Open Grid Service Architecture - Data Access & Integration (OGSA-DAI)

Dave Pearson
OGSA-DAI Programme Manager
dave.pearson@oracle.com

OGSA-DAI Programme

OGSA-DAI Motivation

- UK e-Science Core Programme
  - Database Taskforce
  - Grid data requirements scoping study
  - Data Access and Integration BoF at GGF4
  - DAIS Working Group in Global Grid Forum
- Exploit emerging Grid computing concepts (OGSA)
- Web Services adoption
- DBMS adoption
- Meet needs of Grid application developers for UK e-Science community
  - Grid Pilot projects

UK Grid Pilot Projects

	Research Focus
myGrid	eScience workbench, Bioinformatics
DAME	Distributed Aircraft Maintenance Engineering
CombiChem	Combinatorial (bio-)chemistry
RealityGrid	Interactive steering of data & compute intensive analysis – chemistry, earth sciences
DiscoveryNet	Bioinformatics, Environmental & Earth sciences
Geodise	Computational Fluid Dynamics
AstroGrid	Virtual Observatory

OGSA-DAI Requirements

• Transient and persistent data	- Data sources & Resources
• Maintain complex data models	- Information/Knowledge
• Record and maintain ‘data about data’	- Metadata
• Establish reliability and quality of data	- Provenance
• Make data more accessible	- Publishing and Discovery
• Restrict who can read and modify data	- Access control
• Receive data in a state ready to use	- Transformation
• Personalise data	- Analysis/Interpretation
• Quality of Service	- Managing data

DAI Programme of Work

- Two phases
  - Phase 1 Feb – Sept 2002 complete
  - Phase 2 Oct 2002 – Jul 2003
  - Maintenance activities
- Funding for next phase approved - DAIT
- Grid data service reference implementations
  - OGSA compliant
  - DB2, MySQL, ORACLE, XINDICE
- Integration with Globus toolkit
- Input to Global Grid Forum on standards recommendations
- Collaborative Programme

Participating Organisations

Edinburgh Parallel Computing Centre

North East eScience Centre

National eScience Centre

North West eScience Centre

IBM UK & IBM US

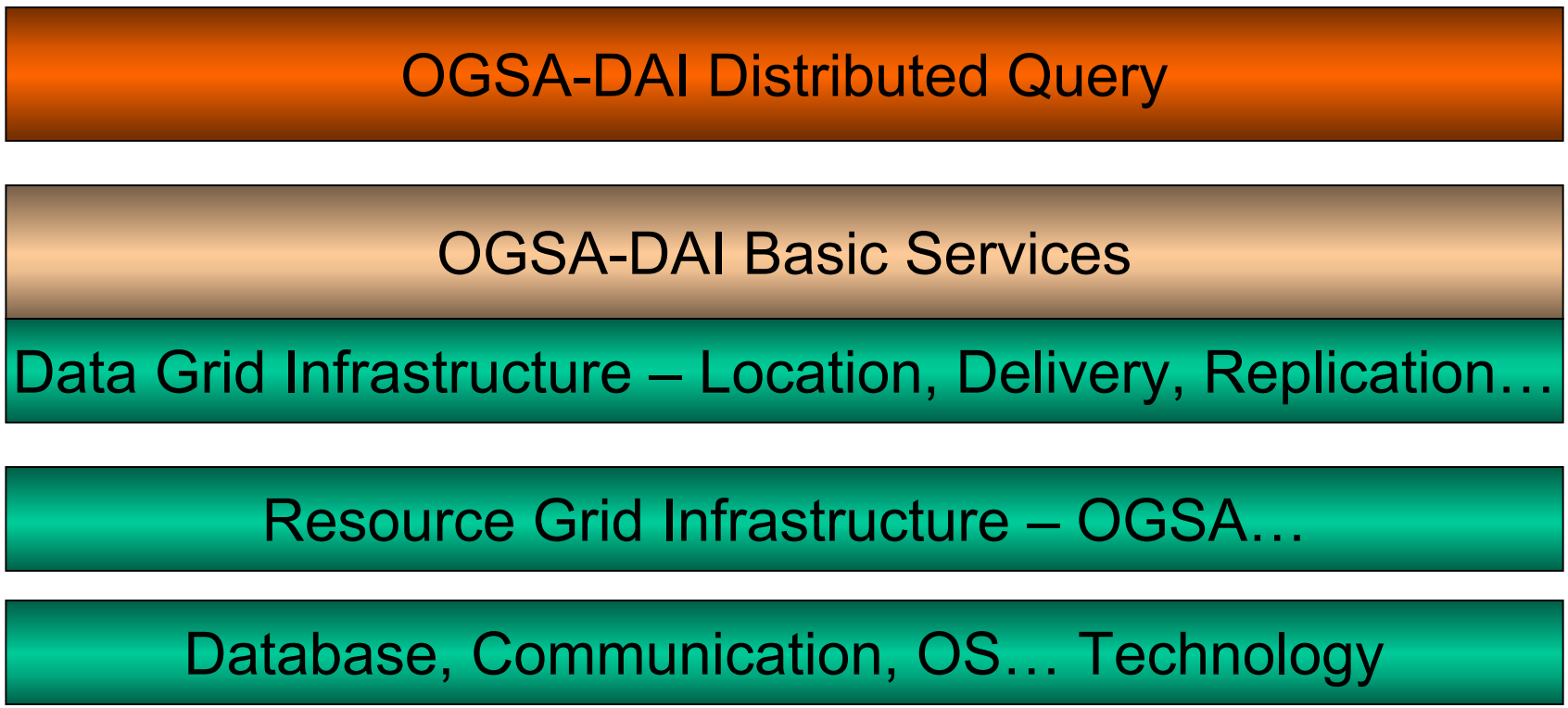
Oracle Corporation UK

Open Grid Services Architecture

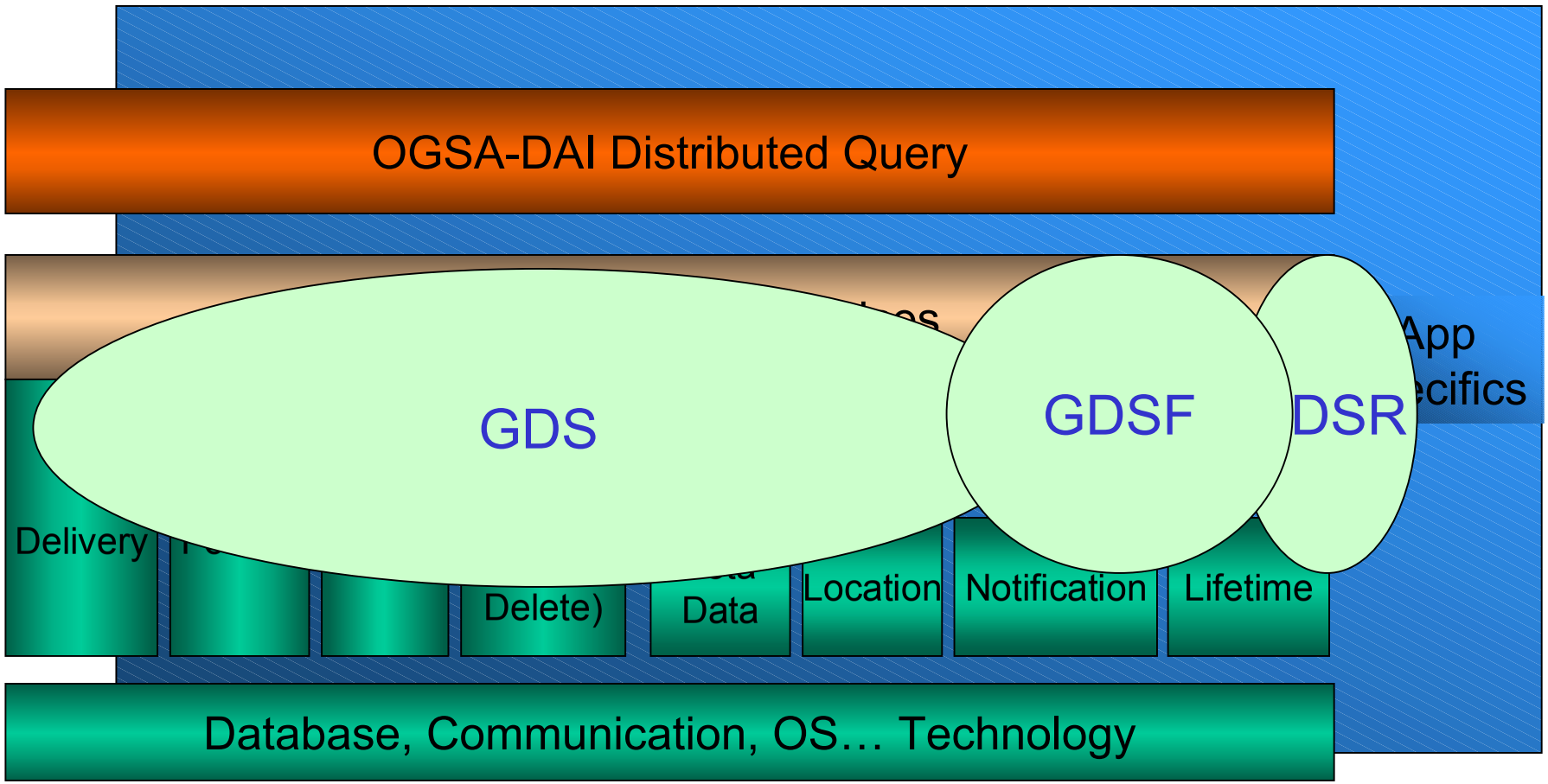
- Web services
  - Standard interface definition mechanisms: multiple protocol bindings, multiple implementations, local/remote transparency
  - Address discovery & invocation of services
- Grid Service extensions
  - Transient services
  - State management in a distributed environment
- Resource virtualisation
- Multiple hosting targets: J2EE, .NET, “C”, …

Data Access and Integration

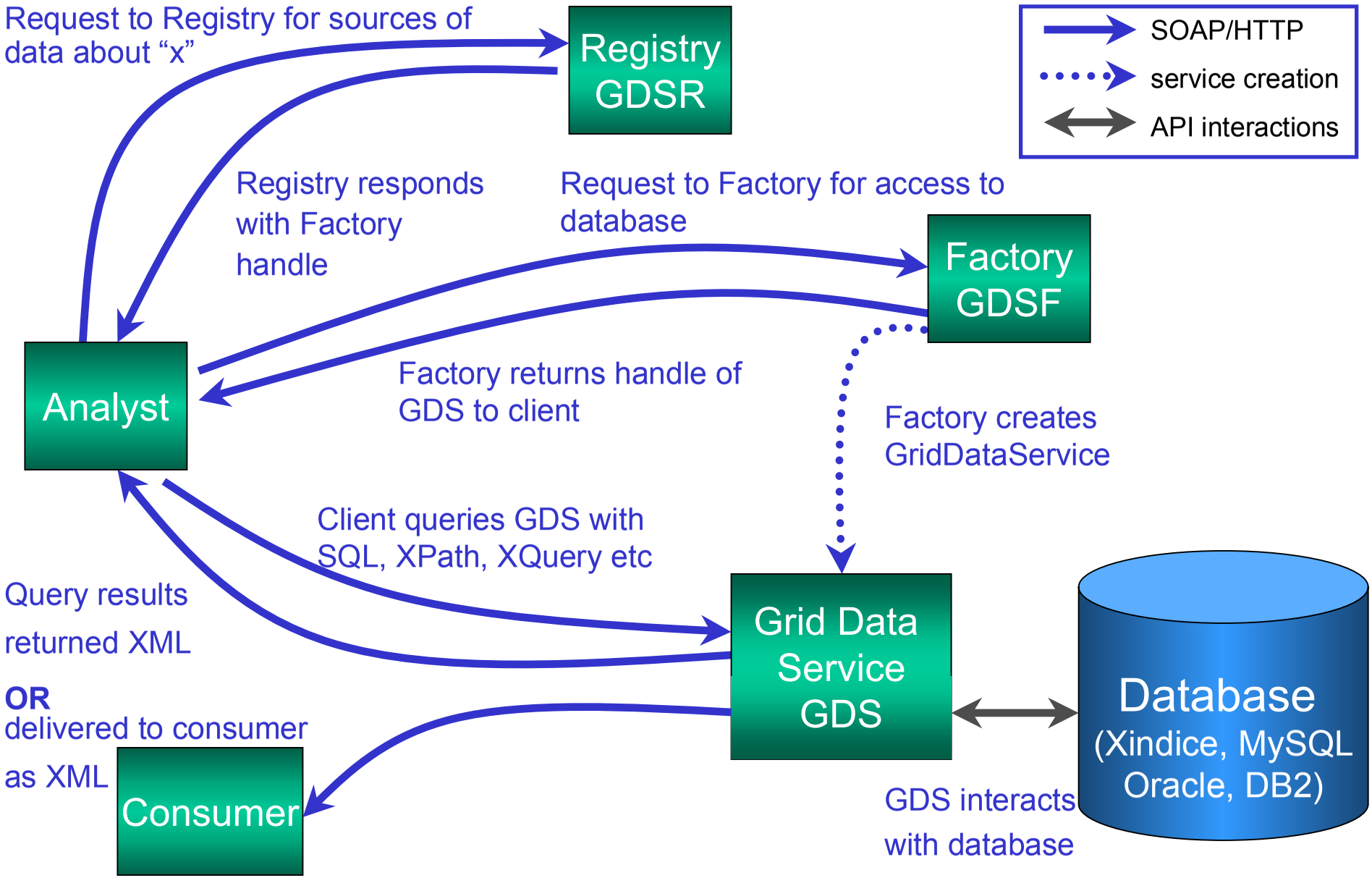
OGSA-DAI Positioning - Vision



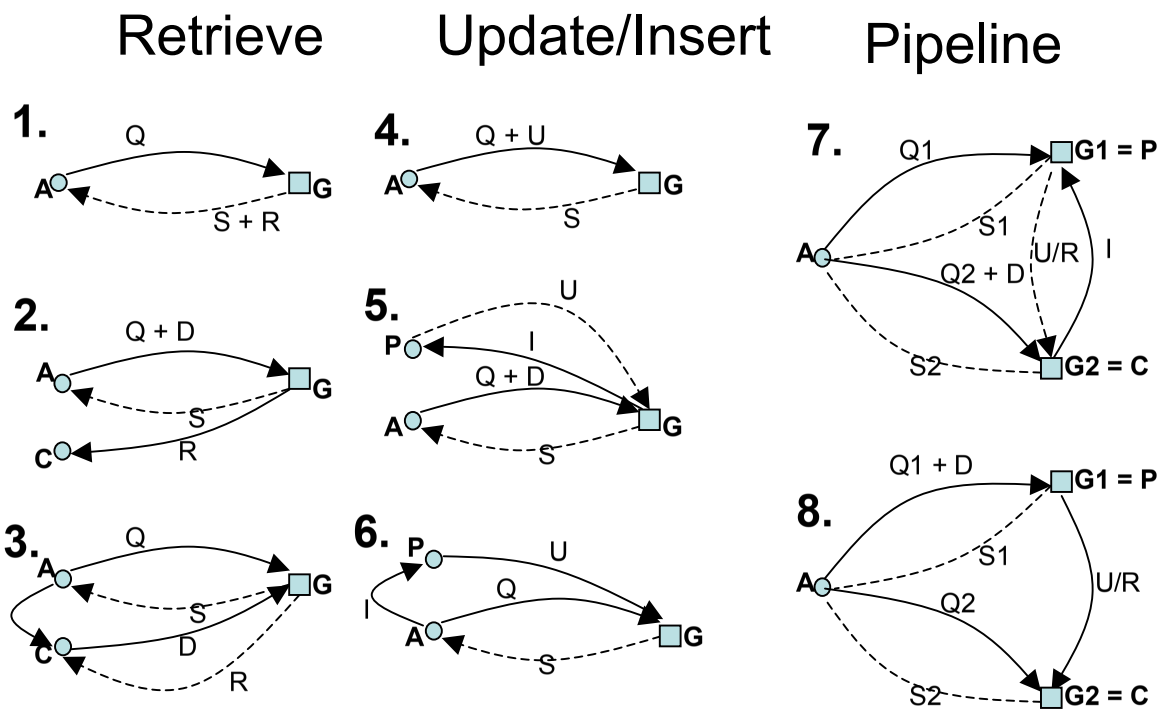
OGSA-DAI Positioning - Today



OGSA-DAI Behaviours

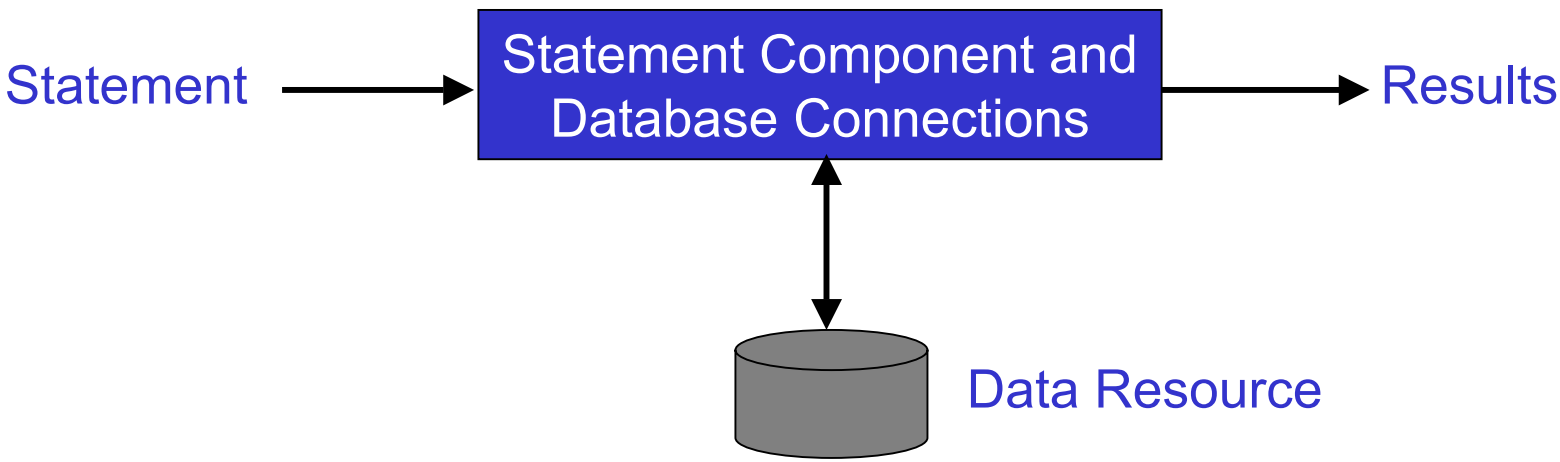


Delivery Patterns



Statement Operations

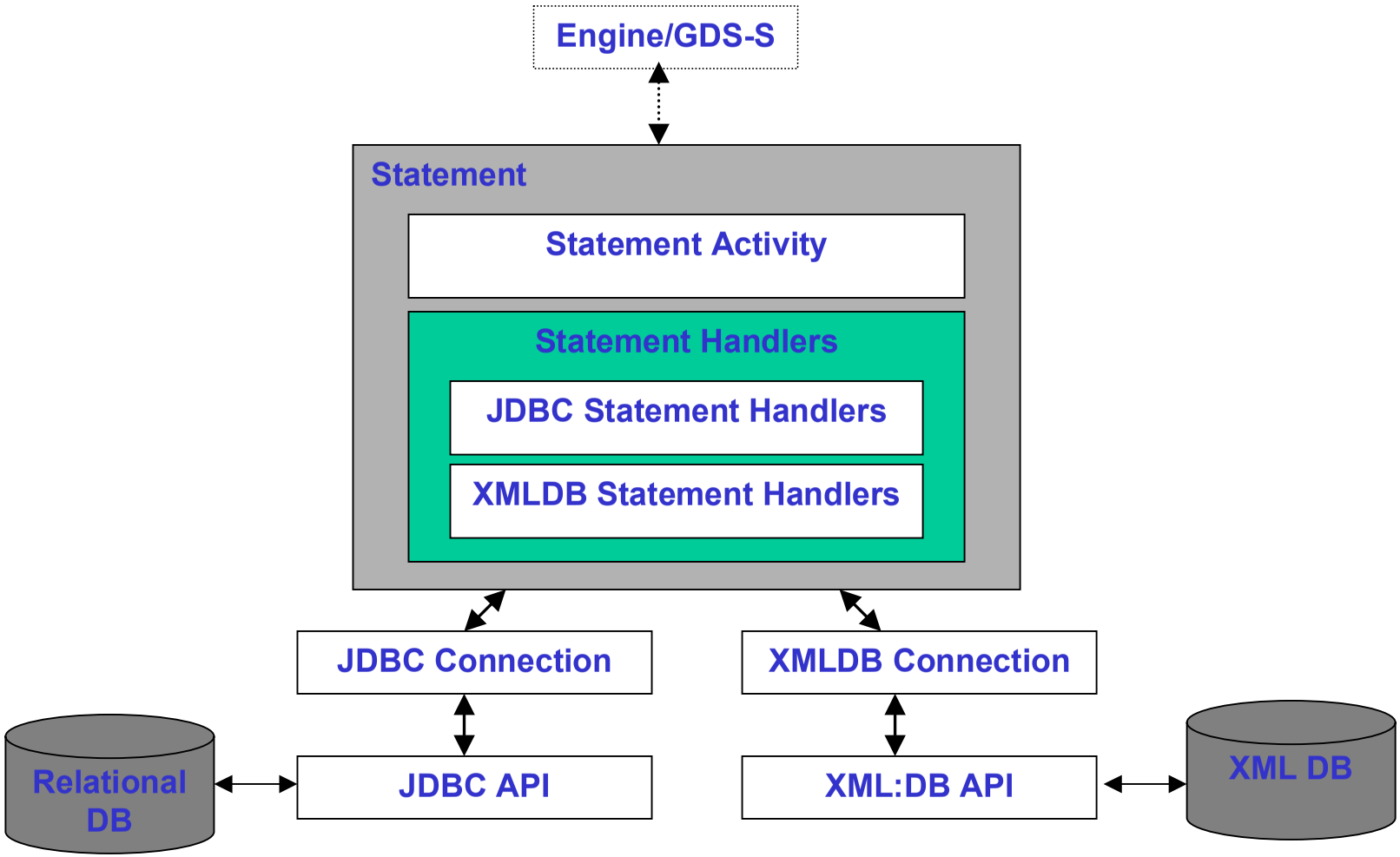
- Receives statement
- Performs statement against a data resource
- Returns results



Statement Document

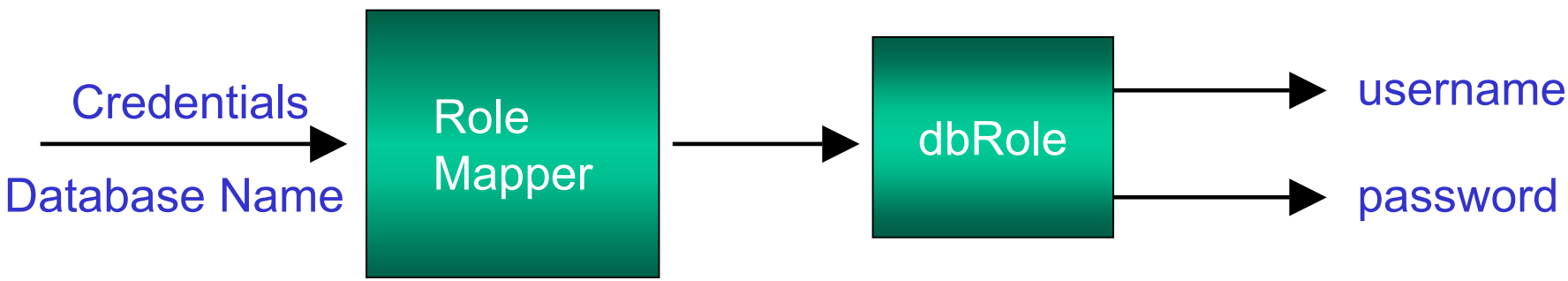
```
<?xml version="1.0" encoding="UTF-8" ?>
<GridDataServiceScript>
<Header> … Name, versioning, configuration specification </Header>
  <Body>
    <Documentation>Update with data delivered with the script</Documentation>
    <DefineParameter name="table">myimages</DefineParameter>
    <DefineParameter name="id">id = 2</DefineParameter>
    <Statement name="xyz" dataResource="MyDataResource">
        SELECT * FROM  <UseParameter reference="table" />
        WHERE <UseParameter reference="id" />
    </Statement>
    <Delivery name="delivery1">
        <Mechanism type="bulk" />
        <Mode type="full" />
        <From>xyz</From>
        <To>response</To>
    </Delivery>
    <Execute name="execute1">xyz</Execute>
  </Body>
</GridDataServiceScript>
```

Statement Architecture



OGSA-DAI Security

- Grid Security Infrastructure Authentication
- The role mapper takes credentials and a database name and returns a dbRole
- dbRole contains username and password



OGSA-DAI Product

OGSA-DAI Product

- Current release 3.0.2
- OGSA-DAI
  - 1183 downloads as at 27th Oct 2003
  - 461 downloads for R3 & R3.0.2
  - >350 UK downloads
- DQP prototype
  - 77 downloads since 1st Sept 2003
- Web site
  - 471 registered users
  - Transfers to users gigabytes / month
    - Aug 1.38GB Sep 1.63GB Oct 1.81GB

Binary, source and documentation available for download from:

www.ogsa-dai.org.uk

DAIT
Next Phase of Development

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