Globus – GTK and Grid Services

Michael Rokitka
SUNY@Buffalo
CSE510B
9/2007
OGSA – The Open Grid Services Architecture

• What are some key requirements of Grid computing?
  – Interoperability: Critical due to nature of heterogeneous environments
  – Virtualization: virtual representation of organizations, resources, applications

• What is OGSA?
  – OGSA defines a Grid system architecture based on well defined standards and Web service technologies
  – Service Oriented Architecture based Grid services provide well-defined interfaces for clients
  – All grid resources (logical & physical) are modeled as services
Web Services Overview

• Web Services employ the client-server architecture and provide interoperable interfaces to applications

• Intended for use by other software, NOT directly by users.
  – Request->Response/Producer->Consumer
Web Services cont.

Diagram:

1. **Client** asks: Where can I find a "weather service"?

2. **Server A** responds: There's a "weather service" in Server B.

3. **Client** asks: How exactly should I invoke you?

4. **Server B** responds: Take a look at this: WSDL.

5. **Client** invokes: `getWeatherInfo()` with parameter "12345".

6. **Server B** responds: "Cloudy with a chance of meatballs".
OGSI – Open Grid Services Infrastructure

• What is OGSI?
  – Formal specification of concepts described by OGSA
  – Specifies extensions to standard Web services needed for Grid services
    • Stateful Web services
    • Extended WSDL
  – Defined Grid service interfaces, lifecycle, concept of Grid service instances, references, and handles
OGSI Model

Grid Service Reference

Grid Service Instance(s)

Grid Service Handle (URI)
The Downfall of OGSI

• Problems found with OGSI
  – Too much stuff all in one specification
  – Doesn’t work well with existing web service tools (too heavy on XSD, etc.)
  – Too object oriented
    (instances, handles, etc.)

• OGSI (base for GT3) replaced by WSRF (base for GT4)
  – WSRF separates the service and its state (resources)
# WSRF – Web Service Resource Framework

<table>
<thead>
<tr>
<th>OGSI</th>
<th>WSRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid Service Reference</td>
<td>WS-Addressing Endpoint Reference</td>
</tr>
<tr>
<td>Grid Service Handle</td>
<td>WS-Addressing Endpoint Reference</td>
</tr>
<tr>
<td>HandleResolver portType</td>
<td>WS-RenewableReferences</td>
</tr>
<tr>
<td>Service data defn &amp; access</td>
<td>WS-ResourceProperties</td>
</tr>
<tr>
<td>GridService lifetime mgmt</td>
<td>WS-ResourceLifeCycle</td>
</tr>
<tr>
<td>Notification portTypes</td>
<td>WS-Notification</td>
</tr>
<tr>
<td>Factory portType</td>
<td>Treated as a pattern</td>
</tr>
<tr>
<td>ServiceGroup portTypes</td>
<td>WS-ServiceGroup</td>
</tr>
<tr>
<td>Base fault type</td>
<td>WS-BaseFaults</td>
</tr>
</tbody>
</table>
Globus Toolkit

• What is the GTK?
  – Framework for developing OGSA Grid services

• What’s included?
  – Basic infrastructure services
    • Job submission/management
    • File transfer
    • Data management
    • Monitoring
  – Grid development environment
    • Programmatic APIs
    • Security mechanisms
  – Tools and examples
Infrastructure Services

- “Connectivity Layer” Solutions
  - Service Management (WS Core)
  - Monitoring/Discovery (WS Core)
  - Security (GSI and WS-Security)
  - Communication (XIO)
- “Resource Layer” Solutions
  - Computing / Processing Power (GRAM)
  - Data Access/Movement (GridFTP, OGSA-DAI)
  - In development: Telecontrol (GTCP)
- “Collective Layer” Solutions
  - Data Management (RLS, DRS, RFT, OGSA-DAI)
  - Monitoring/Discovery (Index, Trigger, Archiver services)
  - Security (CAS, MyProxy)
GT4 Components
Security Mechanisms

Features:

- A "plug in" framework for authorization decisions
- Single sign-on support implemented as "proxies"
- Grid-wide identities implemented as PKI
- Local security administration & enforcement
- Ability to map between Grid & local identities
- Ability to delegate credentials to agents
- Transport-level and message-level authentication
- Grid-wide identities implemented as PKI

Security Mechanisms
Overview of Components - Security

• SimpleCA
  – Tool used for small test grid setups, typically not used in production environments as certificates must be signed by well known CA to ensure authenticity

• MyProxy
  – Remote service used to store user credentials
    • Can use CA, Kerberos, etc.
    • Simplifies certificate management
Information Services

• Index Service
  – Provides registry capability, caching

• Trigger Service
  – Used for monitoring, e-mails alerts when pre-defined conditions are met

• Archive Service
  – Logs values of resource property info over time, useful for auditing

• WebMDS
  – Web browser interfaces to resource properties and index services
    • Collects monitoring info from pluggable sources
    • Applies XSLT to XML to produce human-friendly HTML (uses Tomcat)
Execution Management

• **GRAM**: Grid Resource Allocation & Management
  – Provides basic job submission and control
    • Uniform service interface for remote submission and management of jobs
    • **NOT** a scheduler (fork, Condor-G, PBS, etc. are used for this)
    • Often used as a front end to schedulers and to simplify metaschedulers & brokers
Data Tools

• GridFTP
  – FTP enhanced for security and high performance
    • Multiple data channels for parallel transfer
    • Partial file transfer

• RFT: Reliable File Transfer
  – Provides mechanism for file transfer queuing

• OGSA-DAI: OGSA Data Access & Integration
  – Interface for accessing relational/XML data stores
    • Registry
    • Grid Data Service Factory

• MCS: Metadata Catalog Service (integrated in OGSA-DAI)
  – Stores system or user defined attributes
  – Supports manipulation and query
References

• The physiology of the Grid

• OGSI specification v1

• Globus Toolkit 4 Primer
  http://www.globus.org/toolkit/docs/4.0/key/GT4_Primer_0.6.pdf

• Globus Primer – A tutorial at GlobusWORLD 2006
  http://www-unix.mcs.anl.gov/~liming/primer/Primer-Slides-2.0.ppt

• Wikipedia
  http://en.wikipedia.org/wiki/Web_service