Open Science Grid Software Stack, Virtual Data Toolkit and Interoperability Activities

D. Olson, LBNL for the OSG
www.opensciencegrid.org

International Symposium on Grid Computing, Taipei
7-11 April 2008
Abstract

We will present the current status and plans for the Open Science Grid Software Stack, Virtual Data Toolkit and OSG Interoperability Activities. During 2008 our focus in the VDT is to update our support for Condor, VOMS/gLite, Globus, and other components, and add full support for resource and service validation software, storage management access through BestMan and dCache. We are improving the ability to update versions of VDT "at the side" of existing installations, and reduce the time and effort for updating to new versions of the OSG software. Support for new variants of Linux, MacOSX and AIX are also planned. The Open Science Grid software stack supports the OSG model of interoperation, which provides service bridges and/or adaptors between distributed infrastructures and not necessarily direct compatibility between service interfaces - which we regard as an impossible goal. We have successfully demonstrated this model for interoperation with the EGEE, GLOW, Clemson and FermiGrid campus infrastructures, and ongoing work with TeraGrid.
Contents

• The software stack – emphasis for 2008
  – Updates for Condor, VOMS/gLite, Globus, …
  – Support of Resource Validation
  – Storage elements
  – In-place installation upgrades
  – Additional OS support

• Interoperability
  – Interoperation model
  – Examples

• Acknowledgements
The VDT

An internally consistent build of a large suite of middleware across variety of platforms. Condor, Globus, VOMS/gLite, …

What is in VDT 1.8.1?

What Software is Supported on each Platform?

Legend

- Supported
- Unsupported

<table>
<thead>
<tr>
<th>CentOS 5</th>
<th>CentOS Debian 3.1</th>
<th>Fedora Core 4</th>
<th>Red Hat Enterprise Linux 3</th>
<th>Red Hat Enterprise Linux 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Apache Ant 1.6.5
- Apache HTTPD 2.2.4
- Apache Tomcat 5.0.28
- Apache Tomcat 5.5.25
Emphasis in 2008

- Updates for
  - Condor
    - Currently 6.8.8
    - Improvements for WS-GRAM
  - Globus
    - Currently 4.0.5
    - Adjustments for OSG deployment of WS-GRAM
    - Expecting GT4.2
  - VOMS/gLite
    - Converging on common user registration function of voms-admin and VOMRS
    - Multiple certificates/user
    - More administrative roles for registration workflow
  - authZ
    - Interoperability update (later slide)
- Full support for Resource Service Validation
  - See talk in Operations & Management I session
More Emphasis in 2008

- **Storage**
  - BeStMan
  - dCache
  - Xrootd/Scalla
    - Code is in VDT, deployment configuration is in progress

- **Better support of in-place upgrades**
  - Carrying over previous configuration to new installation

- **Requests for RPM and Debian packages**

- **Goal of same version of VDT packages across OSG, TeraGrid, EGEE**

- **Additional platforms**
  - Mac OS X
  - AIX
  - Additional linux flavors
**BeStMan SE in VDT**

- Light-weight SRM using unix disk
- One step install for default configuration
- Client access to data:
  - available as posix filesystem
  - optional as gateway to Xrootd
  - optional interface to HPSS and adaptable to other MSS
dCache SE in VDT

Key Features of VDT-dCache

- Convenient and easy deployment of SRM/dCache to create an OSG Storage Element (SE)
  - A single configuration file for multi-node installations
  - Reduces installation steps to 20% for regular install
- Hardware and configuration recommendations
- Dry run option for the installation
- dCache Gratia probe
- Step-by-Step installation cleanup, pnfs upgrade & Gratia setup instructions
- Installation and debugging support
Interoperability / Interoperation

- RSV-SAM project for resource info interoperability
- authZ interoperability
  - Effort underway to have common authZ framework across
    - Involves OSG, EGEE, Globus, Condor
    - Policy components - GUMS, SAZ, LCAS/LCMAPS, …
    - Resource gateways – Gatekeeper, SRM, gLExec, …
  - Expect to be in testing by this fall
- Model of interoperation is to support bridges and gateways between differing infrastructures and not at the low level service API level.
  - Reduces dependencies to a manageable level.
- Interoperations today
  - EGEE
  - GLOW
  - Clemson
  - Fermigrid
  - working on TeraGrid
Resource and Service Validation schematic

OSG sites can publish to WLCG via RSV-SAM bridge

Schematic of RSV
(Proposed Phase II work shown in dotted-dashed line).
OSG-EGEE Interoperations

• Weekly WLCG operations meeting
• Trouble ticket interchange
• Publishing service info to WLCG/SAM for those sites participating in WLCG
• VDT builds of Globus, Condor used by EGEE
• CMS analysis jobs run across multiple grids from central submission point (next slide)
Example: CMS Analysis Snapshot

~ 1/3 of worldwide jobs running on OSG
Submitting jobs through OSG to UW Campus Grid

Routing Jobs from UW Campus Grid to OSG

Combining both worlds:
- Simple, feature-rich local mode
- When possible, transform to grid job for traveling globally
**OSG-Clemson**

- **Clemson campus grid**
  - Windows Condor pool
  - Unix applications in Colinux or Cygwin
  - No shared filesystem
  - Linux gatekeeper running NFSlite jobmanager
OSG-TeraGrid Gateway
Planned for summer 2008

OSG-TG Gateway
Web Services Architecture
Acknowledgements

• Deploying a functional and evolving software stack is a large effort of many dedicated people
  – The VDT team
  – The integration team
  – The documentation team
  – Feedback from the administrators and users