

# A new Architecture for OGSA-DAI

Malcolm Atkinson, K. Karasavvas, M. Antonioletti,  
R. Baxter, A. Borley, N. Chue Hong, A. Hume,  
M. Jackson, A. Krause, S. Laws, N. Paton, J. M. Schopf,  
K. Turlas and P. Watson

19<sup>th</sup> September 2005



## Contents: What we talk about

- Background
- Requirements
- Overview of new Architecture
- Status and Future Work



# Background

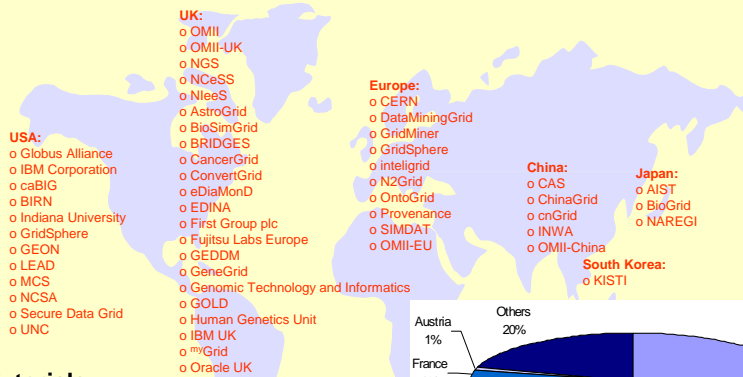
- **OGSA-DAI**

- **Middleware to provide uniform *extensible* access to data resources**
  - ▶ Relational, XML, Files and Indexed Files
  - ▶ Translation and delivery activities
- **Old releases**
  - ▶ Globus Toolkit 3 - OGSF
- **New releases**
  - ▶ Axis - WS-I
  - ▶ OMII - WS-I+
  - ▶ Globus Toolkit 4 - WSRF
- **Experience**
  - ▶ Previous assumption: web services are lightweight
  - ▶ Factory pattern: making a home for session state
  - ▶ Cannot rely on all data resources having sufficient functionality

## Old Architecture

- **One GDS per client interaction**
- **One Data Resource per GDS**
- **No explicit Sessions**
- **No explicit Transactions**
- **Depend on data resources for concurrency management**

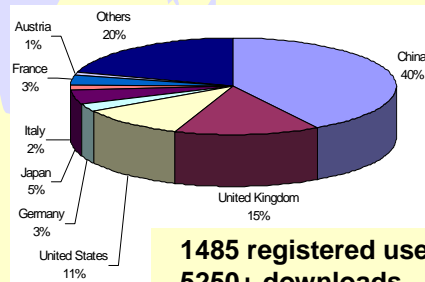
# International Collaboration & Use



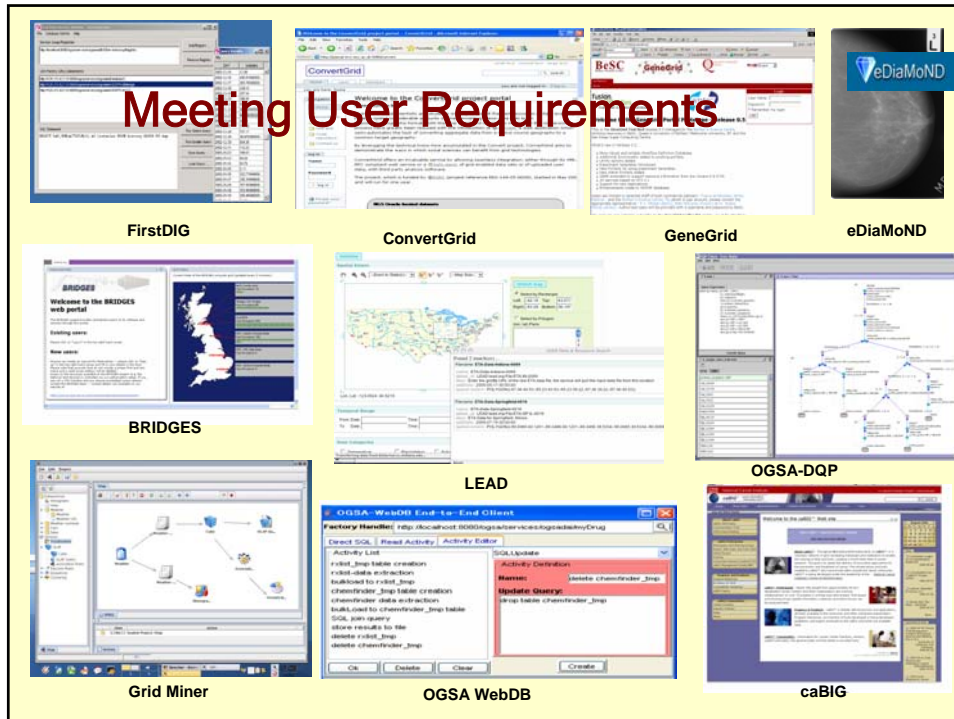
## Tutorials

Boston  
 CERN  
 Edinburgh  
 San Francisco  
 Seoul  
 Tokyo

Cambridge  
 Chicago  
 London  
 Seattle  
 Singapore  
 ISSGC 03 to 05

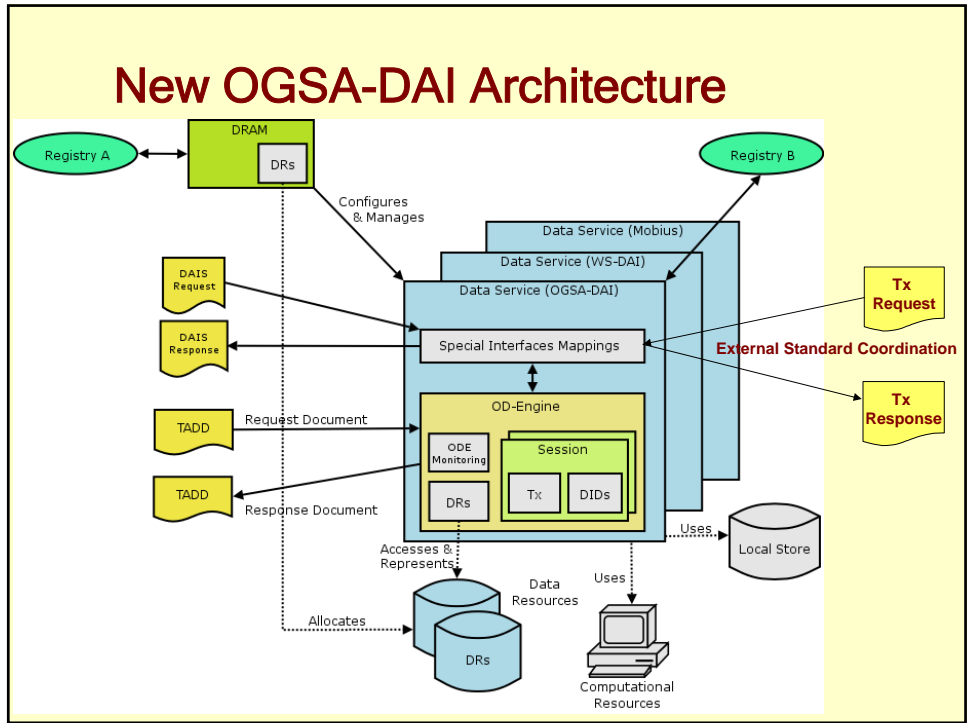


**1485 registered users**  
**5250+ downloads**



## Requirements

- **Web Services now long-running**
  - Manage state lifetime within WS
  - Manage concurrency within WS
  - Manage resources within WS
- **Continued support of functionality**
  - All previous tasks
  - All previous extensibility
  - Single task & composition
- **Accommodate small and *large* tasks**
  - Standard activities
- **Manageable and Configurable**
  - Monitoring and control
- **Persistent & resilient to failures**
  - Persistent operational state
  - Recovery
- **Performance**
  - Streamed pipelines of tasks
- **Concurrent sessions & transactions**
  - Standard coordination protocols
  - Embedded in request documents
  - Multi-request sessions
- **Extensible framework & Increased Power in Requests**
  - Application-defined activities
  - Additional data sources
  - Extensible client libraries
- **Standard interfaces - Preserve application investment**
  - Client library stability
  - Activity API stability
- **Naming**
  - Multiple Data Resources per Data Service
  - Results, streams, sessions, transactions, ...



## Components (1)

- **Data Resources (DRs)**
  - **Currently works with**
    - Relational, XML DBMS, file systems & indexed files
  - **Extensibility point**
    - Any kind of data resources can be added
    - Need not be a single resource - could be a federation
- **Data Services**
  - **Zero or more DRs**
  - **Client Toolkit**
    - Protects developers from changes
    - Common API for all OGSA-DAI platforms

8

## Components (2)

- **Activities (and Tasks)**
  - **Activity is the unit of work**
    - ▶ E.g. SQL or X\* statement, translation, delivery
    - ▶ Types: Core, Supplied and Extension activities
  - **Extensibility point**
- **TADD, Requests and Results**
  - **Composite format**
    - ▶ Reduce round-trip latencies, status & results, delegation of work
  - **Control flow**
  - **Client toolkit**

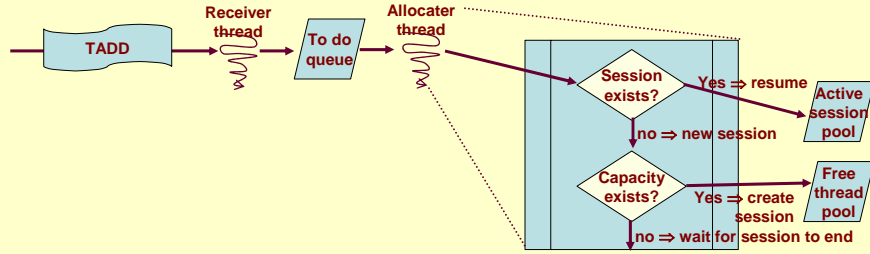


## Components (3)

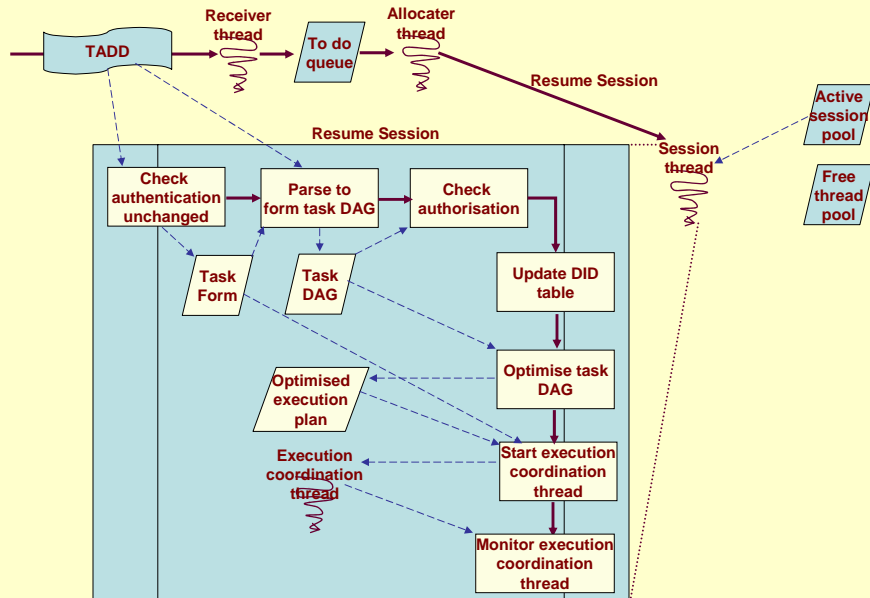
- **Sessions**
  - **Sharing and managing state between TADDs**
    - ▶ External decisions, retain results for collection/delivery, transactional and security information
  - **Implicit, explicit and new sessions**
- **Transactions**
  - **Recovery from failures**
  - **Consistent transactional framework**
    - ▶ Transactional metadata for activities
    - ▶ Logical rules for combining the behaviour of transactional activities
- **Data Identifiers (DIDs)**
- **OGSA-DAI Engine (ODE)**

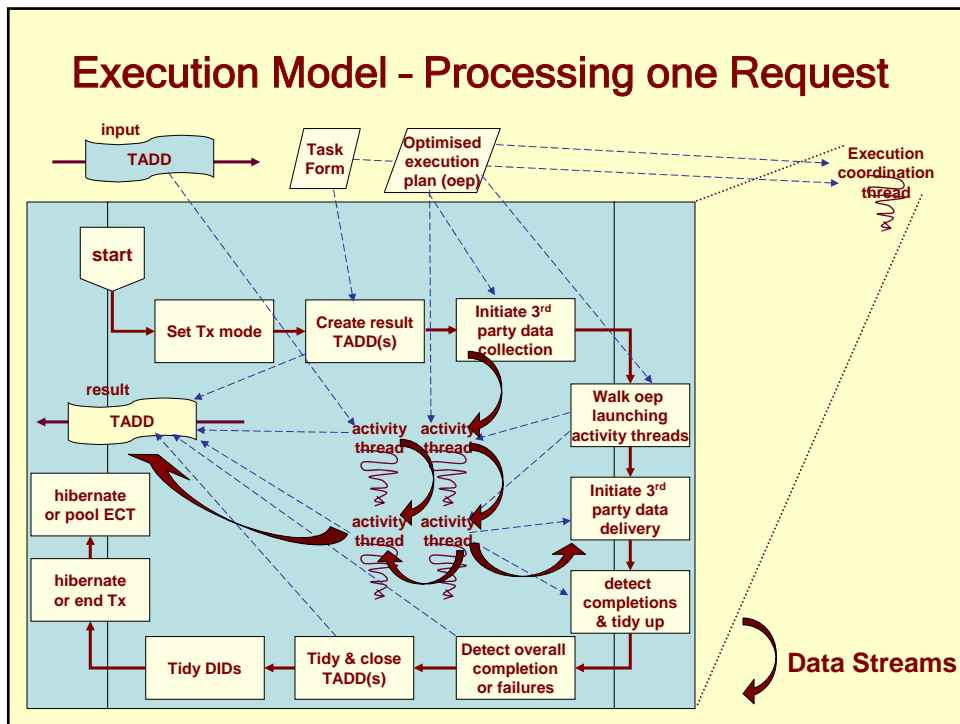


## Execution Model - Workload Throttle & Session manager




## Execution Model - Execution Planning







## Components (4)



- **Special Interfaces**
  - WS-DAI
  - Transaction coordination protocols
    - ▶ WS-AtomicTransactions
- **Data Resource Access Manager (DRAM)**
  - Configuration and Monitoring
  - Gateway to a set of DRs - via a registry
  - Data Service Description Language (DSDL)
- **Naming**
  - Accommodate naming schemes
  - Multiple DRs - TADDs contain DR name



14

## Status and Future Work

- **Release 7 - Partial move to new architecture**
  - New DR extensibility point
  - New activities, e.g. multi-source union
  - Reconfigurable services
  - Concurrency within data service
  - Sessions
  - Parts of DQP
- **DQP Release 3**
- **Releases for Friends & Family**
  - Transactions (tech. preview of phases)

## Status and Future Work 2

- **Following releases**
  - **Performance**
    - ▶ Execution planning, concurrency, streamed pipelines, better data transfer
  - **Functionality**
    - ▶ Completed Transactions
    - ▶ Distributed queries
    - ▶ Resilient multi-source unions
    - ▶ Schema and data integration
    - ▶ More data extraction from files & files in transactions
  - **New extensibility & configuration**
    - ▶ Activity database
    - ▶ Dynamically installing extension activities



## Further information

- **The OGSA-DAI Project Site:**
  - <http://www.ogsadai.org.uk>
- **The DAIS-WG site:**
  - <http://forge.gridforum.org/projects/dais-wg/>
- **OGSA-DAI Users Mailing list**
  - [users@ogsadai.org.uk](mailto:users@ogsadai.org.uk)
  - General discussion on grid DAI matters
- **Formal support for OGSA-DAI releases**
  - <http://www.ogsadai.org.uk/support>
  - [support@ogsadai.org.uk](mailto:support@ogsadai.org.uk)
- **OGSA-DAI training courses**

**Questions  
Please**