontogrid.net/*). S-OGSA is proposed as a Reference Architecture for the project and has been created with the aim to become also a reference framework for the Semantic Grid.

### A principled approach to Semantic-OGSA

OGSA aims to define a core set of capabilities and behaviours for Grid systems [4]. OntoGrid extends OGSA by explicitly defining a lightweight mechanism that will allow for the explicit use of semantics and defining the associated knowledge services to support a spectrum of service capabilities. S-OGSA is guided by seven design principles (see Figure 1), which have emerged from our observations on fundamental issues in Semantic Grid research [6]:

1. **Parsimony**: the architectural framework should be as lightweight as necessary, minimise the impact on legacy Grid infrastructure and tooling, and not dictate the definition of the contents of the descriptions – these will be application or middleware dependent. We believe this crucial to adoption of our approach.
2. **Extensibility**: rather than define a complete and generic architecture, define an extensible and customisable one. Generality is the enemy of applicability.
3. **Uniformity**: Semantic Grids are Grids, so any S-OGSA entity included in the architecture will be OGSA-compliant. OGSA compliance brings about the following expectations:
    a. Similar to the Grid resources they are associated with, knowledge and metadata should exhibit manageability aspects. Semantic descriptions could have state, and have soft state characteristics – they have a lifetime and may change during their life.
    b. During their lifetime, Grid entities can incrementally acquire, lose and reacquire explicit semantics.
    c. Knowledge services in S-OGSA are OGSA-compliant Grid services. Moreover, as metadata stores and ontology services are just special kinds of data services, we have adopted the OGSA-DAI specification for their deployment and can potentially exploit other data grid capabilities.
    d. S-OGSA must encapsulate both stateless and stateful Grid services, as OGSA does.

4. **Diversity**: a dynamic ecosystem of Grid services ranging over a spectrum of semantic capabilities will coexist at any one time. Grid entities do not *need* to be Semantic Grid entities. Semantic capability may be possible for some Grid resources all of the time, and maybe all Grid resources some of the time, not all resources all of the time. Entities in the Semantic Grid are thus classified as:
   a. *Ignorant* of the associated semantics of another entity.
   b. *Aware* that another entity has explicit associated semantics but *incapable* of processing it.
   c. *Aware* that another entity has explicit associated semantics and *capable* of processing it, partially or completely.

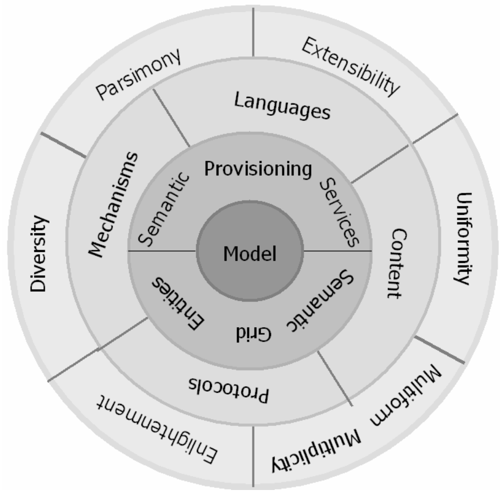


Figure 1: Semantic-OGSA. The outer ring gives the principles; the next ring the main architecture aspects; the next the two major classes of OntoGrid service capability and at the centre is the S-OGSA model.

5. **Multiform + Multiplicity**: the same semantic description may be captured in many representational forms (text, logic, ontology, rule) and any resource's property may have many different descriptions.
6. **Enlightenment**: services should have a straightforward migration path that enables them to become knowledgeable, and minimise the cost of doing so. Cost involved in the migration to the Semantic Grid must be minimised in order to improve the impact and uptake of Semantic Grid, and to take advantage of current tooling and services. Thus:
   a. S-OGSA should have minimal impact on adding explicit semantics to current Grid entity interfaces or on Grid services that are ignorant of Semantic Grid entities;
   b. Grid entities should not break if they can consume and process the Grid entity but cannot consume and process its associated semantics;
   c. If a Grid entity understands only part of the knowledge it consumes it should be able to use it as a best effort.
7. **Conceptual**: Our aim is to develop S-OGSA as a conceptual architecture. That is, it should apply equally with groundings to WSRF [7], to WSDM [8], to Microsoft's WS-Management specification [9] etc.

**S-OGSA: Models, Capabilities and Mechanisms**

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 delivery when deploying the architecture in an application, grounded to a Grid platform).

***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 GGF such as CIM[1], DFDL[2] and JSDL[3], and a glossary associated with OGSA [10].

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 to include the Semantic Provisioning Services and Semantically Aware Grid Services (Figure 2).

---

[1] Common Information Model (CIM) https://forge.gridforum.org/projects/cgs-wg/
[2] Data Formal Definition Language (DFDL) http://forge.gridforum.org/projects/dfdl-wg
[3] Job Submission Description Language https://forge.gridforum.org/projects/jsdl-wg/

***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. These services are classified into two major categories, namely Knowledge Provisioning Services and Semantic Binding Provisioning Services, reflecting the S-OGSA model.

> *Knowledge provisioning services* include *ontology services*, which are in charge of the storage and access to the conceptual models of representing knowledge, and *reasoning services*, in charge of computational reasoning with those conceptual models.
>
> *Semantic binding provisioning services* include *metadata services*, in charge of the storage and access to semantic bindings, normally considered as sets of ontology instances, and *annotation services*, in charge of generating metadata from different types of information sources, like documents, databases, provenance information, credentials, etc.

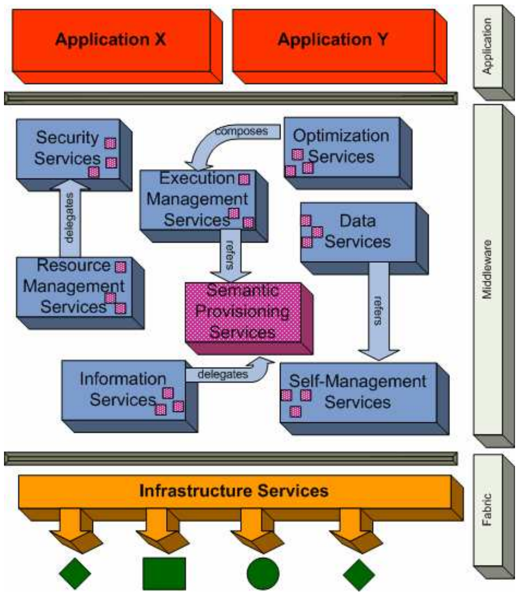


Figure 2: The S-OGSA semantic provisioning services positioned in the OGSA services

***Semantically Aware Grid Services*** exploit knowledge technologies to deliver their functionality. They are able to consume semantics bindings and take actions based on knowledge and metadata. Examples of such actions are:

- Metadata aware authorization of a *given identity* by a **VO Manager service**;
- Execution of a *search request* over entries in a **semantic resource catalogue**;
- Incorporation of a *new concept* in to an ontology hosted by an **ontology service**;

***S-OGSA Mechanisms.*** The conceptual work described above has to be first *represented* with concrete Grid modeling elements then *grounded* into infrastructure specific resource representations and finally *accessed* through patterns we lay out as part of S-OGSA mechanisms.

**Representation** of the S-OGSA model is done at a level that is independent of any Grid implementation system, as required by the *parsimony principle.* For representation we have chosen the Common Information Model (CIM), which is currently undergoing standardization through GGF. Based on S-OGSA's uniformity principle, it is crucial ***to treat Knowledge Entities and Semantic Bindings as Grid Resources*** within S-OGSA's representation. These entitites are modelled as first class citizens in S-OGSA. And consequently, when grounded to a particular infrastructure this status is reflected in the technology and paradigm specific layers. This is because we want to explicate the existence of semantics at appropriate abstraction levels: viz. the abstraction levels at which Grid implementation systems operate.

The representation and a sample grounding is depicted in Figure 3. S-OGSA entities are represented in an extension to the CIM Model. Within the project we ground the extended CIM model to WSRF. The S-OGSA extensions to CIM are as follows

- Grid Entities are represented with the class `CIM-ManagedElement` in the CIM Model.
- Knowledge Entities are represented with the new class S-OGSA-KnowledgeEntity, which is an indirect subclass of `CIM-ManagedElement`. This is how, in the model, Knowledge Entities are themselves Grid Entities..
- Finally, the association between a Grid Entity (`CIM-ManagedElement`) and a Knowledge Entity (`S-OGSA-KnowledgeEntity`), which in our model is a Semantic Binding, is represented with the new class `S-OGSA-SemanticBinding`.

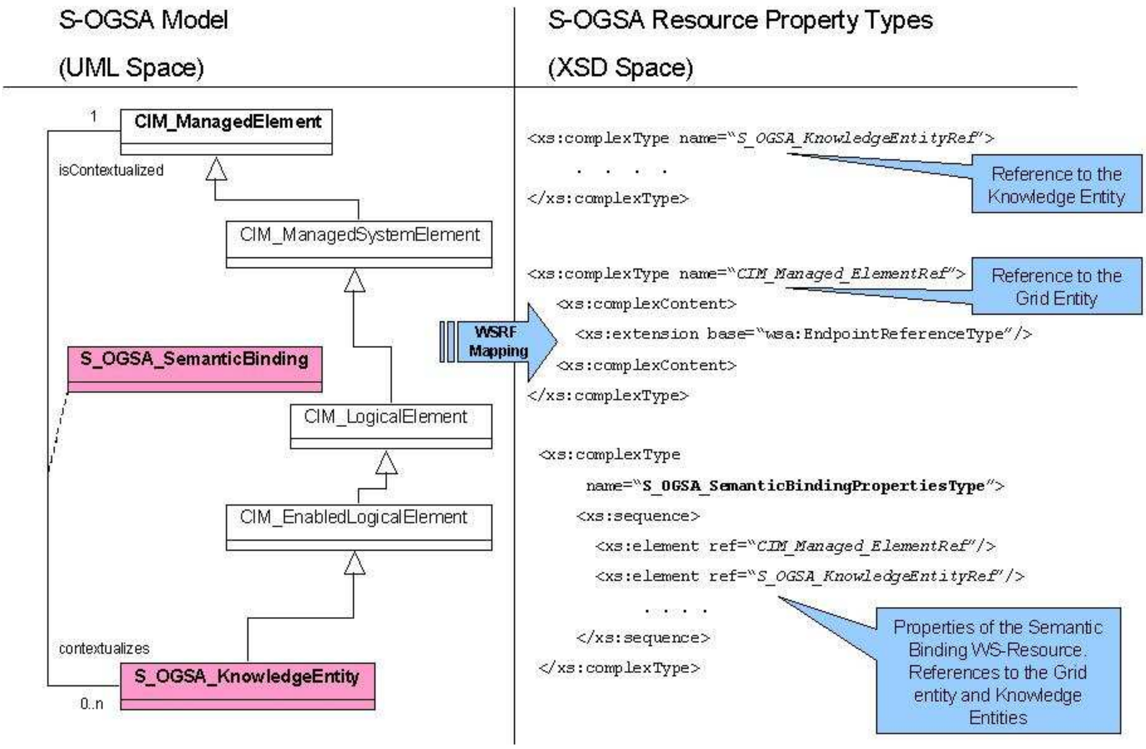


**Figure 3. Representation of S-OGSA in CIM and its Grounding to WSRF Resource Type Spaces**

**Groundings** of CIM to different Grid realization infrastructures are currently being developed. For the WSRF grounding we have used the mapping rules in [11]. Figure 3 gives an extract of the result of the WSRF grounding of the UML class `S_OGSA_SemanticBinding` to the XML Schema complex type `S_OGSA_`SemanticBindingPropertiesType that characterizes a WSRF specific representation of resources that are Semantic Bindings (SB). Since SBs represent associations between two entities, in the properties of a WSRF resource that is a Semantic Binding we can observe the end point references to the Knowledge Entities and Grid Entities (see the complex type named `S_OGSA_SemanticBindingPropertiesType` in the figure).

With S-OGSA , we also introduce the concept of ***S-Stateful Services***, which are those that provide **access to** (or virtualize) Grid resources that are affiliated with explicit metadata (that is, which have Semantic Bindings). Similar to XML-based simple metadata (e.g.. XML-based WSRF Resource Properties), the semantically encoded metadata about a resource can also retrieved and queried. The access pattern we introduce for delivering non-trivial metadata is depicted in Figure 4. The interaction of components in this figure can be summarized as follows:

- Semantic Bindings are managed as resources of their own, and since they represent metadata with respect to the entities it is associated to, it is managed via the Metadata Service (see figure 4). For instance, in a WSRF based realization, the Metadata Service would be a WSRF compliant service that provides access to resources typed with the XSD complex type named `S_OGSA_SemanticBindingPropertiesType` in Figure 3 . Knowledge, or the schema for the metadata would be treated similarly with its corresponding complex type and hosted by the Ontology Service as depicted in Figure 4.
- Upon request, Grid resources can provide end point references of their associated Semantic Binding Resources (that they know about) as part of their Resource Property set. These properties are accessed through infrastructure specific delivery operations (see steps 1 and 2 in Figure 4). For example, in the case of WSRF-based implementations, these operations are `GetResourceProperty` and `QueryResourceProperties`.
- Clients interested in exploiting Semantic Bindings can interact with metadata services through any operation that a Semantic Binding resource might support (e.g., query, retrieve value, etc.), as shown in steps 3 and 4 of the figure. The query evaluation process may also involve interaction with Knowledge Services (e.g. ontology servers, reasoners, etc.), as shown in step 5.

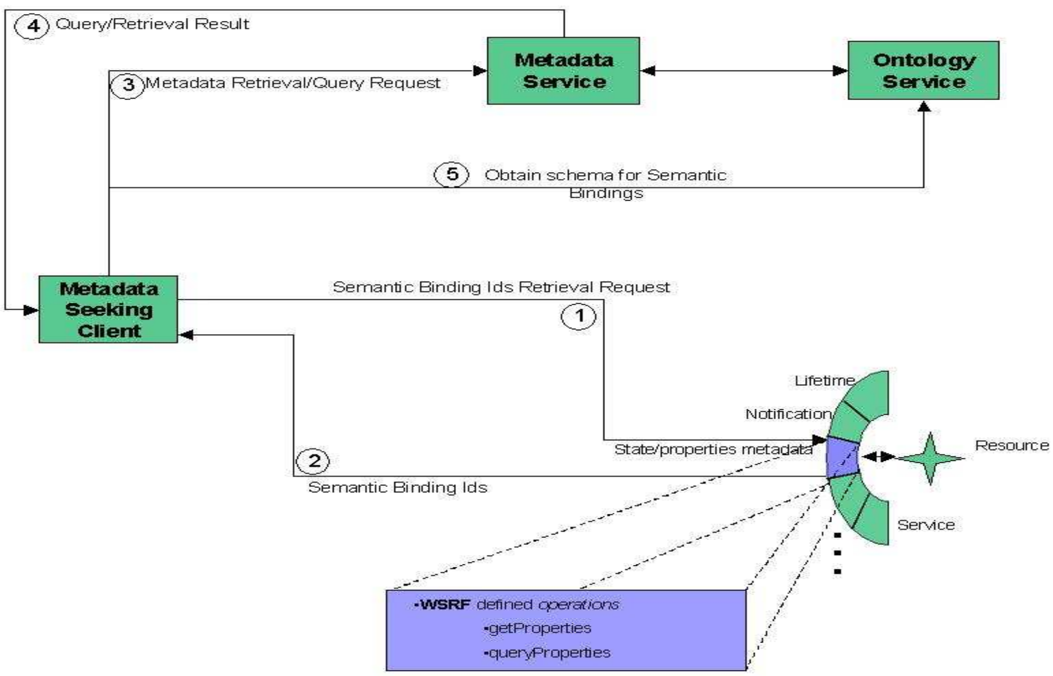


**Figure 4.** Retrieving and Querying Semantic Bindings of Resources

## Experiments

Two business case studies will be used to evaluate the S-OGSA architecture by deploying a set of Semantic Provisioning and Semantically Aware Grid services to address certain business requirements. The two case studies are concerned with the applicability of the semantic grid in the international insurance settlement and satellite data management domain respectively. As a first step of introducing a semantic grid solution in the insurance settlement domain, we are implementing Semantically Aware Virtual Organisation Management System. A preliminary version of this system is expected by mid 2006.

## Acknowledgements

The OntoGrid Consortium: Universidad Politécnica de Madrid, Spain (Coordinator), The University of Manchester, UK, The University of Liverpool UK, Technical University of Crete (TUC), Greece, Intelligent Software Components, Spain, Y'all B.V., The Netherlands, Deimos Space, S.L, Spain, Boyd International, B.V. The Netherlands. This work is supported by the EU FP6 OntoGrid project (STREP 511513) funded under the Grid-based Systems for solving complex problems.

## 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. J. 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. I. 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. The Dagstuhl Seminar No 05271; Semantic Grid: The Convergence of Technologies, organised by C. Goble, C.Kesselman and Y. Sure, July 2005.Proceedings available from http://drops.dagstuhl.de/portals/index.php?semnr=05271
7. W. Vambenepe, "Web Services Distributed Management: Management Using Web Services (MUWS 1.0)." OASIS, Standard, March, 2005.
8. K. Czajkowski, D. Ferguson, I. Foster, J. Frey, S. Graham, I. Sedukhin, D. Snelling, S. Tuecke, and W. Vambenepe, "Web Services Resource Framework (WSRF)." Globus Alliance and IBM, Technical report, 5 March, 2005.
9. Web Services for Management (WS-Management) by A. Arora et. al. http://msdn.microsoft.com/library/en-us/dnglobspec/html/ws-management1004.pdf
10. J. Treadwell, "Open Grid Services Architecture Glossary of Terms." Open Grid Services Architecture WG, Global Grid Forum, January 25, 2005.
11. W. Vambenepe, "Web Services Distributed Management: Management Using Web Services (MUWS 1.0)." OASIS, Standard, March, 2005.