

UK e-Science Report on OGSA, OGSi & OGSA-DAI

Malcolm Atkinson

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

28th October 2002

**Meeting of the UK e-Science
Technical Advisory Group
Renaissance Hotel
Heathrow, London**

Overview

- **OGSA, OGSi @ GGF6**
 - ▶ Powerful force & progress
- **DAIS-WG @ GGF6**
 - ▶ International, strong UK lead, making good progress
 - ▶ Influencing OGSi
- **OGSA-DAI**
 - ▶ Phase 1: Complete and successful
 - But worrying structural issues
 - ▶ Phase 2: Energetic and effective development collaboration
 - But unresolved structural issues
- **Take home message**
 - ▶ Address these structural issues

OGSA & OGSi prognosis

- 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
- Alpha 15th January '03
- Major engineering effort
 - Complex & large code
- More demanding of software environment
 - Web services
 - Messaging
- Depends on planned extensions to standards
- First APIs for Java



Still gets my vote
Sustain your investment

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

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

- **Goals**

- Agree Standards for Database Access & Integration
- Freely available reference implementations
 - ▶ OGSA-DAI one source & focus for discussions

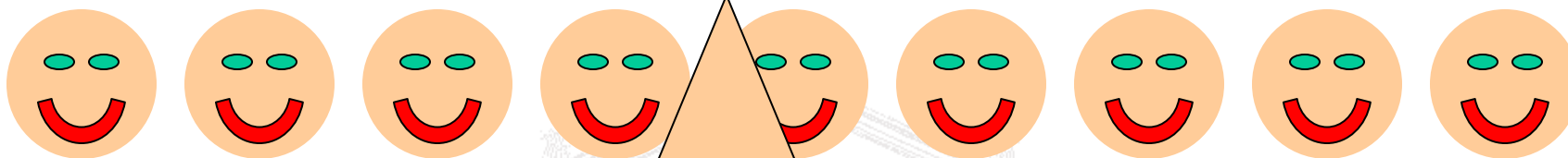
- **Data Provenance and Derivation Workshop**

- Significant UK contribution

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

The OGSA stake holders

Scientific Users



Scientific Application

Application
Developers

**Keep all the
stake holder groups
satisfied**

Scheduling

Authorization

Grid F

Infrastructure

Owners

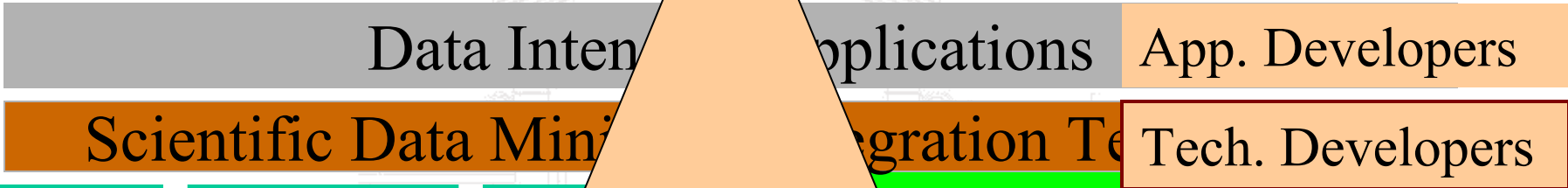
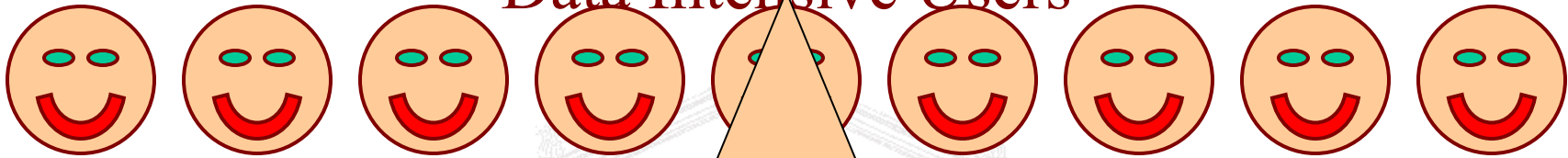
Data & Compute Resources

Operations
Team

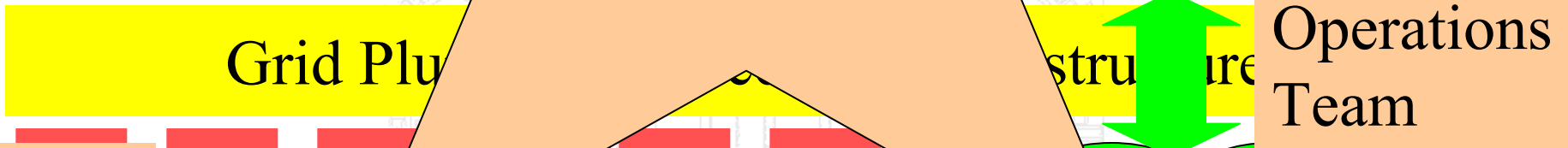
Distributed

OGSA-DAI Stake Holders

Data Intensive Users



Keep all the stake holder groups satisfied



Operations Team
Data Providers
Data Curators

← Distributed →



Data Access & Integration



- **Central to e-Science**
Astronomy, Earth Sciences, Ecology,
Biology, Medicine, ...
 - **Collaboration**
 - ▶ Shared Databases
 - ▶ Curated Knowledge
 - ▶ Accumulated Observations
 - ▶ Accumulated Simulations
 - **Computation**
 - ▶ Data mining
 - ▶ Input to models
 - ▶ Calibration of models
 - **Presentation**
 - ▶ Publication of results
 - ▶ Visualisation



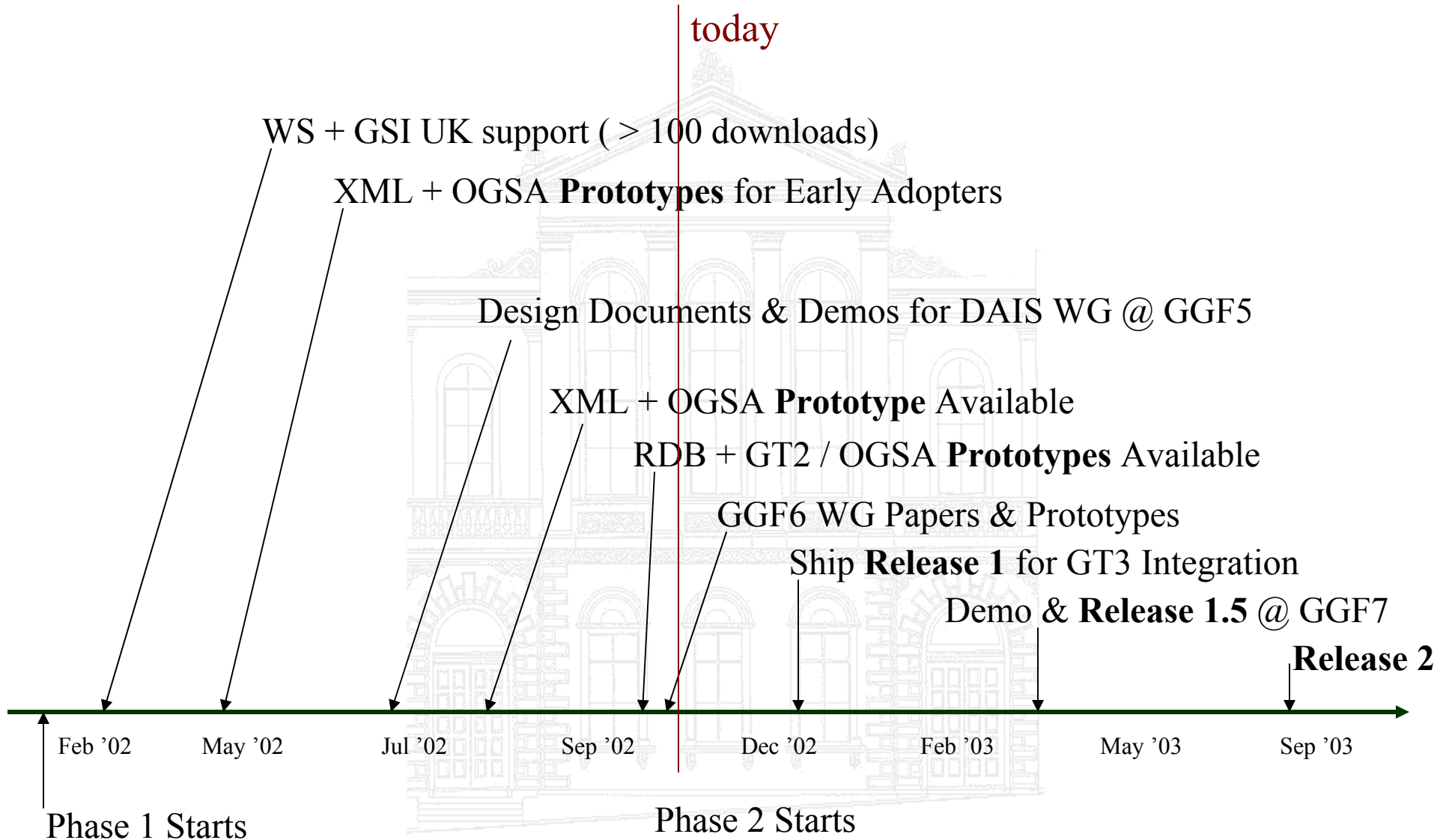
OGSA-DAI project



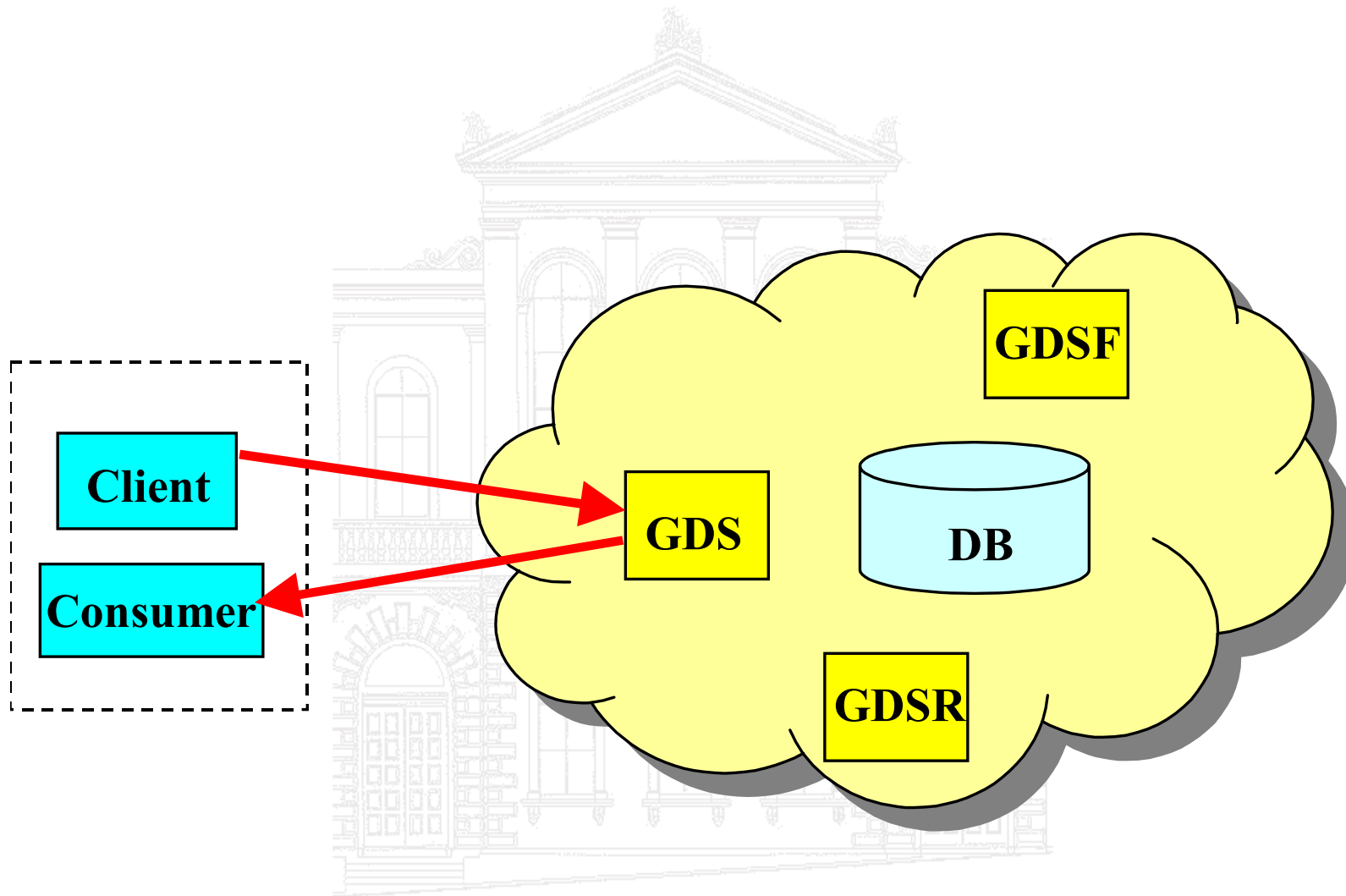
- **Lego kit for Data Access & Integration**
 - Components for e-Science Applications
 - Accelerated Application Development
 - Multiple Data Models
 - Distributed Data
 - Access via Grid & Proxies
 - Integration, Translation & Transformation
- **Open Source Reference Implementation**
 - For DAIS-WG standard
- **Trigger for DAI Component Construction**
 - Stimulate a community

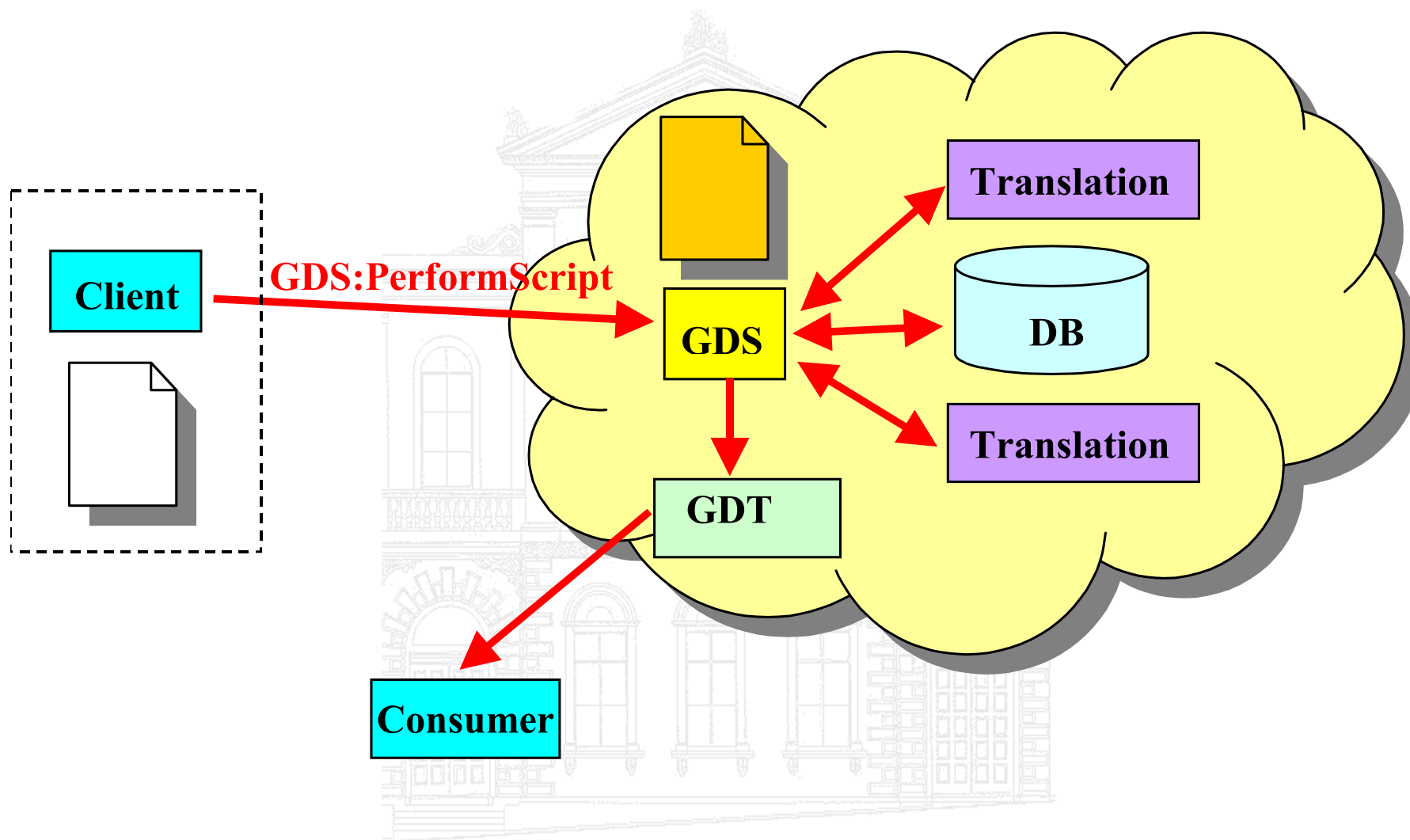


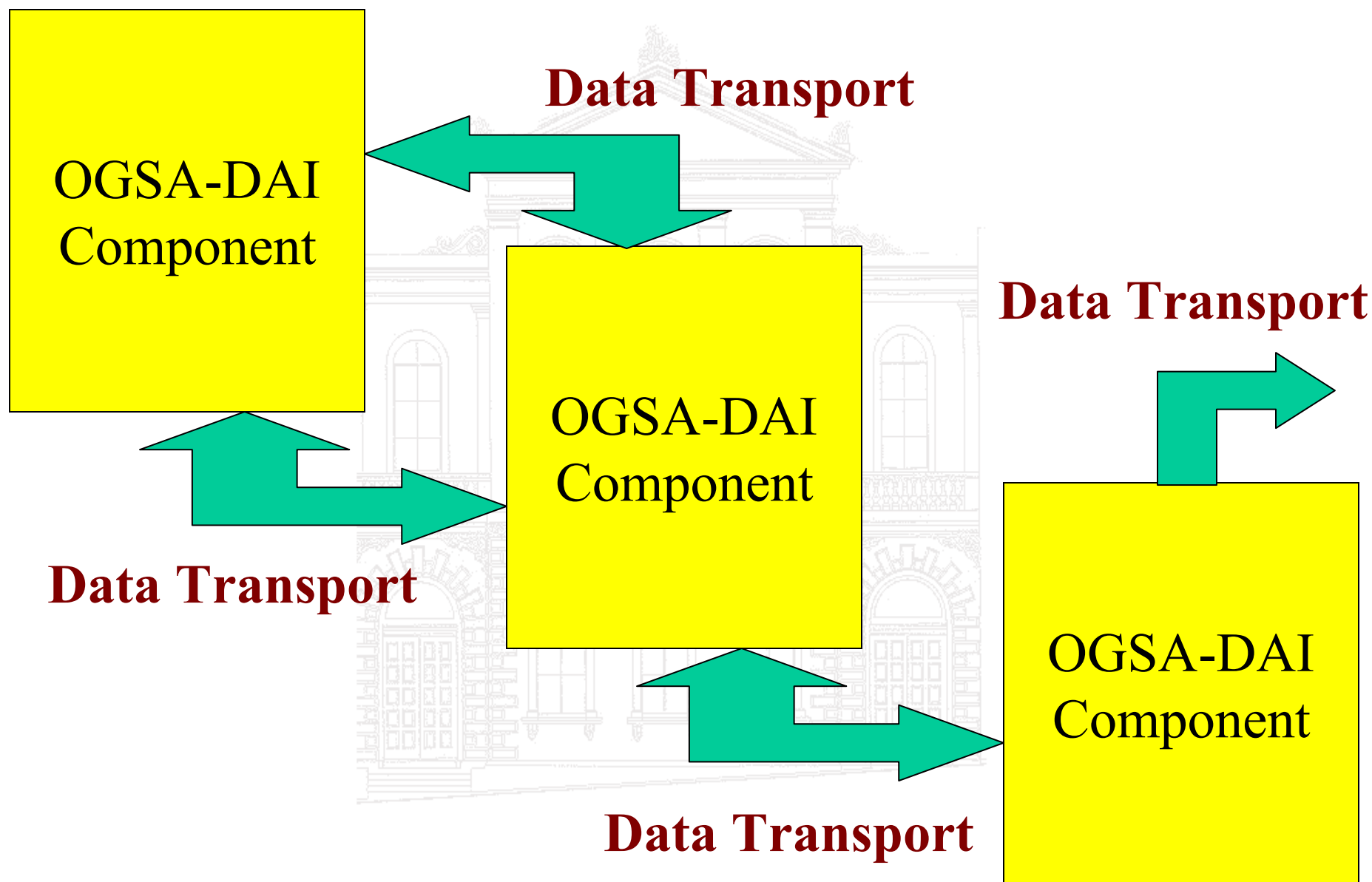
OGSA-DAI Time Line



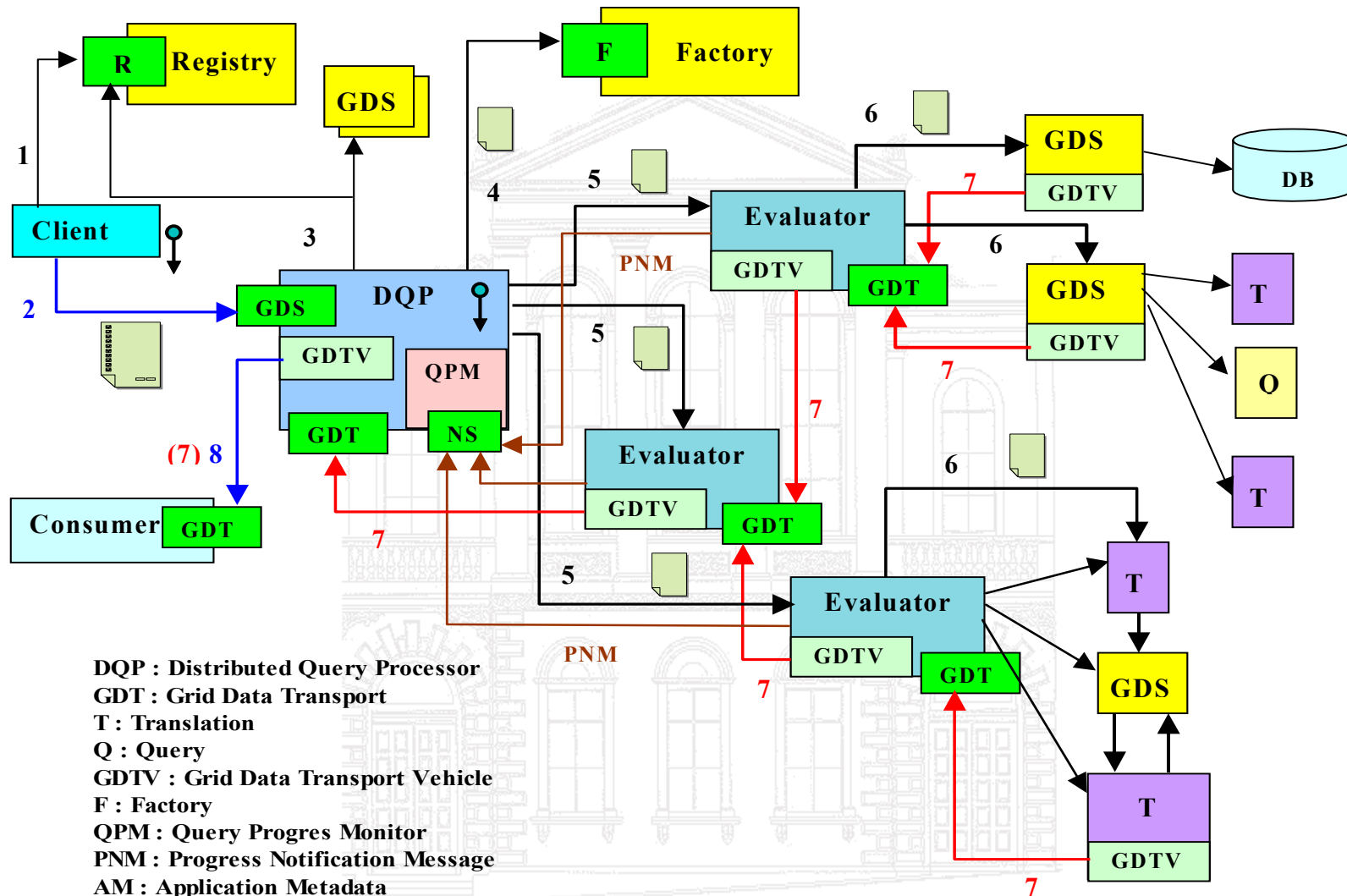
Primary Components







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

Today we have

- **A well-developed architecture**
 - **Issues identified and prioritised**
- **An emerging standard spec.**
- **A high level design**
- **Active cross-site development teams**
- **A functional spec for release 1**
- **A plan for the remainder of the project**
- **Effective operational arrangements for development**
- **Architecture needs rewriting**
 - **Doc. Structure**
 - **Emerging ideas**
- **Standard needs negotiation and to be led from reference implementation**
- **High-level design still being “polished”**
- **Resources not fully committed by managers**
- **Development resources still being set up**

Today we need

- **Commitment of and by Managers**
 - Members of teams allocated for an agreed period
 - Managers' and leaders' time and schedule committed
- **Completion of collaboration agreement**
 - Sign off source licence
 - IPR and liability agreement
- **Agreed strategy for Globus copyright**
 - General issue for many UK projects
- **Agreed plan for UK hand over and support**

Take Home Message

- **Developers working hard and well**
 - Quality of input and work very good
 - Good quality team
- **Plans, designs, proto-products emerging**
- **Committed to release to GTS Alpha**

Requires Project Board Action

- **Unresolved high-level issues are major risk**
 - **Could block release**
 - **Could disrupt effective teams**
 - **Could leave users without support**

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



OGSA-DAI Summary



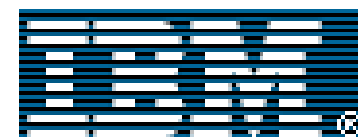
- **On Schedule & Going Well**
- **Contributions via DAIS-WG @ GGF5 & 6**
- **Releases with GT3 Releases scheduled**
- **Status: Early Days**
 - Released prototypes
 - Tested Architectural Design
 - Using OGSA
 - Working with Early Adopter Pilot Projects
 - ▶ AstroGrid & MyGrid
 - First PRODUCT release Dec '02
- **Influence OGSA-DAI direction**
 - Via DAIS-WG & Direct messages to us



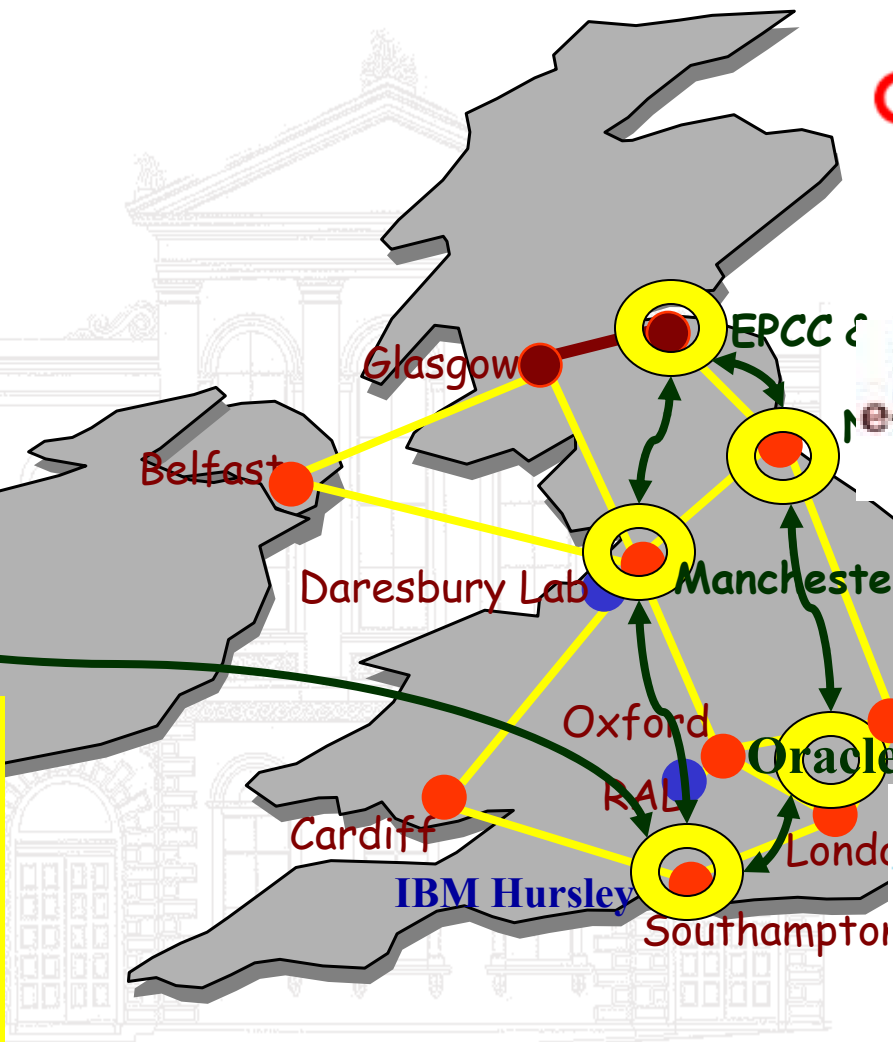
OGSA-DAI Partners



ORACLE®

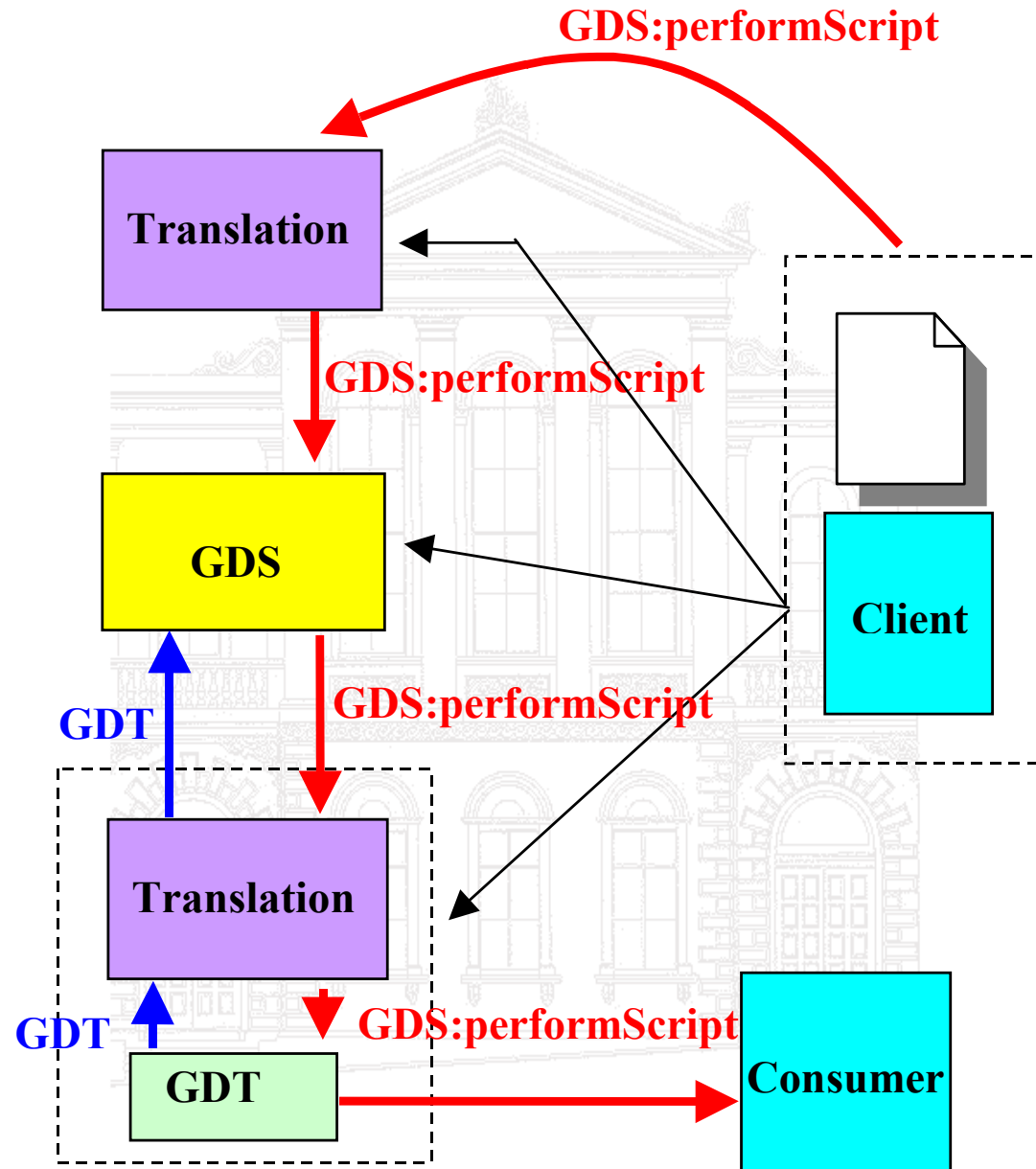


EPCC & NeSC
IBM UK
IBM USA
Manchester e-SC
Newcastle e-SC
Oracle



£3 million, 18 months, started February 2002

Composed Components





DAI Key Components



GridDataService GDS Access to data & DB operations

GridDataServiceFactory GDSF Makes GDS & GDSF

GridDataServiceRegistry GDSR Discovery of GDS(F) & Data

GridDataTranslationService Translates or Transforms Data

GridDataTransportDepot GDTD Data transport with persistence

Relational & XML models supported

Role-based Authorisation

Binary structured files



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	