A semantically-enhanced grid registry: Work in progress

Sylvia Wong, Victor Tan, Weijian Fang, Simon Miles, Luc Moreau

School Electronics and Computer Science
University of Southampton, UK

www.grimoires.org
Grimoires Grid Registry Project

- Grimoires
  - Grid Registry with Metadata Oriented Interface: Robustness, Efficiency, Security
Semantically Discovery of Services

- Service discovery difficult on the Grid
  - Large number of services advertised

- Semantic descriptions of services
  - Filter out most suitable services
  - Augment service descriptions with extra information (metadata) useful to discovery
  - Service providers
    - Access polices, contract negotiation details
  - Users
    - Quality of service, reputation metrics
Requirements of Metadata
Attachments

- Annotation to all concepts that influence discovery
  - Services
  - Operations supported by services
  - Input and output types of operations

- Multiple attachments

- Third party metadata
  - Users to enrich descriptions not foreseen by providers

- Efficiency in updates
  - Some metadata can change frequently, eg user ratings
  - Can be updated without republishing the entity or other metadata attached
Metadata Representation

- Semantic descriptions as RDF statements
  - Subject: entity to be annotated
  - Predicate: type of annotation
  - Object: value of annotation
    - string, URI or RDF graph
  - Examples
    - (service, mygrid:NumericRating, 8.5)
    - (input, mygrid:SemanticType, mygrid:Nucleotide_Sequence)

- Provenance information
  - date, author
Metadata Inquiries

- Multiple query patterns, from simple to complicated
  - List of all metadata attached to a service
  - List of all entities with metadata that match a list
  - RDQL (RDF query language)
    - For query patterns not exposed in inquiry interface
    - Example: Metadata data values are exact matches currently, use RDQL to find all services with user ratings > 8.5
Architecture

- UDDI compatible
- Multiple web services containers
  - Tomcat, Apache Axis, Globus Toolkit 4, OMII
- Multiple triple store memory backends
  - In-memory: Faster, enough for 1 million services
  - PostgreSQL, MySQL, Berkeley DB

[Diagram showing the architecture, including connections between UDDI compliant client, Grimoires client, uddi4j, etc, and components like UDDIv2 interface, metadata interface, authorisation module, Jena triple store, and Registry.]
Signature-based Authentication

- UDDI v2 and v3
  - Username/password credential scheme
  - Authentication tokens

- Grid environments typically use certificate-based authentication schemes
  - Eg, Globus, OMII

- Grimoires in OMII container
  - WS-Security standards compliant SOAP message signing and verification
  - Authentication using Distinguished Name (DN) extracted from submitted X509 client certificate

- Benefits
  - Easy integration into existing Grid security infrastructure
  - First step to single sign-on
Performance

- Preliminary performance tests
  - Identify problem areas in efficiency and scalability
  - Service publication and inquiry
  - Metadata inquiry
Performance Test: Publish

Publishing 100 Services against Registry Data Size

GRIMOIRE, WS
GRIMOIRE, BL
jUDDI 0.9rc4

Average publication time: 30ms per service
Performance Test: Inquiry

Average inquiry time: 100ms per service
Future work

- RDFS and OWL support
  - Ontology aware service discovery

- Access control on metadata attachments
  - Third-party publication leads to more complicated access patterns
    - Who can annotate a service?
    - Who can update a piece of metadata?
    - Querying only a subset of metadata?

- Performance improvements
  - Different triple store implementations
  - Distributed registry
Summary

- Presented a semantically-enhanced grid registry
  - Metadata interface
  - Signature based authentication
  - Preliminary performance tests
  - Future work