OGSA-DAI & DAIT projects

Update for TAG

www.ogsadai.org.uk

Prof. Malcolm Atkinson
Director

www.nesc.ac.uk

30th October 2003
Contents

- OGSA-DAI Product report
- OGSA-DAI Project report
- DAIT Project report
- OGSA-DAI Product road map
OGSA-DAI Product

- **Brand name:** OGSA-DAI
  - Continues as established
- **Current release 3.0.2**
  - OGSA-DAI: 1183 downloads (461 R3 & R3.0.2)
    - >379 in UK
  - 50 downloads by users of R3.0.0 of R3.0.2 within a week
  - Recent performance analysis ⇒ R3.0.3 Nov 03
  - DQP prototype: 77 downloads
    - Since 1st September 2003
- **Web site**
  - 471 registered users
  - Transfers to users gigabytes / month
    - Aug 1.38GB  Sep 1.63GB  Oct 1.81GB
Downloads By Country - All Releases

- United Kingdom: 417 downloads
- United States: 220 downloads
- China: 198 downloads
- Japan: 150 downloads
- Germany: 42 downloads
- Austria: 30 downloads
- Korea, Republic of: 29 downloads
- Brazil: 198 downloads
- India: 198 downloads
- Canada: 198 downloads
- Hungary: 150 downloads
- Sweden: 198 downloads
- Australia: 198 downloads
- Switzerland: 198 downloads
- Italy: 198 downloads
- Taiwan: 198 downloads
- France: 198 downloads
- Poland: 198 downloads
- Netherlands: 198 downloads
- Romania: 198 downloads
- Russian Federation: 198 downloads
- Singapore: 198 downloads
- Ireland: 198 downloads
OGSA-DAI Project

- EPCC Team continued to 31st Oct 03
  - Then continues as the DAIT R&D team
- Globus Alliance Formed
  - Edinburgh, PDC, Chicago University & University of Southern California
    - 1 Sept 03
  - Collaboration meetings
    - Agreeing schedule, standards, procedures for OGSA-DAI ↔ GT3.4 source integration
- GGF DAIS WG — major investment — UK led
  - 3 face-2-face meetings, > weekly telcons & authoring for GGF9
- ODD-Gene Demo
  - http://www.epcc.ed.ac.uk/~oddgenes/
DAIT Project

- **Product / Brand remains**: OGSA-DAI
- **Planning meeting at NeSI 20th October**
  - ESNW, IBM & NEReSC by Access Grid
- **4 Papers produced Public after**
  - Release 4 (April 04) Specification
  - OGSA-DAI road map
  - Project management structure revision
  - Draft research plan

- **DAIT Contract Status**
  - Completed EPSRC forms submitted to DTI
  - No information yet on how to proceed contractually
    - No comfort letters
    - We urgently need contracts
DAIT Project Status

• Partners formally proceeding from 1 Oct. 03
  • EPCC charging from 1\textsuperscript{st} November 03
    ▶ Team currently continues unchanged – \textit{working well}
  • NeSC PDRA not yet advertised
  • ESNW PDRA appointed on \textit{other} funds
  • NEReSC PDRA not yet advertised
  • IBM financial planning processes underway
    ▶ Expectation of engagement continuing
    ▶ But not yet formally agreed or quantified
  • ORACLE will remain active in standards

• Project transition meeting OGSA-DAI $\rightarrow$ DAIT
  • 25\textsuperscript{th} November 2003
OGSA-DAI road map 1

- **R3.1.0** Jan 04  Tech. Preview part of R4
- **User Group: inaugural meeting** Q1 04
- **R4.0.0** April 04
  - Performance & monitoring
  - Additional DBMS’s supported
  - Additional SQL supported
  - DBMS management operations
    - archive, restore, bulk load
  - File access
  - Client libraries
  - Installation wizard
  - User support, courses, training material, performance report
OGSA-DAI road map 2

- **R5 October 04**
  - Compliance with DAIS standards proposal
  - Distributed Relational Query Processing
  - Improved dependability and security integration
  - Extended & integrated XML and relational facilities
  - Distributed transaction participation
  - Coordinated OGSA-DAI contributor community

- **R6 April 05**
  - Integrated with GT3
  - New facilities depend on user priorities, context and research
  - OGSA-DAI components from contributor community

- **R7 October 05**
  - Maintainable release for the user community
Preliminary Organisation

OGSA-DAI Phase 3
Project Structure

Programme Board
Tony Hey, eBRC (chair)
Jim Fleming, EPSRC
Malcolm Atkinson, NeSC
Norman Paton, ESRNW
Paul Watson, NEReSC
Arthur Trew, EPCC
Dave Pearson, Oracle
Patrick Darnessingle, IBM
David Snelling, Fujitsu

Technical Review Board
Malcolm Atkinson
Rob Baxter
Neil Chue Hong
Simon Laws
Norman Paton
Paul Watson

Malcolm Atkinson
PI, NeSC

Norman Paton
PI, Manchester

Paul Watson
PI, Newcastle

Neil Chue Hong (0.5)
Project Manager

DAIS-WG

User Group

BBSRC
OGSA-DAI

OMII

Globus Alliance

Research Team
Desmond Fitzgerald
Manchester PDRA

Mario Antonioletti
System Architect
Neil Chue Hong (0.25)
Software Developer
Mike Jackson (0.75)
Software Developer
Amy Krause
Software Developer
Tom Sugden
Software Developer
Martin Westhead (0.5)
Software Developer

Development team
automatically store data into database. MEG is a device which can detect the change in minute magnetic fields generated from the brain activity.

**Design Strategy**

Resources of neuroinformatics such as data, tools and models are located in various research institutes. We focus on metadata-driven access and service-oriented grid environment as key technologies for organic linking of various resources.

**Metadata-driven access**

Metadata takes a central role in our framework. All the data, tools and models are explained with metadata formatted as XML document. Users’ requests for computation are processed by Portal Service, which retrieves information about appropriate resources and organizes them.

**Service-oriented grid environment**

In widely distributed grid environment, data and metadata should be provided through commonly accepted interfaces. Currently we are examining Globus Toolkit 3 and OGSA-DAI in terms of their functionality, working speed and so on.

Acknowledgement: This work was supported in part by a Grant-in-Aid for Scientific Research on the Priority Area, "Informatics Studies for the Foundation of IT Evolution" (13224059) by the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the IT-program (Construction of Supercomputer Network) of the Ministry of Education, Culture, Sports, Science and Technology.

Researchers: Takahiro Kosaka (tak-k@ais.cmc.osaka-u.ac.jp) 
Susumu Date (date@ais.cmc.osaka-u.ac.jp) 
Yuko Mizuno-Matsumoto (yuko@ais.cmc.osaka-u.ac.jp) 
Shinji Shimojo (shimojo@cmc.osaka-u.ac.jp)

Supported by:
OGSA-DAI is the retained *product* brand
- There is growing user community engagement
- And evidence of use

The OGSA-DAI $\Rightarrow$ DAIT *project* transition
- Slightly delayed
- Uncertainties about contracts & industry remain

But work on the *product* and standards
- Continues at a steady pace
- With enhancements being delivered to users

Support (part of GSC) is vital
Comments & Questions

www.ogsadai.org.uk
Reserve & Informational Slides

www.ogsadai.org.uk
Data Services

- GGF Data Access and Integration Svcs (DAIS)
  - OGSI-compliant interfaces to access relational and XML databases
  - Needs to be generalized to encompass other data sources (see next slide...)

- Generalized DAIS becomes the foundation for:
  - Replication: Data located in multiple locations
  - Federation: Composition of multiple sources
  - Provenance: How was data generated?
“OGSA Data Services” (Foster, Tuecke, Unger, eds.)

- Describes conceptual model for representing all manner of data sources as Web services
  - Database, filesystems, devices, programs, ...
  - Integrates WS-Agreement
- Data service is an OGSI-compliant Web service that implements one or more of base data interfaces:
  - DataDescription, DataAccess, DataFactory, DataManagement
- These would be extended and combined for specific domains (including DAIS)
OGSA-DAI Approach

- Reuse existing technologies and standards
  - OGSA, Query languages, Java, transport
- Build portTypes and services which will enable:
  - controlled exposure of heterogenous data resources on an OGSI-compliant grid
  - access to these resource via common interfaces using existing underlying query mechanisms
  - (ultimately) data integration across distributed data resources
- OGSA-DAI (the software) seeks to be a reference implementation of the GGF DAIS WG standard
  - Can’t keep up with frequent standard changes, so software releases track specific drafts
- See http://www.ogsadai.org.uk/ for details.
1a. Request to Registry for sources of data about “x”

1b. Registry responds with Factory handle

2a. Request to Factory for access to database

2b. Factory creates GridDataService to manage access

2c. Factory returns handle of GDS to client

3a. Client queries GDS with XPath, SQL, etc

3b. GDS interacts with database

3c. Results of query returned to client as XML

SOAP/HTTP

service creation

API interactions

Data Access & Integration Services
Third Party Delivery

1. Requestor Stub
2. Data Set
3. Data Set
4. Consumer Stub

CLIENT API
REQUESTOR STUB

dr

CLIENT API

Data Set

Data Set