

UK e-Science

OGSA-DAI

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

Neil Chue Hong
Technical Manager OGSA-DAI
www.epcc.ed.ac.uk

November 2002

SuperComputing 2002

Baltimore



OGSA-DAI

Data Access and Integration for the New Grid

Uniform Service Interfaces

for Accessing Multiple Data Sources

within the Open Grid Services Architecture.

A UK e-Science Programme's Contribution to GT3



A Multi-Centre, Multi-Company Collaboration

**Funded by the UK e-Science Grid Core
Programme**

**to Develop Data Access and Integration
Services for OGSA**

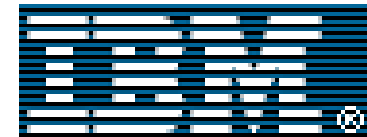
- Access to XML Databases -**
- Access to Relational Databases -**
- Distributed Query Processing -**
- XML Schema Support for e-Science -**



OGSA-DAI Partners

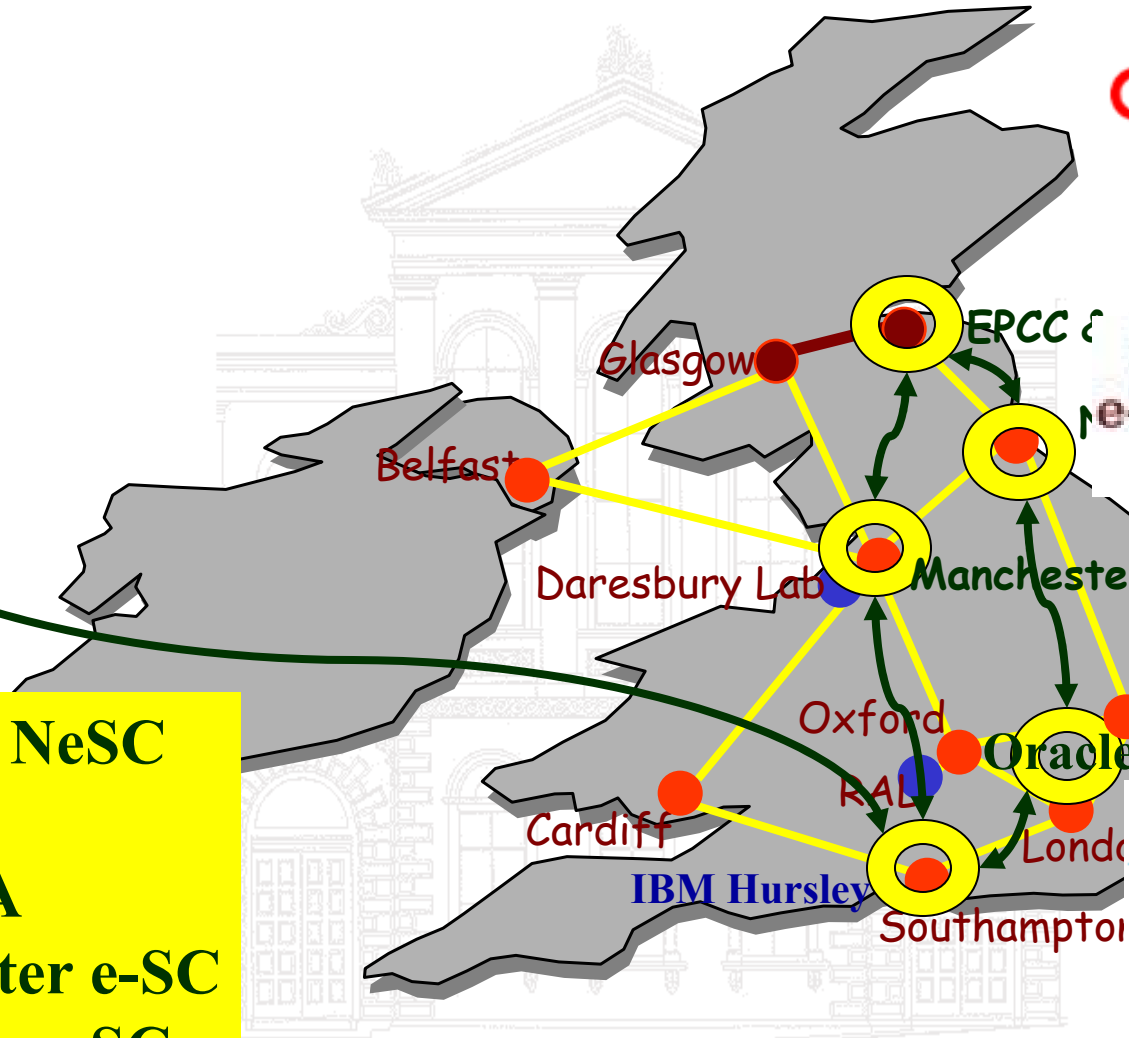


ORACLE®



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

\$5 million, 18 months, started February 2002



Building on OGSA & OGSi

- 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 out
- Alpha 15th January '03
- 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



Gets our vote

GGF DAIS WG

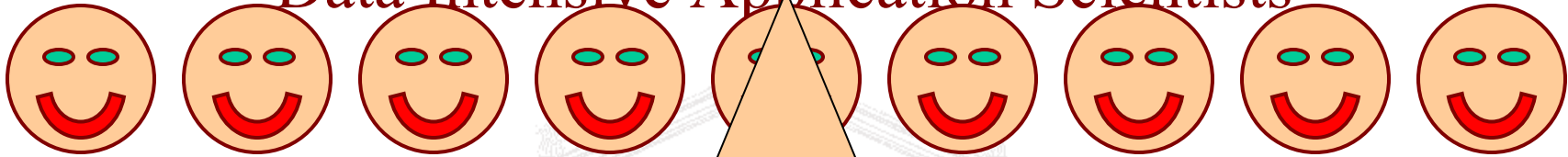
- **Chairs**
 - Norman Paton (Manchester Uni.)
 - Leanne Guy (CERN) ... **dropped out**
 - Dave Pearson (Oracle UK)
- **Activity**
 - BoF GGF4 Toronto
 - WG Meeting GGF5 Edinburgh
 - Papers for GGF6
 - Workshops & Mail lists
- **Goals**
 - Agree Standards for Database Access & Integration
 - Freely available reference implementations
 - ▶ OGSA-DAI one source & focus for discussions
- **Data Provenance and Derivation Workshop**
 - Significant UK contribution

Norman Paton,
Inderpal Narang,
Susan Maliaka,
Greg Ricardi, ...

<http://www.cs.man.ac.uk/grid-db/>

OGSA-DAI Stake Holders

Data Intensive Application Scientists



Data Intensive Applications App. Developers

Scientific Data Mining Integration Teams Tech. Developers

Monitoring

Scheduling

Accounting

**Keep all the
stake holder groups
satisfied**

Data Access

Grid Plug-ins

Structure

Operations
Team

Owners Data & Storage Resources

Data Providers

Data Curators

Distributed

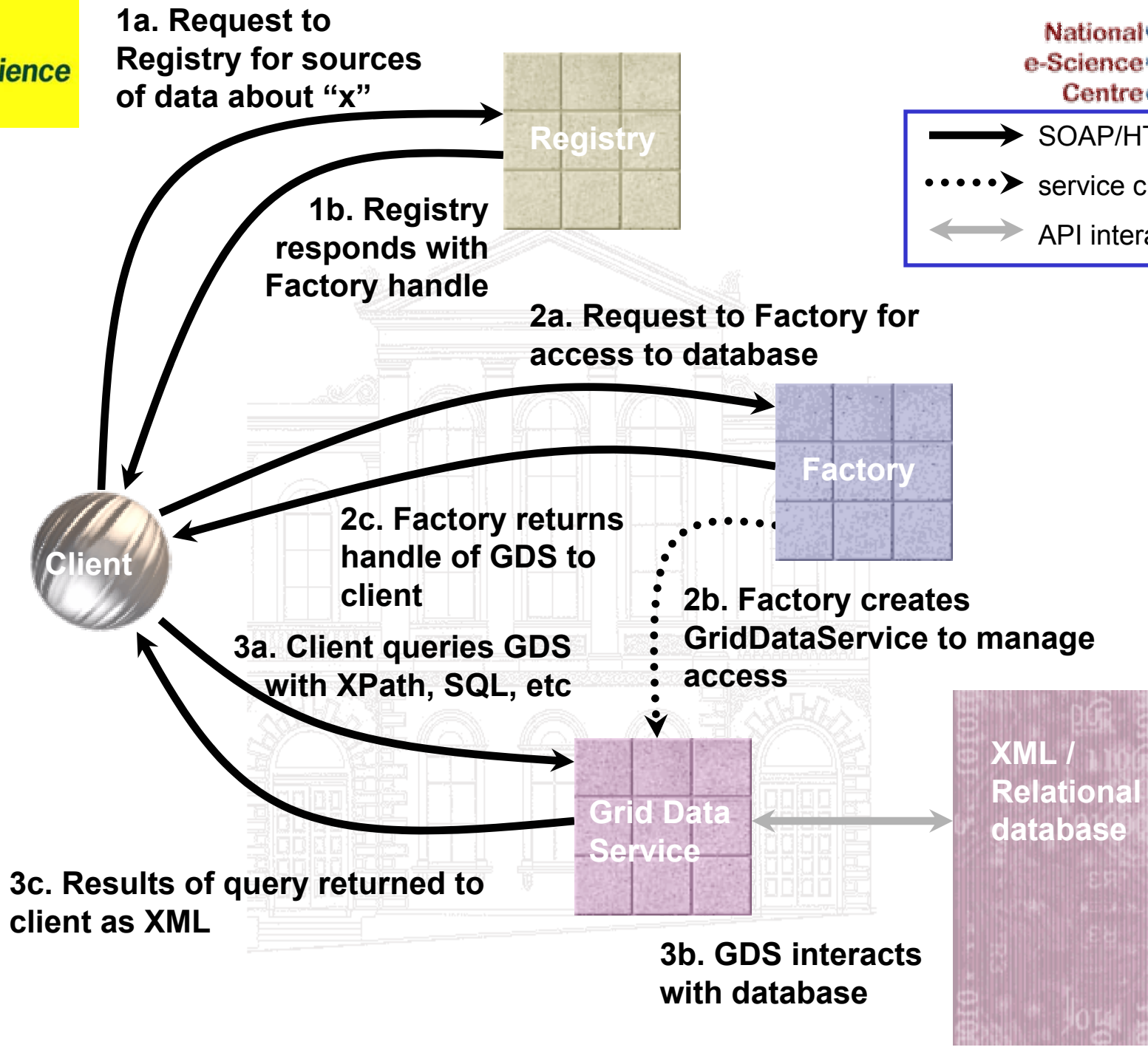
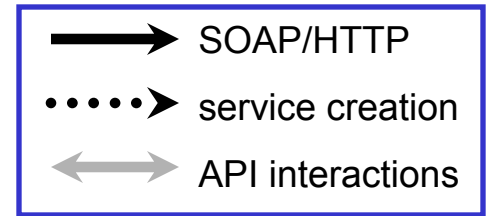


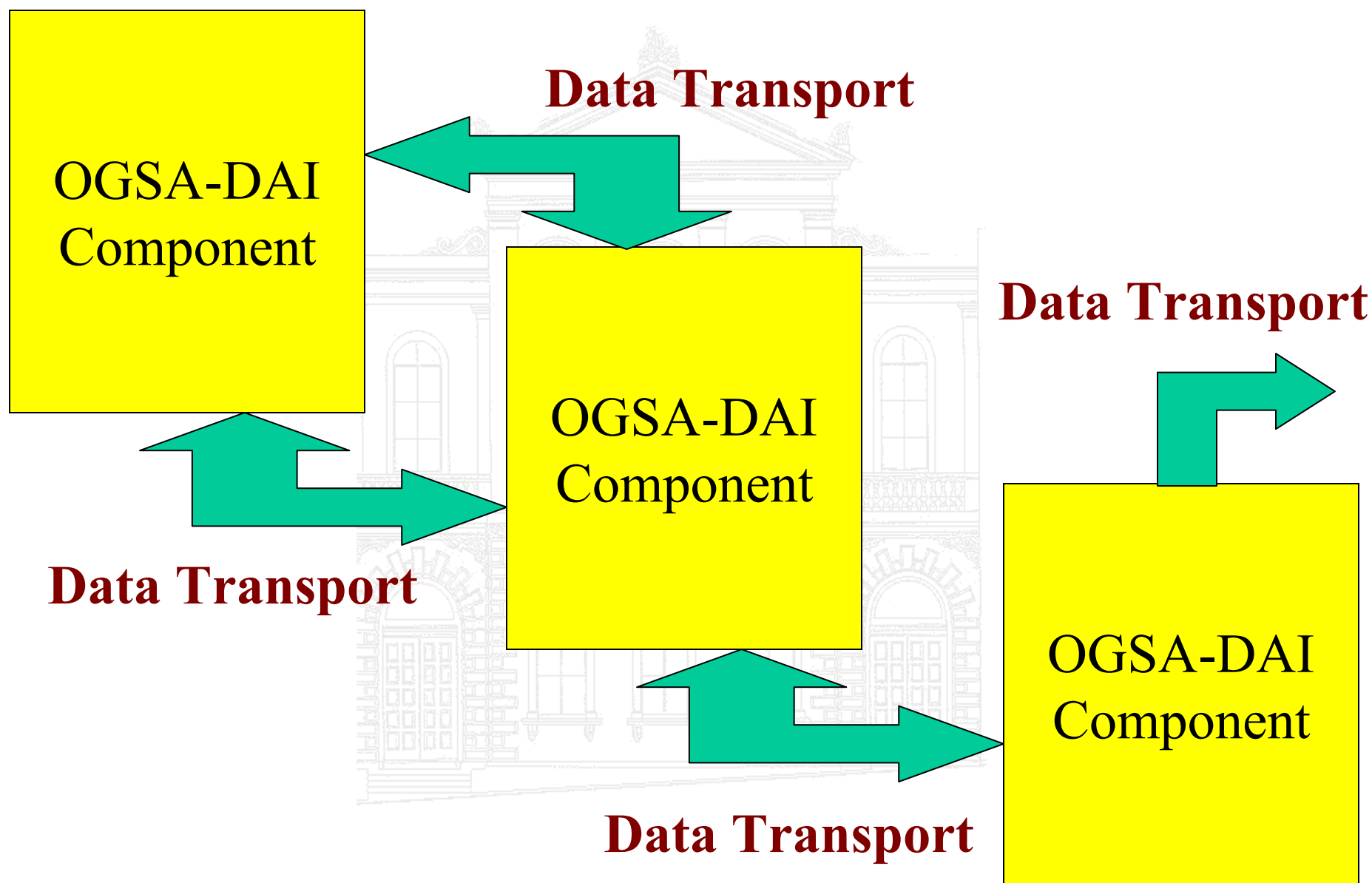
DAI Key Services



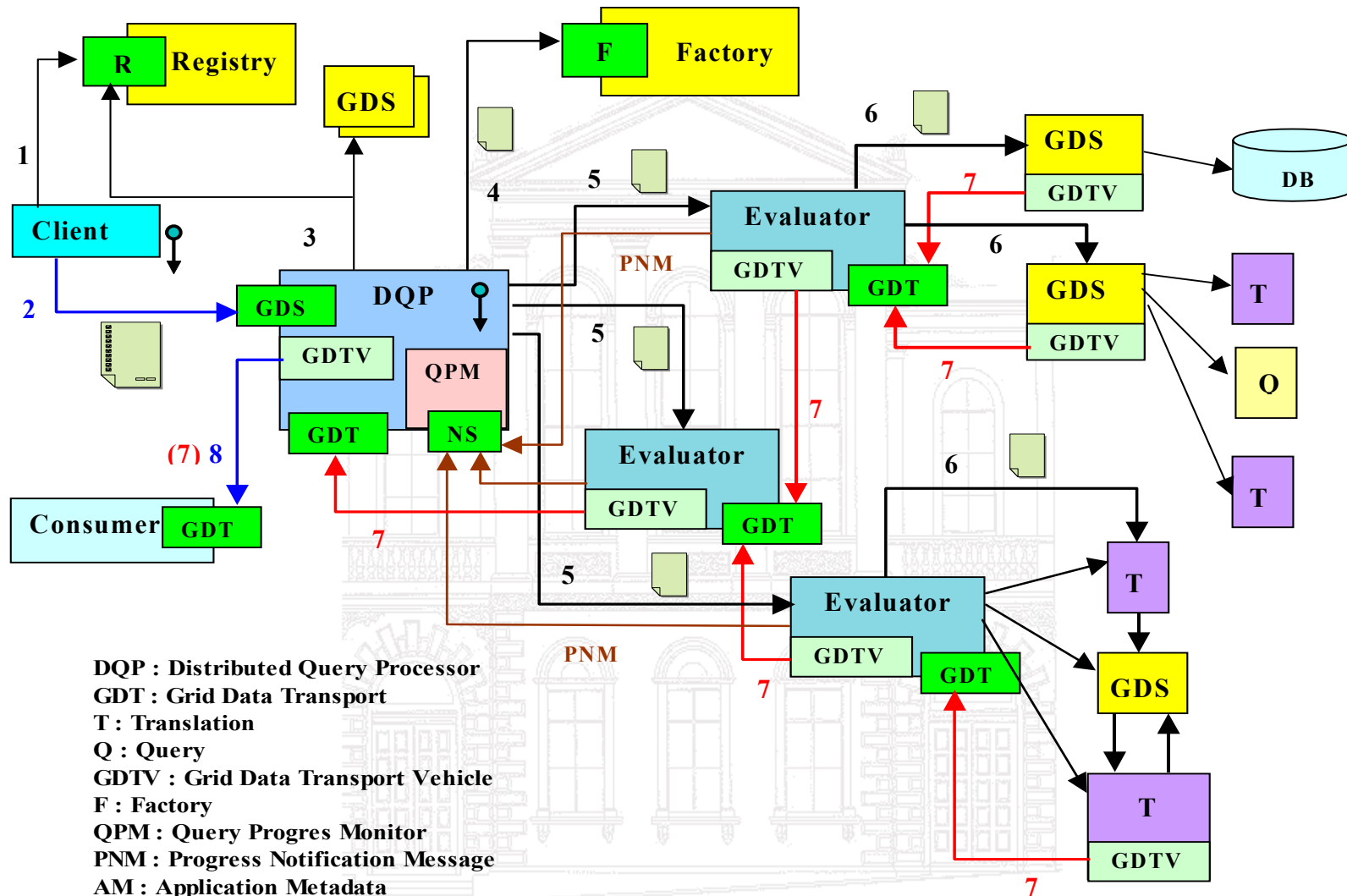
GridDataService	GDS	Access to data & DB operations
GridDataServiceFactory	GDSF	Makes GDS & GDSF
GridDataServiceRegistry	GDSR	Discovery of GDS(F) & Data
GridDataTranslationService	GDTS	Translates or Transforms Data
GridDataTransportDepot	GDTD	Data transport with persistence

Integrated Structured Data Transport
Relational & XML models supported
Role-based Authorisation
Binary structured files (later)

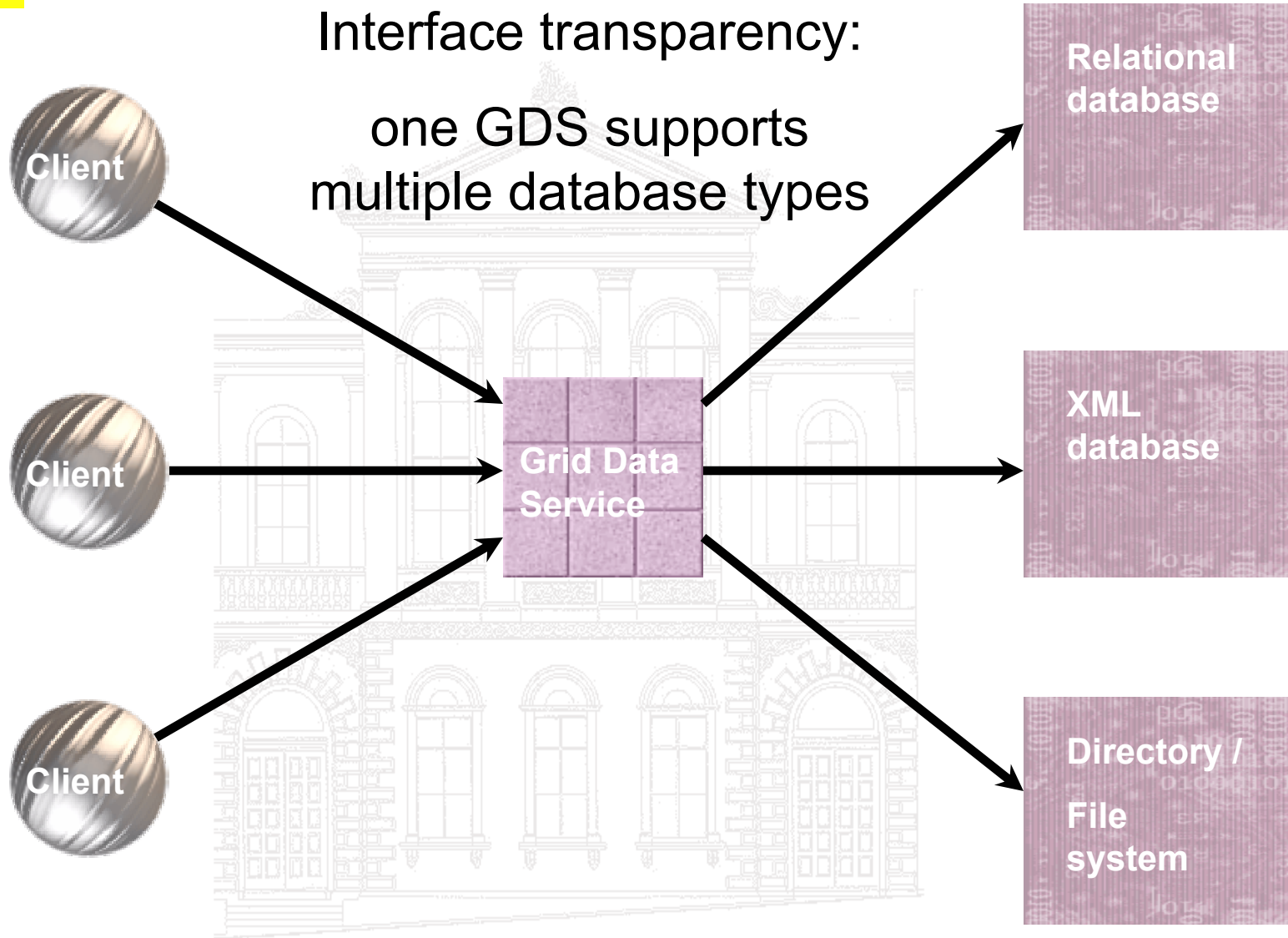




Distributed Query

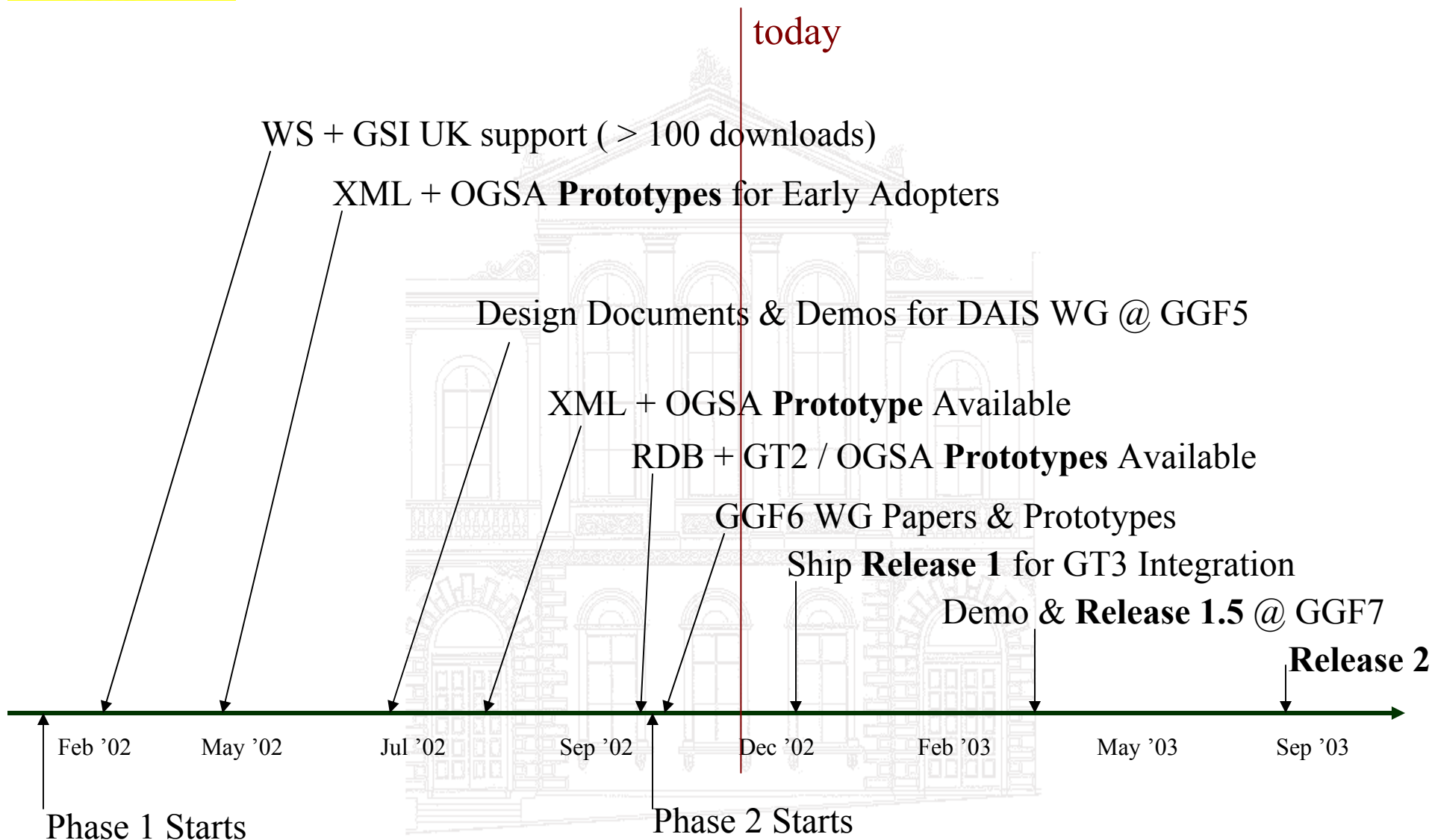


DQP : Distributed Query Processor
GDT : Grid Data Transport
T : Translation
Q : Query
GDTV : Grid Data Transport Vehicle
F : Factory
QPM : Query Progress Monitor
PNM : Progress Notification Message
AM : Application Metadata
CRM : Computational Resource Metadata
NS : Notification Sink





OGSA-DAI Time Line



A faint, light-colored architectural drawing of a classical building facade is centered in the background. It features a prominent pediment supported by columns, with a series of arched windows and doorways below. The drawing is rendered in a simple line-art style.

End of Presentation

Questions Please



Software Availability



Available now

- ❖ Phase 1 prototype of GDS, GDSF & GDSR for XML

Java implementations for the axis/tomcat platform and the Xindice database

- Globus-2 Relational database support

- ❖ BinX Schema v0.2

www.epcc.ed.ac.uk/gridserve/WP5

An XML Schema for describing the structure of binary datafiles – the power of XML for terabyte files

Software Q1 2003

- ❖ Reference implementation 1

- ❖ Access & Update

- XML databases
- Relational databases

- **To be released as Basic Services in Globus Toolkit 3**

umbriel.dcs.gla.ac.uk/NeSC/general/projects/OGSA_DAI/products



OGSA Relationship



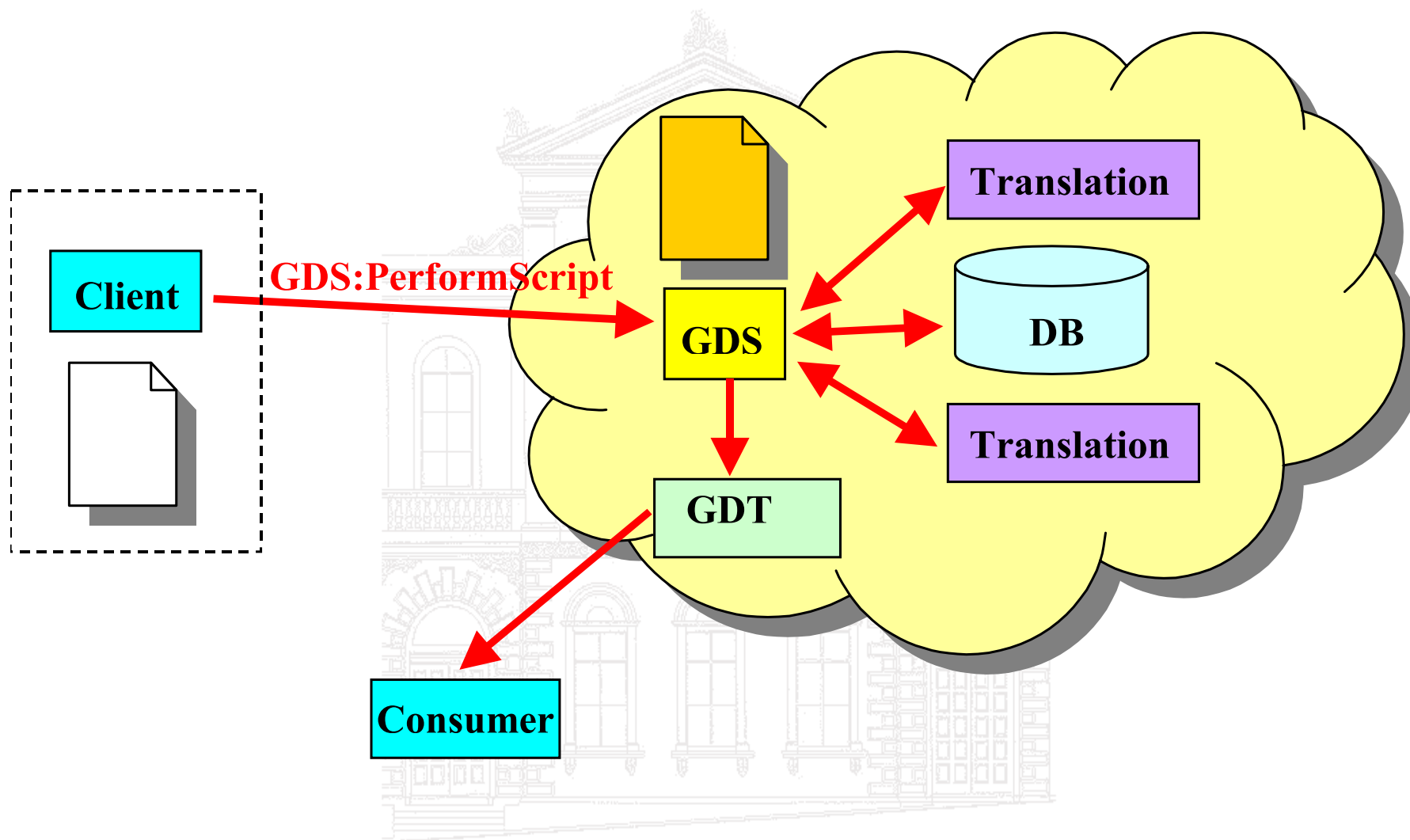
Class	GridService	Registry	NotificationConsumer	NotificationProducer
GDS	Mandatory		Optional	Normal
GDSF	Mandatory		Optional	Normal
GDSR	Mandatory	Mandatory		Normal
GDTS	Mandatory			
GDTD	Mandatory		Optional	Normal



DAI portType Usage



Class	GridDataService	DataTransport	Factory
GDS	Mandatory	Normal	
GDSF	Optional	Normal	Mandatory
GDSR	Optional		
GDTS	Optional	Mandatory	
GDTD	Optional	Mandatory	



Composed Components

