

GridSAM - A Standards Based Approach to Job Submission Through Web Services

William Lee and Stephen McGough

London e-Science Centre

Department of Computing, Imperial College London

- GridSAM Overview
- Job Submission Description Language
- Job Submission and Monitoring Port Types
- GridSAM Architecture
- GridSAM Implementation
- Planned Works
- Summary

- What is GridSAM?
 - Part of the OMII managed programme
 - Client's Perspective
 - Job Submission and Monitoring Web Service
 - Standardised Job Description
 - Virtual file input and output sandboxes
 - Client-side submission clients
 - Developer's Perspective
 - Extensible JobManager API interfacing with existing Distributed Resource Managers (DRM)
 - Used as an embedded library
 - Job Submission Portal
 - Grid Applications
- What not?
 - A job scheduling system
 - Replacement of existing Grid resource management system.

- Job Submission Description Language (JSDL)
 - Job Description Markup Language (JDML) is one of the feeder language developed in the EU-DataGrid project.
 - Standardising through the Global Grid Forum JSDL-WG

Requirements

Application

DataAttributes

JSDL

```

<?xml version="1.0" encoding="UTF-8"?>
<jSDL:job xmlns:jSDL="http://www.gridforum.org/JSDL"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.gridforum.org/JSDL jSDL.xsd">
  <jSDL:JobIdentification>
    <jSDL:JobName>This is my job - a short description
    </jSDL:JobName>
  </jSDL:JobIdentification>
  <jSDL:Application>
    <!-- type should be xsd:token not xsd:string -->
    <jSDL:ExecutableName type="jSDL:onPath">mycommand</jSDL:ExecutableName>
    <jSDL:Argument>-somearg</jSDL:Argument>
    <jSDL:Argument>inputfile.txt</jSDL:Argument>
    <jSDL:StdOut>output.txt</jSDL:StdOut>
    <jSDL:StdErr>error.txt</jSDL:StdErr>
  </jSDL:Application>
  <jSDL:DataAttributes>
    <jSDL:File>
      <jSDL:FileName>inputfile.txt</jSDL:FileName>
      <jSDL:Source>http://mysite/some/public/file.txt</jSDL:Source>
    </jSDL:File>
    <jSDL:File>
      <jSDL:FileName>output.txt</jSDL:FileName>
      <jSDL:Target>gridftp://mygridserver/incoming/output.txt</jSDL:Target>
    </jSDL:File>
    <jSDL:File>
      <jSDL:FileName>error.txt</jSDL:FileName>
      <jSDL:Target>ftp://myftpserver/incoming/output.txt</jSDL:Target>
    </jSDL:File>
  </jSDL:DataAttributes>
</jSDL:job>

```

JobSubmissionSOAPPort

JobMonitoringSOAPPort

GridSAMService

submitJob
Input: JobDescription
Output: JobIdentifier
Fault: SubmissionFault

JobSubmissionPortType

JobSubmissionSOAPBinding

getJobStatus
Input: JobIdentifier
Output: JobStatus
Fault: UnknownJobFault

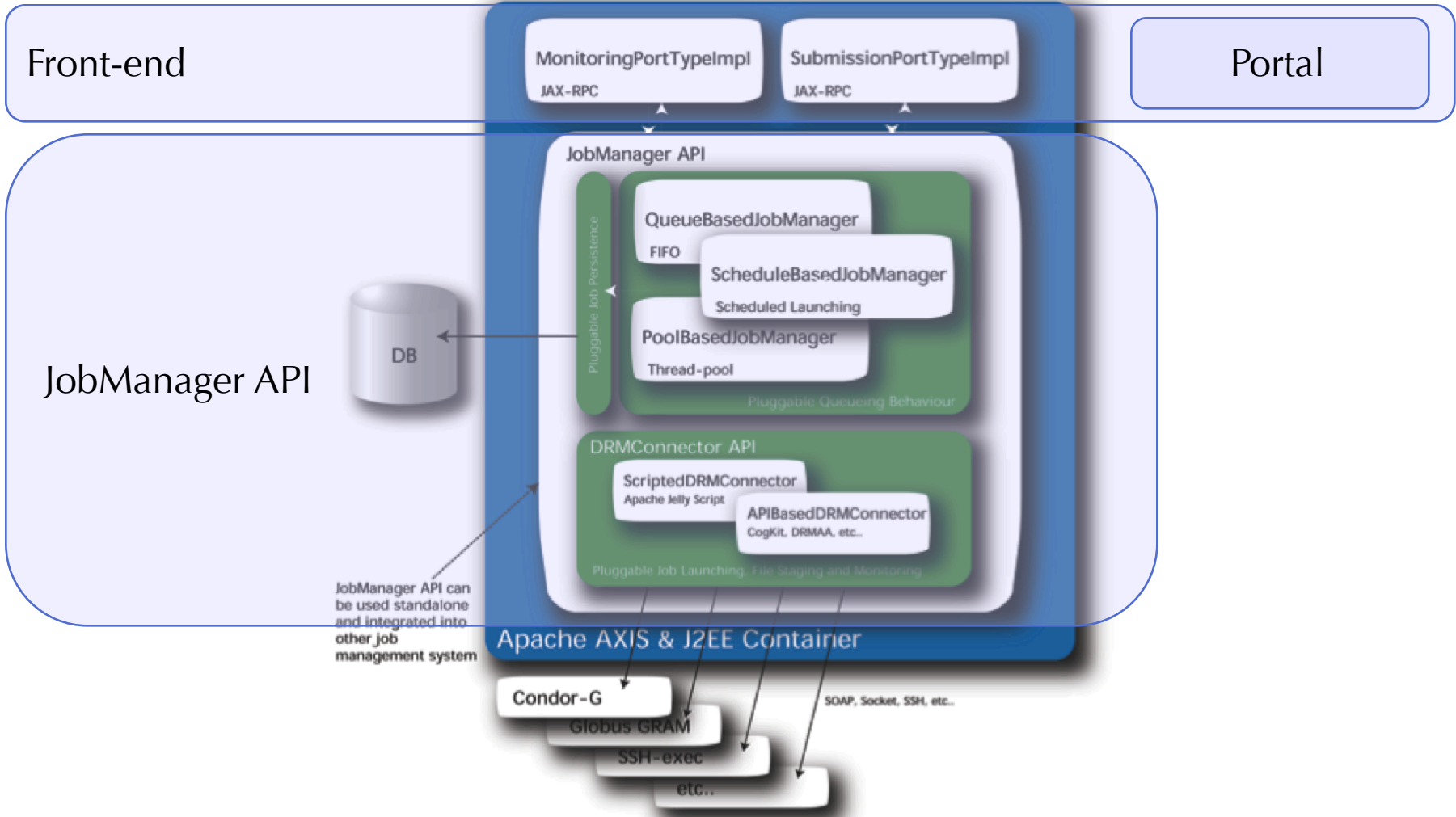
JobMonitoringPortType

JobMonitoringSOAPBinding

WS-Security X509 Certificate Profile
Signed and optionally Encrypted SOAP
Messages

MonitoringPortType

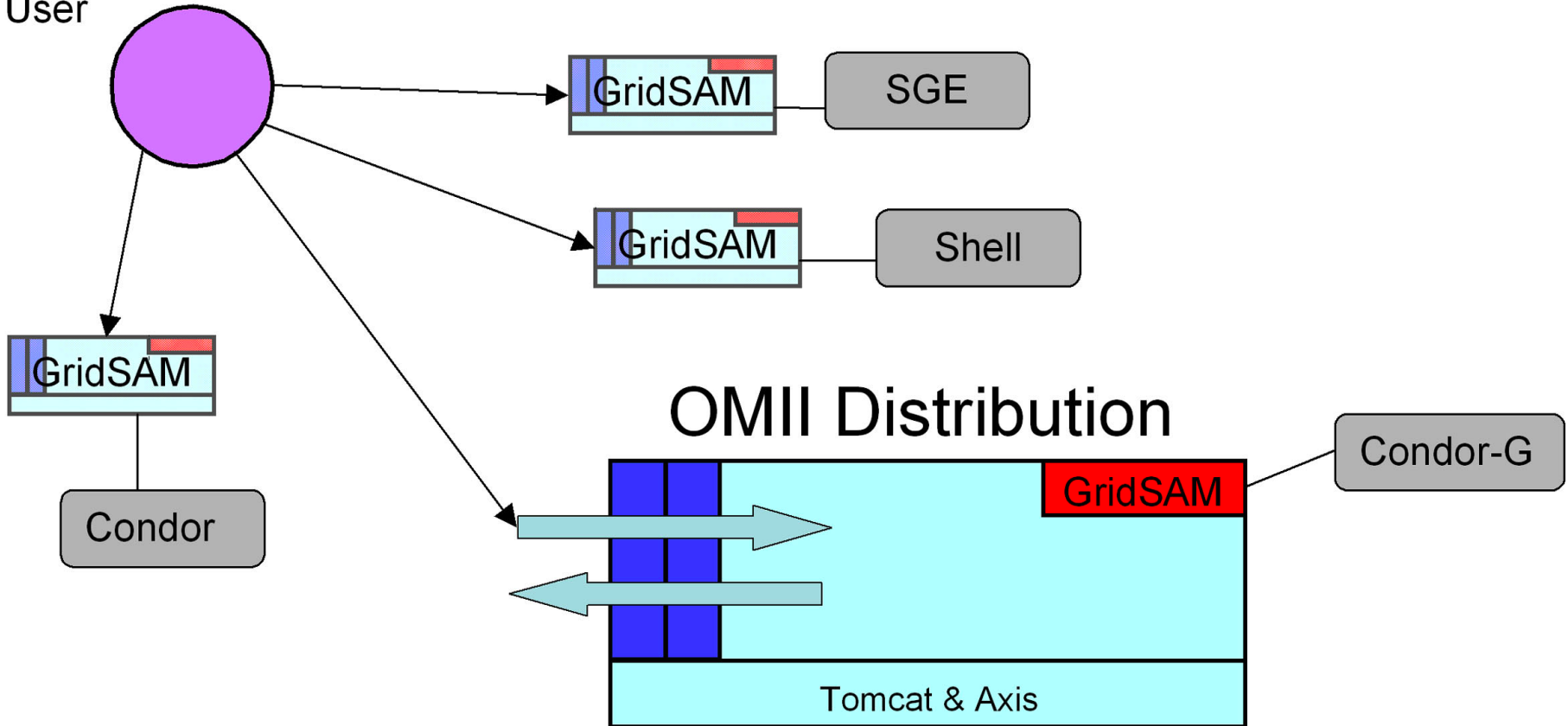
SubmissionPortType



- WS-I Basic and Security Profiles
 - XML, XSD
 - WSDL (document/literal)
 - SOAP
 - WS-Security
- WS-Notification*
- JSDL*

- J2EE
 - Implemented using the J2EE JAX-RPC Web Service API.
 - Tested on the Sun Application Server v.8
 - Exploits clustering and Java Message Queue
 - Cross-deployable on other J2EE compliant servers
- Apache Axis on Tomcat Server
 - Porting to Apache Axis Web Service integrating with the OMII base distribution
- Core functionalities reside in JobManager API, front-ends are pure wrappers providing front-line security to the underlying system.

Application/
User



- JobManager implementation - pluggable queuing behaviour

- CalendarBasedJobManager

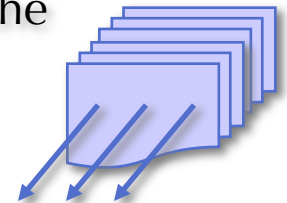
- Passes job to DRMConnector according to a timetable

- ThreadPoolBasedJobManager

- Pool of threads taking jobs off the queue passing them to the DRMConnector

- JMSJobManager

- Java Message Queue based job queuing. Allows DRMConnectors to be clustered using J2EE features



Implementation

DRMConnector API

- Interacting with existing DRM or providing new launching mechanism.
 - ScriptBasedDRMConnector
 - An helper implementation third-party developers can extend.
 - Converts declarative JSDL into an executable set of actions in the form of an *Apache Jelly* script.
 - Staging in files
 - Generating lower-level script
 - Invoking external commands
 - Staging out files
 - In-line XML parsing of JSDL using XPath or Xquery
 - Exploits the wealth of “Jelly Tags” (plugins) available from the open-source community.
 - Aids diagnostic

- DRMConnector implementation - DRM-specific plugins
 - ForkDRMConnector
 - Spawns job locally
 - Provides limited requirement specification
 - CondorDRMConnector
 - Launches job through Condor
 - Exploits requirement matching capability in Condor
 - Submits jobs to Globus resources through Condor-G
 - Others
 - SGE
 - PBS
 - etc..

- Job Persistence Layer
 - JobQueue
 - InMemoryJobQueue
 - For testing or embedded use
 - JDBCJobQueue
 - Job information and state changes are persisted to RDBMS using Java JDBC API
 - JMSJobQueue
 - Use Java Messaging Queue to persist durable job information.

- Currently supported file transfer protocols
 - Input: http(s)://, ftp://, file://
 - Output: ftp://, file://
- Planned support
 - Gridftp: gridftp://
- Others
 - Secure FTP: sftp://
 - WebDAV: webdav://
 - etc..

- WS-Security - Message Level Security
 - User signs SubmitJobRequest message.
 - Server checks message integrity and authorise access using the J2EE role-based access control or the OMII authorisation mechanism.
- HTTPS - Transport Level Security
 - Encrypted data stream
 - X509-certificate mutual authentication on the portal front-end

- Jobs are submitted under a mapped native user account
- Investigate secure delegation in order to perform secured file staging on behalf of users
- Job state notification
- Continuously track development of JSDL
- Full use of Condor-G functionality through CondorDRMConnector
- Provides experimental supports for other DRM systems

- GridSAM
 - is a Job Submission and Monitoring Web Service
 - is a Job Management API embeddable in other applications
 - uses a standardised Job Description
 - provides a standard-based approach to interface with existing DRM systems
 - supports a variety of file transfer protocols for staging files in-and-out of jobs.

For more information and demonstration

- London e-Science Centre
 - <http://www.lesc.imperial.ac.uk>
- Open Middleware Infrastructure Institute
 - <http://www.omii.ac.uk>

- Director: Professor John Darlington
- Research Staff:
 - Nathalie Furmento, Stephen McGough
 - William Lee, Jeremy Cohen
 - Marko Krznaric, Murtaza Gulamali
 - Asif Saleem, Laurie Young, Jeffrey Hau
 - David McBride, Keith Sephton
- Others:
 - Steven Newhouse, Yong Xie, Gary Kong
 - James Stanton, Anthony Mayer
- Contact:
 - <http://www.lesc.ic.ac.uk/iceni>
 - e-mail: lesc@ic.ac.uk