OGSA-DAI

Policy for re-use

- This presentation can be re-used for academic purposes.

- However if you do so then please let training-support@nesc.ac.uk know. We need to gather statistics of re-use: no. of events, number of people trained. Thank you!!

Acknowledgments

- Matt Ford, NGS Induction Workshop (Dec. 2004, NeSC)

- Neil Chue Hong , OGSA-DAI Tutorial GGF13

- OGSA-DAI website, www.ogsadai.org

Data services on NGS

Simple data files

- Middleware supporting
  - **Replica files**
  - **Logical filenames**
  - **Catalogue**: maps logical name to physical storage device/file
  - **Virtual filesystems**, POSIX-like I/O

- **Storage Resource Broker**

Structured data

- RDBMS, XML databases

- Require *extendable* middleware tools to support
  - Move computation near to data
  - easy access, controlled by AA
  - integration and federation

- **OGSA -DAI**

What is OGSA-DAI?

- The Open Grid Services Architecture Data Access and Integration project is concerned with constructing middleware to assist with access and integration of data from separate data sources via the grid.

- The project was conceived by the UK Database Task Force and is working closely with the Global Grid Forum DAIS-WG and the Globus team.
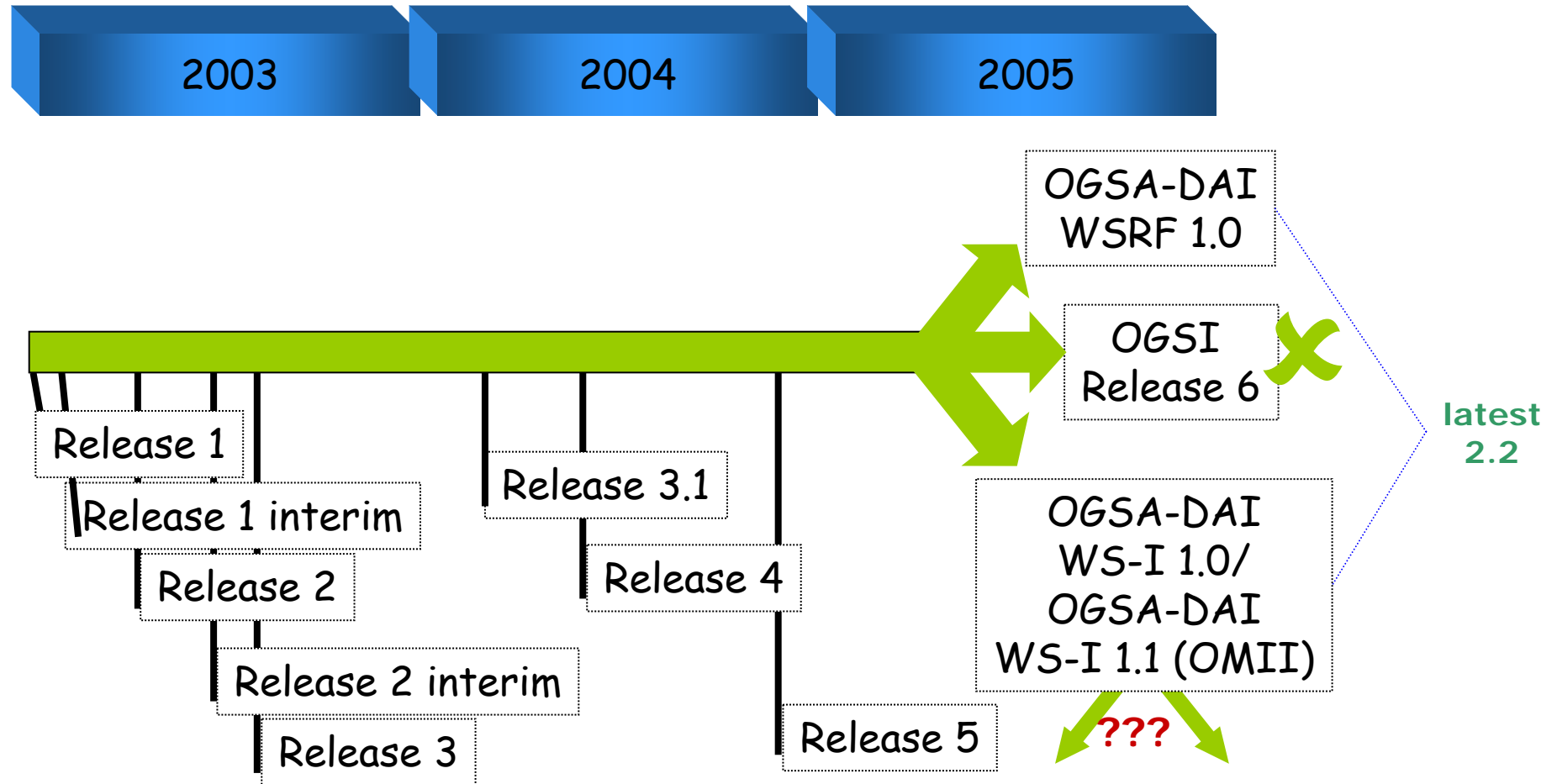
OGSA-DAI Design Principles – I

- Efficient client-server communication
  - Minimise where possible
  - One request specifies multiple operations

- No unnecessary data movement
  - Move computation to the data
  - Utilise third-party delivery
  - Apply transforms (e.g., compression)

- Build on existing standards
  - Fill-in gaps where necessary

OGSA-DAI Design Principles – II

- Do not hide underlying data model
  - Users must know where to target queries
  - Data virtualisation is hard
- Extensible architecture
  - Modular and customisable
  - e.g., to accommodate stronger security
- Extensible activity framework
  - Cannot anticipate all desired functionality
  - Activity = unit of functionality
  - Allow users to plug-in their own

OGSA-DAI Timeline

NGS National Grid Service

2003	2004	2005

OGSA-DAI WSRF 1.0

OGSI Release 6

Release 1

Release 1 interim

Release 2

Release 3.1

Release 2 interim

Release 4

Release 3

Release 5

OGSA-DAI WS-I 1.0/ OGSA-DAI WS-I 1.1 (OMII)

???

latest 2.2

8

OGSA-DAI Motivation

- OGSA-DAI is motivated by the need to:
    - Provide an extensible framework for easily integrating data resources on to Grids.
    - Provide for data discovery from previously unknown locations.
    - Allow different types of data models from distributed data resources to be easily integrated to Grid applications.
    - Allow data to be accessed through uniform interfaces.
    - Facilitate the integration of data from various sources to obtain the required information.
   - …

OGSA-DAI Provides

- Access to and updating of data resources

- Exposure of Data Resources to the Grid

- Additional data manipulation functionality at the service level

- Uniform access to disparate, heterogeneous data resources
  - Does not hide underlying data model

- Data resources exposed through services
 - Clients interact with these services

Interacting with Data Resources

- **Activity**: The data resource manipulation, data transformation and delivery actions that the client wants the service to perform.
  - Think of sending the job to the data not the data to the job.

- **Perform documents**: Used by clients to specify to the services the activities they want executed.
  - Usually don't see/construct explicetly

- **Response documents**: Used by the services to inform clients as to the status of execution of their Perform documents and, often, to also return data to a client.

Interacting with Data Service resources
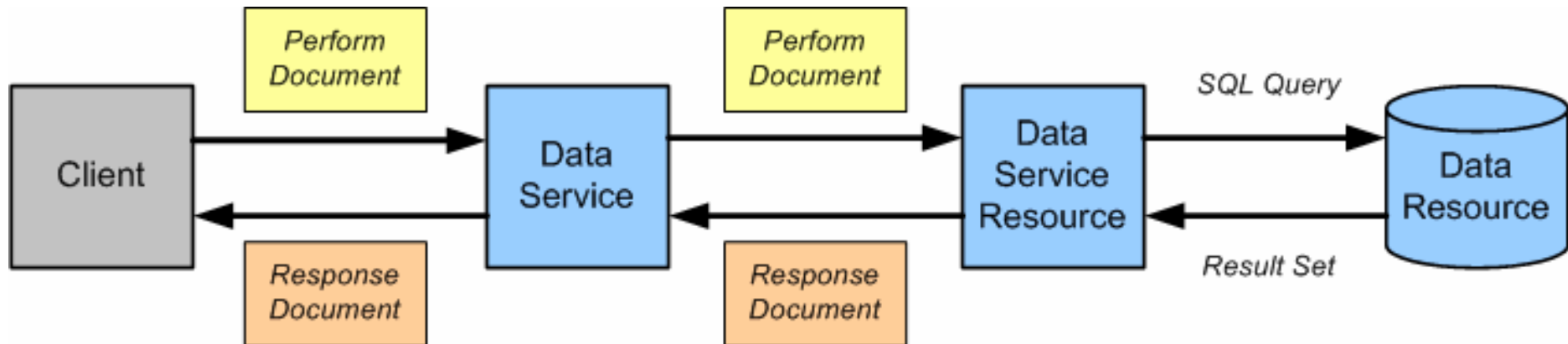
OGSA-DAI Deck of Activities

13
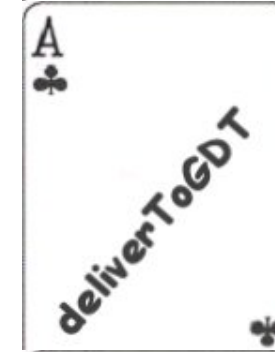
OGSA-DAI and the NGS

- the OGSA-DAI deployment on the NGS is being actively developed

- users should expect that procedures may change – it does not reflect the commitment NGS has to providing a service.

- Initially the Manchester JISC data cluster has been charged with deploying the OGSA-DAI service

Why OGSA-DAI?

- Can embed additional functionality at the service end
  - Transformations, compressions, third party delivery
  - The extensible activity framework
- Avoiding unnecessary data movement
- Common interface to heterogeneous data resources
  - Relational, XML databases, and files
- Language independence at the client end
  - Do not need to use Java
- Platform independence
  - Do not have to worry about connection technology, drivers, etc

Standards Based

National Grid Service

- **OGSA-DAI and DAIS**
  - The *DAIS - Database Access and Integration Services* - Working Group of the *Global Grid Forum* (GGF) is formulating standards for database access and integration services.
  - The development of OGSA-DAI has been occurring in parallel to the development of these specifications. OGSA-DAI influences, and is influenced by, this work. OGSA-DAI is currently based upon the DAIS specifications of March 2003. It is intended that OGSA-DAI will eventually provide a reference implementation of the final version of these standards.
- Further information on the GGF and DAIS are available from:
- DAIS Working Group:
  - http://forge.gridforum.org/projects/dais-wg.
- Global Grid Forum:
  - http://www.ggf.org

Security (1)

- OGSA-DAI uses standard WSRF architecture for security:

  – Transport Level Security

  – Message Level Security

- Either can guarantee privacy, integrity and authenticity

The images in this slide were taken from http://gdp.globus.org/gt4-tutorial/multiplehtml/ch10s02.html

Security (2)

- **Comparison of transport-level and message-level security**

	GSI Secure Conversation	GSI Secure Message (Not used by OGSA-DAI)	GSI Transport
Technology	WS-SecureConversation	WS-Security	TLS (formerly known as SSL)
Privacy (Encrypted)	Yes	Yes	Yes
Integrity (signed)	Yes	Yes	Yes
Anonymous authentication	Yes	No	Yes
Delegation	Yes	No	No
Performance	Good if sending many messages	Good if sending few messages	Best

The Setup (2)

	tc03	tc05	tc07
OGSA-DAI Servers	— LittleBlackBook — Scratch0 — Challenge1	— StaffList — Scratch1 — Challenge2	— Scratch2 — Challenge3 — Challenge4
Additional Servers	Database Server Anonymous FTP Server	HTTP Server Anonymous FTP Server	 Anonymous FTP Server
	OGSA-DAI Client Toolkit (Loaded with "module load java ogsadai_wsrf")		
User Accounts	user01 … user05	user06 … user10	Un-used Today

19