UK e-Science

Report on OGSA, OGSI & OGSA-DAI

Malcolm Atkinson

Director of National e-Science Centre
www.nesc.ac.uk

28th October 2002

Meeting of the UK e-Science Technical Advisory Group
Renaissance Hotel
Heathrow, London
Overview

- **OGSA, OGSI @ GGF6**
  - Powerful force & progress

- **DAIS-WG @ GGF6**
  - International, strong UK lead, making good progress
  - Influencing OGSI

- **OGSA-DAI**
  - Phase 1: Complete and successful
    - But worrying structural issues
  - Phase 2: Energetic and effective development collaboration
    - But unresolved structural issues

- **Take home message**
  - Address these structural issues
OGSA & OGSI prognosis

- Powerful rallying cry
- Well organised
  - OGSI, OGSA, ...
- Responsive to DAI
- Substantial industrial investment
- Multi-national intensely active working groups
- IBM announcement of contribution of its OGSI code to public source
- Technical Previews
- Alpha 15th January ‘03

Still gets my vote

Sustain your investment

- Major engineering effort
  - Complex & large code base
  - More demanding of software environment
  - Web services
  - Messaging
  - Depends on planned extensions to standards
  - First APIs for Java
GGF DAIS WG

- **Chairs**
  - Norman Paton (Manchester Uni.)
  - Leanne Guy (CERN) ... dropped out
  - Dave Pearson (Oracle UK)

- **Activity**
  - BoF GGF4 Toronto
  - WG Meeting GGF5 Edinburgh
  - Papers for GGF6
  - Workshops & Mail lists

- **Goals**
  - Agree Standards for Database Access & Integration
  - Freely available reference implementations
    - OGSA-DAI one source & focus for discussions

- **Data Provenance and Derivation Workshop**
  - Significant UK contribution

http://www.cs.man.ac.uk/grid-db/
The OGSA stake holders

Scientific Users

Keep all the stake holder groups satisfied

Owners

Data & Compute Resources

Grid Planning Infrastructure

Scheduling

Authorisation

Monitoring

Diagnosis

Logging

Accounting

Scheduling

Grid Plumbing & Security Infrastructure

Application Developers

Distributed

Operations Team

Scientific Application
OGSA-DAI Stake Holders

Data Intensive Users

Data Intensive Applications

Scientific Data Mining & Integration Technology

Monitoring

Scheduling

Accounting

Grid Plumbing & Security Infrastructure

App. Developers

Tech. Developers

Operations Team

Owners

Data & Storage Resources

Distributed

Data Providers

Data Curators

Keep all the stake holder groups satisfied
Data Access & Integration

- Central to e-Science
  Astronomy, Earth Sciences, Ecology, Biology, Medicine, ...

- Collaboration
  - Shared Databases
  - Curated Knowledge
  - Accumulated Observations
  - Accumulated Simulations

- Computation
  - Data mining
  - Input to models
  - Calibration of models

- Presentation
  - Publication of results
  - Visualisation
OGSA-DAI project

- Lego kit for Data Access & Integration
  - Components for e-Science Applications
  - Accelerated Application Development
  - Multiple Data Models
  - Distributed Data
  - Access via Grid & Proxies
  - Integration, Translation & Transformation

- Open Source Reference Implementation
  - For DAIS-WG standard

- Trigger for DAI Component Construction
  - Stimulate a community
OGSA-DAI Time Line

WS + GSI UK support (> 100 downloads)
XML + OGSA **Prototypes** for Early Adopters

Design Documents & Demos for DAIS WG @ GGF5
XML + OGSA **Prototype** Available
RDB + GT2 / OGSA **Prototypes** Available
GGF6 WG Papers & Prototypes
Ship **Release 1** for GT3 Integration
Demo & **Release 1.5** @ GGF7

**Phase 1 Starts**
Feb ’02 May ’02 Jul ’02 Sep ’02 Dec ’02 Feb ’03 May ’03 Sep ’03

**Phase 2 Starts**
**Release 2**

**today**
Primary Components

Client
Consumer
GDS
DB
GDSF
GDSR
Composing Components

OGSA-DAI Component

Data Transport

OGSA-DAI Component

Data Transport

OGSA-DAI Component

Data Transport
Distributed Query

DQP : Distributed Query Processor
GDT : Grid Data Transport
T : Translation
Q : Query
GDTV : Grid Data Transport Vehicle
F : Factory
QPM : Query Progress Monitor
PNM : Progress Notification Message
AM : Application Metadata
CRM : Computational Resource Metadata
NS : Notification Sink
Today we have

- A well-developed architecture
  - Issues identified and prioritised
- An emerging standard spec.
- A high level design
- Active cross-site development teams
- A functional spec for release 1
- A plan for the remainder of the project
- Effective operational arrangements for development

- Architecture needs rewriting
  - Doc. Structure
  - Emerging ideas
- Standard needs negotiation and to be led from reference implementation
- High-level design still being “polished”
- Resources not fully committed by managers
- Development resources still being set up
Today we need

- Commitment of and by Managers
  - Members of teams allocated for an agreed period
  - Managers’ and leaders’ time and schedule committed
- Completion of collaboration agreement
  - Sign off source licence
  - IPR and liability agreement
- Agreed strategy for Globus copyright
  - General issue for many UK projects
- Agreed plan for UK hand over and support
Take Home Message

- Developers working hard and well
  - Quality of input and work very good
  - Good quality team
- Plans, designs, proto-products emerging
- Committed to release to GT3 Alpha

Unresolved high-level issues are major risk
- Could block release
- Could disrupt effective teams
- Could leave users without support
OGSA-DAI Summary

- On Schedule & Going Well
- Contributions via DAIS-WG @ GGF5 & 6
- Releases with GT3 Releases scheduled
- Status: Early Days
  - Released prototypes
  - Tested Architectural Design
  - Using OGSA
  - Working with Early Adopter Pilot Projects
    - AstroGrid & MyGrid
  - First PRODUCT release Dec ‘02
- Influence OGSA-DAI direction
  - Via DAIS-WG & Direct messages to us
OGSA-DAI Partners

£3 million, 18 months, started February 2002

EPCC & NeSC
IBM UK
IBM USA
Manchester e-SC
Newcastle e-SC
Oracle
Composed Components

Translation

GDS

Client

GDS:performScript

GDS:performScript

GDS:performScript

GDS:performScript

GDS:performScript

GDT

Translation

GDT

Consumer
DAI Key Components

GridDataService (GDS)  Access to data & DB operations
GridDataServiceFactory (GDSF)  Makes GDS & GDSF
GridDataServiceRegistry (GDSR)  Discovery of GDS(F) & Data
GridDataTranslationService  Translates or Transforms Data
GridDataTransportDepot (GDTD)  Data transport with persistence

Relational & XML models supported
Role-based Authorisation
Binary structured files
## OGSA Relationship

<table>
<thead>
<tr>
<th>Class</th>
<th>GridService</th>
<th>Registry</th>
<th>NotificationConsumer</th>
<th>NotificationProducer</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDS</td>
<td>Mandatory</td>
<td></td>
<td>Optional</td>
<td>Normal</td>
</tr>
<tr>
<td>GDSF</td>
<td>Mandatory</td>
<td></td>
<td>Optional</td>
<td>Normal</td>
</tr>
<tr>
<td>GDSR</td>
<td>Mandatory</td>
<td>Mandatory</td>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td>GDTS</td>
<td>Mandatory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDTD</td>
<td>Mandatory</td>
<td></td>
<td>Optional</td>
<td>Normal</td>
</tr>
</tbody>
</table>
## DAI portType Usage

<table>
<thead>
<tr>
<th>Class</th>
<th>GridDataService</th>
<th>DataTransport</th>
<th>Factory</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDS</td>
<td>Mandatory</td>
<td>Normal</td>
<td></td>
</tr>
<tr>
<td>GDSF</td>
<td>Optional</td>
<td>Normal</td>
<td>Mandatory</td>
</tr>
<tr>
<td>GDSR</td>
<td>Optional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDTS</td>
<td>Optional</td>
<td>Mandatory</td>
<td></td>
</tr>
<tr>
<td>GDTD</td>
<td>Optional</td>
<td>Mandatory</td>
<td></td>
</tr>
</tbody>
</table>