IGE Roadmap

Steve Crouch
Helmut Heller

25 Oct 2010

Introduction and Overview

- Addressing the EGI UMD Roadmap V3 from August 27th, 2010
  - IGE as technology provider
  - Prioritization of capabilities
  - Timelines given where possible
  - UMD Roadmap is a living document - IGE Roadmap will also change over time

Introduction

- Early days for IGE!
  - Start 1st of October, KickOff last week
  - Our requirements gathering process hasn't yet begun; our contributions are preliminary
  - Users are Globus end-users and also EGI, DEISA, PRACE, EU-IndiaGrid2, …
- Target is to have our components **integrated** into UMD

Capability: Information System

- Currently under investigation in IGE
- Globus IIS – not appropriate at this time, although looking into development for the future
- Perhaps strong requirements from IGE within GLUE for this aspect

Capability: Compute

- IGE is actively involved in OGF PGI-WG
  - one use case contributed by IGE
- JSDL (subset, including POSIX Application and HPC Application Profile Extension) already now supported via GridWay (compliance tests needed)
- BES support via GridWay planned for mid 2011
- Plan to develop native HPC-Basic Profile/ OGSA BES/JSDL component for Globus GRAM5

Capability: Compute Job Scheduling

- Meta-Scheduler: GridWay
- File stage-in and stage-out is supported
- Plan to extend GridWay support for standards compliance:
  - GridWay will support submitting jobs to BES endpoints by March 2011
  - GridWay will expose a BES interface by second half of 2011
- **High priority**

Capability: Database Access

- OGSA-DAI supported as access layer
  - Available now
  - Accessing data resources via web type servers in enterprise architectures is well understood and supported well in industry, commerce and academia by solutions such as the JEE
  - OGSA-DAI is alternative solution for *distributed* data access and management using heavier weight "grid services"
  - OGSA-DAI allows data resources - for example MySQL5, DB2, Oracale10, SQL Server, Postgres, and other data sources - to be federated and accessed via web services on the web or within grids or clouds. Via these web services, data can be queried, updated, transformed and combined in various ways
  - OGSA-DAI is fully compliant with the GT4.x; IGE will port to GT 5.x

Capability: Metadata Catalogue

- Work starting for globusonline.org in January/February 2011.
  - Initial: single attribute/value as a start
  - Googledoc spreadsheet integration
  - Driving use cases: pull images off instrument, extract thumbnails, put into Googledoc spreadsheet (as a metadata catalogue), drive analysis off that

Capability: File Transfer

- GridFTP
  - good for transfers inside a Grid
  - difficult for upload/download into/out of the Grid to local resources (laptop, instrument, etc.)
- Lightweight file transfer agent (LTA) for firewall-protected resources, ssh-tunnel based technology, access control provided by SAML/XACML – for, e.g., globusonline.org
  - Essentially a lightweight service endpoint for file transfer
  - Presents a straightforward environment for server-side security
  - Timeline:
    - demonstration at SuperComputing 2010
    - delivery: January 2011
    - Adding of HTTP support by March/April 2011
  - May become a feature of GridFTP

Capability: File Transfer Scheduling

- Now: automatic restarts of failed file transfers in GridFTP and globusonline.org
- No near-term plans for versatile scheduling
  - are there any use cases that need this?
  - if there is enough interest, we could implement this

Capability: Parallel Jobs

- Already supported in GRAM5 (MPI)

Capability: Remote Instrumentation

- Globus Use cases:
  - Argonne has advanced X-Ray Photon Source
    - CT scanner attached to resource
    - Large file oriented data sets (this is also a use case for LTA)
  - Remote telescopes operated via Globus in AstroGrid-D

Capability: Workflow

- GridWay supports DAG workflows and the DRMAA standard
- Taverna2-based workflows work with Globus
  - Added GSI security support
  - Can invoke GSI-enabled web services
  - SCUFL engine can be hosted
- Parallel scripting language Swift
  - High throughput use cases
  - Parameter sweeping
  - Tested for machines up to 2000 core systems
  - http://www.ci.uchicago.edu/swift/

Capability: Authentication

- X.509 supported now, incl. delegation, but HTTPG
- Already now support for non-delegation over HTTPS
- Delegation currently requires HTTPG, but currently investigating methods to do this over https
  - Known issues with HTTPS for delegation (1 extra bit)
  - Aiming for HTTPS with proxy certs with a delegation layer, although plain https requires separate delegation service
  - IGE aims to address this from March 2011 onwards. Two possible approaches:
    - Extend GRAM5/GridFTP protocol
    - Remove extra information from GSI header in GSI library
  - Important to interoperate with other EGI technologies where appropriate
- **Medium priority**

Capability: Credential Management

- SLCS and TCS working fine with Globus already
  - Connection to Shibboleth via GridShib in production use in Germany and Switzerland
- MyProxy for proxy renewal
- Kerberos
  - Where is the use case for this in Europe?
  - KCA working at FermiLab: http://security.fnal.gov/pki/new2kcafaq.html
  - Kerberos integrated into MyProxy through PAM module

Capability: User Management

- Envisaged that VOMS capability will come through EMI
  - IGE will use most modern versions of VOMS libraries that support SAML assertions
- AdHoc from GWT is being investigated as GUI front-end to VOMS for easy VO management
- **Medium priority**

Capability: Authorization

- LCAS (AuthZ)/LCMAPS (Account mapping mechanism)
- LCAS uses plugins to base decisions on authenticated credentials and policy
- LCMAPS: once authorised/policy approved, identity/ affiliation within VO is used to determine local account to use
- Globus working towards direct support for LCAS/ LCMAPS
- Formally working towards this integration within IGE
- Used also within UK NGI (National Grid Service)
- **High priority**