S-OGSA v2.0.
Metadata Management in the Semantic Grid

www.ontogrid.eu

Oscar Corcho

L3S Institute, Hanover, Germany
October 18th, 2006
http://www.cs.man.ac.uk/~ocorcho/InvitedTalks/L3S_October2006.zip

Outline

- Background
  - The Grid and its characteristics
  - Open Grid Services Architecture-OGSA
  - Grid Standardization Activities
  - Semantic Grid
- OntoGrid and Semantic-OGSA (S-OGSA)
  - The S-OGSA model
  - S-OGSA capabilities and mechanisms
  - Lifetime specification
- S-OGSA scenarios of use
- Semantic Provisioning Services
- Conclusions and Future Work
What is the Grid?

- Heavily overloaded term
  - Analogy to On-Demand Access to Electricity

- Vision and definition
  - “When the network is as fast as the computer's internal links, the machine disintegrates across the net into a set of special purpose appliances” [Gilder]
  - Infrastructure that will enable "coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations" [Foster et al.]

Virtual Organizations

- Dynamic confederations organized around common goals
  - Diverse membership & capabilities
    - People, compute resources, data resources, etc.
  - Diverse geographic distribution

- Sharing is well-controlled
- Minimum knowledge about physical characteristics of resources
- Construction of higher level capabilities via composition of existing ones similar to SOA

From http://www.globus.org
CAUTION!

ACRONYM SPILL!!

Open Grid Service Architecture - OGSA

- Cross cutting requirements
  - Interoperable
  - VO level
  - Optimized
  - Reliable
  - Certain QoS Guarantee
  - Scalable
  - Available
  - Extensible

- Characteristics
  - Service Orientation
  - Management operations
    - Resource Representation/State
    - Lifetime

WS-RF
WS-Management
WS++
WS-GAF
Grid Reality

- Requires experts to install, configure and maintain
- Not near the ambitious OGSA landscape in terms of cross-cut requirements
- Heavy use of XML

OGSA - Vision

Realizations

“Virtual Homogeneity”

Metadata Matters

- Particularly for the following activities:
  - Provenance
  - Systems Configuration
  - Policy representation and reconciliation
  - Resource discovery

- Using:
  - Open, flexible and extensible self describing schemas that don’t have to be nailed down
    - “Let’s describe my data set, or the output format of this tool”
    - Lightweight, referable, non-flat schemas
    - Decoupled, interoperable systems, which resist to syntactic changes
  - Data integration
    - Resource models
    - Policy models
  - Formalization & Reasoning support
"The Semantic Grid is an extension of the current Grid in which information and services are given well-defined and explicitly represented meaning, better enabling computers and people to work in cooperation" D. De Roure, et. al

Semantics in and on the Grid

The Semantic Grid Roadmap

- **Pioneering Phase**
  - myGrid, Geodise, UniGrids.....

- **Systematic Investigation Phase**
  - Bridge Building:
    - Dagstuhl Seminar on Semantic Grid (July 2005)
    - Semantic Grid Workshops in OGF
    - Semantic Grid 101
  - OntoGrid and others:
    - Synergetic (Grid ←→ SW) approaches, tackling problems in middleware and application scenarios

- **Demonstration phase** (1-5 years)
  - Deployment of semantic grid systems
Outline

- Background
  - The Grid and its characteristics
  - Open Grid Services Architecture-OGSA
  - Grid Standardization Activities
  - Semantic Grid
- OntoGrid and Semantic-OGSA (S-OGSA)
  - The S-OGSA model
  - S-OGSA capabilities and mechanisms
  - Lifetime specification
- S-OGSA scenarios of use
- Semantic Provisioning Services
- Conclusions and Future Work

EU-STREP Project OntoGrid

- SEMANTIC OGSA
  - Capabilities & Behaviors for Semantic Grids
  - Principled way of realization
- Applications
  - Insurance Settlement
  - Satellite Image Quality Analysis
- Middleware for the Semantic Grid
  - Metadata Storage & Querying
  - Ontology Access
  - Annotation
    - Data and provenance
    - Services
  - Business Process Monitoring
  - Negotiation and Coordination

And other applications being analysed
S-OGSA Design Principles

- **Conceptual**: reference architecture that can be applied to any grounding (WSRF, WS-Man, WS-I+, etc.)
- **Parsimony**: Architecture as lightweight as possible: minimise the impact on tooling, not dictate content
- **Extensibility**: Extensible and customisable as opposed to complete and generic architecture
- **Diversity**: Mixed ecosystem of Grid and Semantic Grid services. Semantics Ignorant, Semantics aware but incapable, Semantics aware and capable
- **Uniformity**: Everything is OGSA compliant. Our services are Grid services, knowledge and Metadata are Grid Resources.
- **Multiform-Multiplicity**: Any resource can have multiple descriptions and any description can be in different formalisms
- **Enlightenment**: Straightforward migration path

Semantic-OGSA (S-OGSA) is...

- **Our proposed Semantic Grid reference architecture**
- **A low-impact extension of OGSA**
  - Mixed ecosystem of Grid and Semantic Grid services
    - Services ignorant of semantics
    - Services aware of semantics but unable to process them
    - Services aware of semantics and able to process (part of) them
  - Everything is OGSA compliant

**Defined by**
- Information model
  - New entities
- Capabilities
  - New functionalities
- Mechanisms
  - How it is delivered
S-OGSA Patterns. Semantic-ignorant service

Metadata Seeking Client

Access/Query Metadata

Ontology Service

Metadata Service

S-OGSA Patterns. Semantic Aware but Incapable Service

Metadata Seeking Client

Get Semantic Binding Pointers

Ontology Service

Metadata Service

1

2

S-OGSA Patterns. Semantic Aware but Incapable Service

1 Get Semantic Binding Pointers

Metadata Service

Ontology Service

Access/Query Semantic Bindings

Resource

Model

Capabilities

Mechanisms

Properties

Lifetime

Resource

Other...

client

Resource

Properties

Lifetime

Resource

Other...

Model

Capabilities

Mechanisms
S-OGSA Patterns. Semantic Aware and Capable Service

- Semantic aware interface
- Access/Query Semantic Bindings
- Farm out request
- Resource
- Other
- Semantic Service
- Metadata Service
- Metadata Seeking Client
- Ontology Service

S-OGSA Grounding. Grid Ontology and S-OGSA Ontology

- Grid Ontology
  - Common set of ontologies to describe Grid entities (resources and services)
  - Based on work from UniGrids
- Effort to be continued by OntoGrid
  - Available in OntoGrid’s CVS
What happens if...
- any or all of the Grid entities it refers to disappears?
  - Instrument and planning files for satellites do not disappear
  - Insurance contracts, cars, repair companies, etc., may disappear
- the Knowledge entities disappear or evolve?
  - Ontologies may change
- a SB is no longer available (its content is not useful any more)?
  - Damage claims: add witness reports, improve info about location, create new hypothesis...

When do/should SBs become invalid? How often should this be checked?

What is the status of the content of a SB (e.g., content checked, stable, unchecked, etc.)?

Is a SB always active or can it be archived after a period of time?
- Satellite data that is not used after some time

Lifetime specification based on WS-ResourceLifetime

Extension with
- Resource properties (state)
- Updates
- Archive
- Notifications

SB Housekeeping service
**Basic Operations**

- createSemanticBinding (Factory)
- addGridEntityReference/removeGridEntityReference
- addKnowledgeEntityReference/removeKnowledgeEntityReference
- getContents
- updateSBContent
- query
- queryWithInference

**Update Notifications**

**From entities to Semantic Binding**

Polling of resource property [lastModificationTime]
Update Notifications

- Semantic Binding Update
  - Description: Updates in the content or in the state of a Semantic Binding
  - Message content:
    - updateTime
    - updateType [stateChange, contentChange]
    - newState [any of the ones defined in the state machine]
    - updateReason

Outline

- Background
  - The Grid and its characteristics
  - Open Grid Services Architecture-OGSA
  - Grid Standardization Activities
  - Semantic Grid
- OntoGrid and Semantic-OGSA (S-OGSA)
  - The S-OGSA model
  - S-OGSA capabilities and mechanisms
  - Lifetime specification
- S-OGSA scenarios of use
- Semantic Provisioning Services
- Conclusions and Future Work
S-OGSA Scenario. Satellite Image Quality Analysis

Scenes:
- Routine operations
- Metadata generation
- Report retrieving

Satellite Lifecycle:
- Launch and Early Orbit Phase (~ 3 days)
- Calibration and Validation campaign (~ 6-9 months)
- Routine operations (~ 5-9 years)
- Satellite de-orbiting. Product processing continues

Satellite Routine Operations

Satellite with instruments

WebDAV

WebDAV client

e.g. MS Windows Explorer

Copy satellite XML summary file

Convert time to canonical representation

UTC2Seconds

Soaplab

Input criteria

Convert time to canonical representation

QUARC-SG client

Query

Metadata Service

Metadata generation process

Metadata querying process

Grid-KP

Obtain ontology

Type metadata

WS-DATInt

Satellite Domain Ontology

XML Summary File

HTTP PUT

Annotate file

Store

Oscar Corcho. L3S, Hanover, 18 Oct 2006
S-OGSA Scenario. Insurance settlement

- Data and resources scenarios
  - Register Repair Co. contract at CarRepairGrid.
  - Select Repair Companies for negotiation

- Metadata scenarios
  - Calculate offer by a Repair Company (damage report)
  - Judge Invoice sent by Repair Company

- Process management scenarios
  - Multi issue negotiation between Repair Companies (repair)
  - Multi issue negotiation between >3 insurance companies (Recovery)

- Services scenarios
  - Provide Policy Information
  - Check coverage

- Security scenarios
  - Check client registration at insurance companies
  - Check Car Theft - automatic check on car identity i.e. frame numbers and parts

Oscar Corcho. L3S, Hanover, 18 Oct 2006
International Insurance Settlement Scenario

- Insurance Security scenario cast as role based Grid Access Control Scenario.
- **Role based Access Control Policy** is:
  - Good Reputation Drivers are allowed to ask for an insurance policy. Bad Reputation ones are not.
- **VO ontology** based on
  - KaOS ontologies (Actors, Groups and Actions)
- **Role definitions**
  - Extend ontology with domain-specific classes and properties
  - Define roles wrt these extensions
    - E.g., a blacklistedDriver is a driver that has had at least 3 accident claims in the past
    - E.g., a goodReputationDriver is a driver that has been insured at least by one trusted company and that has had at most 2 accident claims
- The **Access Control Function** uses a DL classifier to obtain roles of a Subject.

---

S-OGSA Scenario. Authorisation

(Covered in the S-OGSA slide.)

Oscar Corcho. L3S, Hanover, 18 Oct 2006
S-OGSA Scenario. Authorisation

1. User requests a policy.
2. The service checks the user's role.
3. It obtains semantic bindings of the user.
4. It classifies the user with respect to the ontology.
5. It looks up whether the role permits the action.

Semantic Binding Service

John Doe has had 2 distinct accidents.

Pellet Reasoner

Obtain Semantic Bindings of John Doe
Obtain all classes that are subclass of ROLE
Classify John Doe wrt VO ont
Lookup whether the ROLE that is inferred permits or not

VO Ontology

XACML_AuthZService (PDP)
CarFraudService (PEP)
XACML_AuthZResponse

S-OGSA Scenario. Authorisation
S-OGSA Scenario. Authorisation

1. Get insurance policy
2. Result or exception
3. Semantic binding service
4. Obtain semantic bindings of John Doe
5. Pellet reasoner
6. Obtain all classes that are subclass of ROLE
7. Classify John Doe wrt VO ont
8. Look up whether the ROLE that is inferred permits or not

Semantic Binding Service

Role
- BlockedDriver
- GoodReputationDriver

Pellet reasoner

VO Ontology Class Hierarchy -RDFS

John Doe has had 2 distinct accidents

S-OGSA Scenario. Authorisation

1. Get insurance policy
2. Result or exception
3. Semantic binding service
4. Obtain semantic bindings of John Doe
5. Pellet reasoner
6. Obtain all classes that are subclass of ROLE
7. Classify John Doe wrt VO ont
8. Look up whether the ROLE that is inferred permits or not

Semantic Binding Service

Role
- BlockedDriver
- GoodReputationDriver

Pellet reasoner

VO Ontology Class Hierarchy -RDFS

John Doe has had 2 distinct accidents
Outline

- Background
  - The Grid and its characteristics
  - Open Grid Services Architecture-OGSA
  - Grid Standardization Activities
  - Semantic Grid
- OntoGrid and Semantic-OGSA (S-OGSA)
  - The S-OGSA model
  - S-OGSA capabilities and mechanisms
  - Lifetime specification
- S-OGSA scenarios of use
- Semantic Provisioning Services
- Conclusions and Future Work
Motivation for a Grid Compliant Ontology Service

- Current Semantic Grid Applications access ontologies using Semantic Web technologies
  - No Grid compliant ontology access mechanisms for RDF(S) and OWL ontologies that exploit resource sharing capabilities
  - No specialized ontology registries
  - No Grid compliant instance access mechanisms in RDF

WS-DAIOnt-RDF(S) Implementation Architecture

- Web Services Data Access and Integration – The Ontology Realization
  - First implementation focused on RDF(S): WS-DAIOnt-RDF(S)
    - Sesame and Jena
      - Future: Oracle RDF-store
    - Query languages: SPARQL, SeRQL, RDQL, RQL
  - Two-tier architecture
    - Web Service tier
      - Upper service layer
      - Intermediate service layer
      - Lower service layer
    - RDF(S) access tier
      - RDFSConnector
      - Storage layer
Outline

- Background
  - The Grid and its characteristics
  - Open Grid Services Architecture-OGSA
  - Grid Standardization Activities
  - Semantic Grid
- OntoGrid and Semantic-OGSA (S-OGSA)
  - The S-OGSA model
  - S-OGSA capabilities and mechanisms
  - Lifetime specification
- S-OGSA scenarios of use
- Semantic Provisioning Services
- Conclusions and Future Work

Conclusions

- A principled Semantic Grid reference architecture
  - Low-impact extension of OGSA
  - Mixed ecosystem of Grid and Semantic Grid services
- Ontology and metadata technology...
  - ... can be used in Grid applications
  - ... has to be adapted for its use in Grid environments
    - Grid-compliant (provide Grid protocols, interfaces, etc.)
    - Grid-aware (use of Grid technology)
- First use cases being deployed
  - Still far from large-scale (production) deployment
Our future work

- Semantic Binding Service
  - Lifetime management
  - Fine-grained AuthZ
- Prototypes demonstrating Knowledge-Aware Grid Services
  - OGSA-DAI semantic extensions
  - EGEE information service consolidating heterogeneous information sources
  - Meta-scheduler using semantic technologies
- Enlightenment
- Guidelines about how and when to apply semantic technologies in Grid systems

More information

- Publications
  - http://www.ontogrid.eu/, Deliverable D1.2v2
- Source code
  - http://www.ontogrid.eu/, For Downloading Distributions
  - Access to CVS
    - Connection type: pserver
    - user: ontogrid
    - password: not needed
    - Host: rpc262.cs.man.ac.uk
    - Port: 2401
    - Repository path: /local/ontogrid/cvsroot
    - module: prototype
Thank you for your attention!

Questions?

Acknowledgements

- OntoGrid Consortium
  - Pinar Alper, Ioannis Kotsiopoulos, Paolo Missier, Sean Bechhofer, Carole Goble
  - Miguel Esteban and Asunción Gómez-Pérez (WS-DAIOnt)

Oscar Corcho. L3S, Hanover, 18 Oct 2006