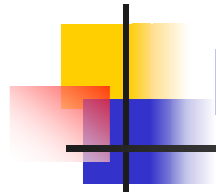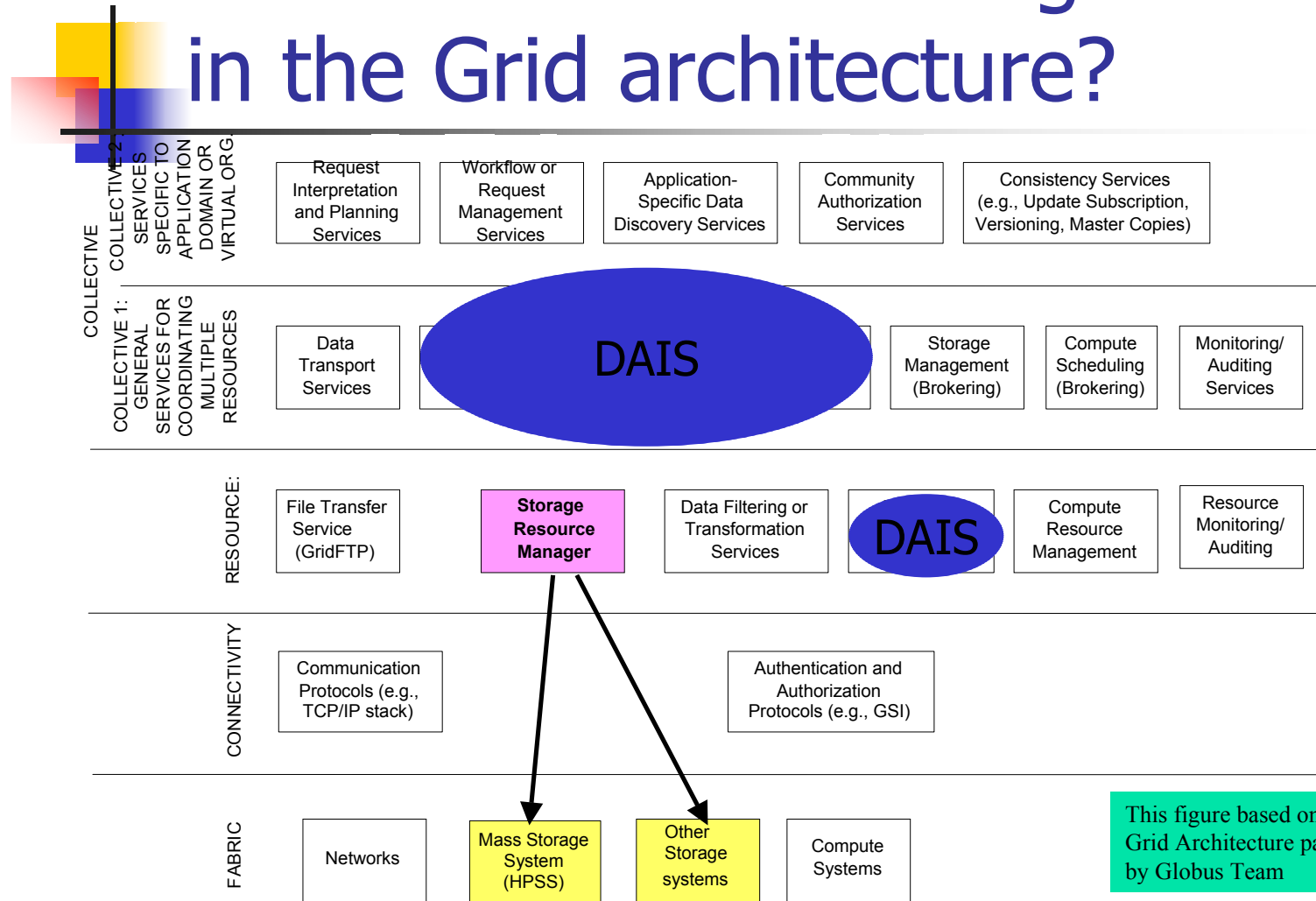# DAIS Overview
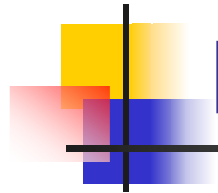
Dave Pearson

Oracle

DAIS Working Group Goals

- To identify and document the requirements of database and information management applications within a Grid setting.

- To develop a proposal for a standard interface for grid database services.

- To work with other groups within the GGF to ensure that the database service standard proposal conforms with other emerging Grid standards.

- To foster the development of reference implementations of Grid database services.
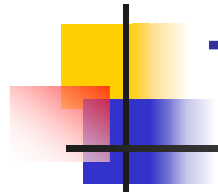
Where do SRMs belong in the Grid architecture?

	Request Interpretation and Planning Services	Workflow or Request Management Services	Application-Specific Data Discovery Services	Community Authorization Services	Consistency Services (e.g., Update Subscription, Versioning, Master Copies)

COLLECTIVE 2: SERVICES SPECIFIC TO APPLICATION DOMAIN OR VIRTUAL ORG.

COLLECTIVE

	Data Transport Services	**DAIS**		Storage Management (Brokering)	Compute Scheduling (Brokering)	Monitoring/ Auditing Services

COLLECTIVE 1: GENERAL SERVICES FOR COORDINATING MULTIPLE RESOURCES

	File Transfer Service (GridFTP)	**Storage Resource Manager**	Data Filtering or Transformation Services	**DAIS**	Compute Resource Management	Resource Monitoring/ Auditing

RESOURCE:

	Communication Protocols (e.g., TCP/IP stack)		Authentication and Authorization Protocols (e.g., GSI)

CONNECTIVITY

	Networks	Mass Storage System (HPSS)	Other Storage systems	Compute Systems

FABRIC

This figure based on the Grid Architecture paper by Globus Team
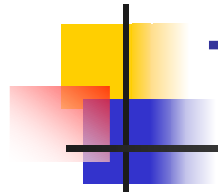
DAIS Specifications

- Web Services Data Access and Integration (WS-DAI):
    - A paradigm-neutral specification of descriptive and operational features of services for accessing data.
- The WS-DAI Realisations:
    - WS-DAIR: for relational databases.
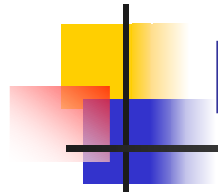   - WS-DAIX: for XML repositories.

Terminology - 1

- *Data Service*: a web service that conforms to one of the DAIS specifications.
- *Data Resource*: any system that can act as a source or a sink of data (e.g. a relational database).
- *Consumer*: an application that exploits a Data Service to access a Data Resource.
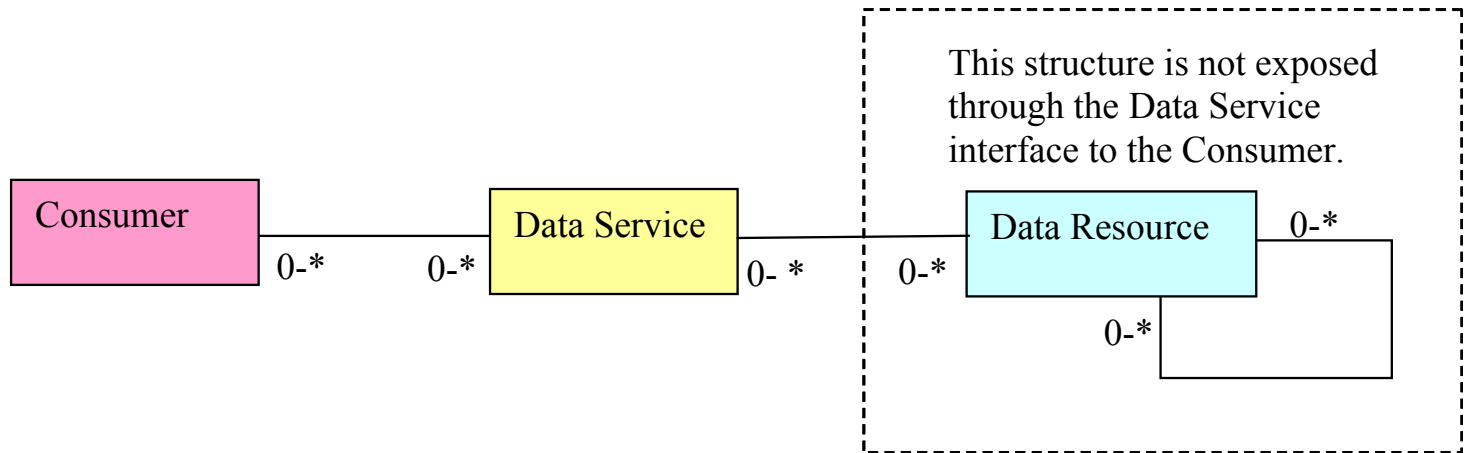- *Data Set*: an encoding of data suitable for externalising data outside a Data Resource.

Terminology - 2

- *Data Description*: provides metadata about a data resource and the relationship of that resource to a service interface.
- *Data Access*: provides access to data through a service interface.
- *Data Factory*: exposes derived data through a service interface.
- *Data Management*: manages the relationship between a service and the data resource that it exposes.
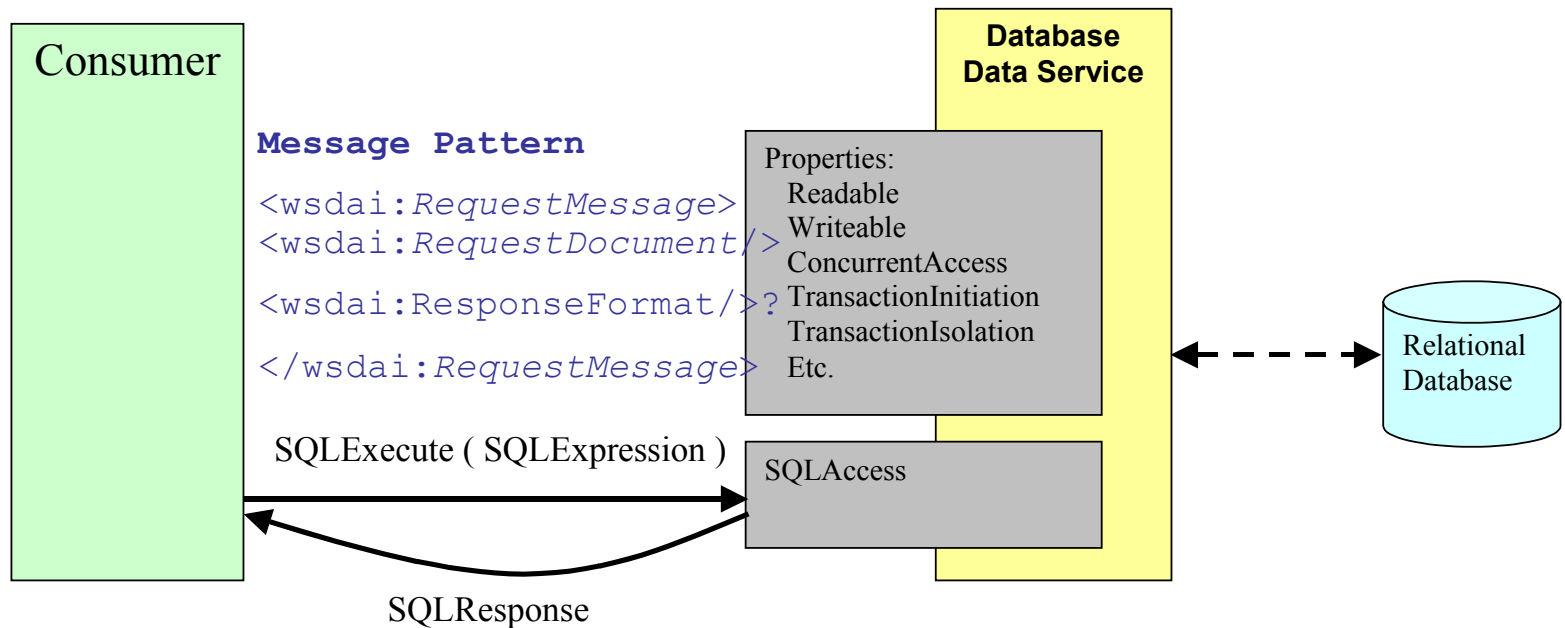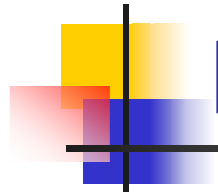
Data Service Model

Consumer

0-*   0-*   Data Service

0- *   0-*   Data Resource   0-*

This structure is not exposed through the Data Service interface to the Consumer.

0-*

Data Access Capabilities

- Data Access collects together messages that access and/or modify a resource.

Consumer		Database Data Service

Message Pattern

```
<wsdai:RequestMessage>
<wsdai:RequestDocument/>

<wsdai:ResponseFormat/>?

</wsdai:RequestMessage>
```

Properties:
 Readable
 Writeable
 ConcurrentAccess
 TransactionInitiation
 TransactionIsolation
 Etc.

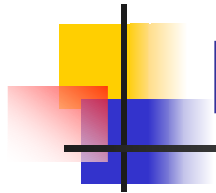SQLExecute (SQLExpression)

SQLAccess

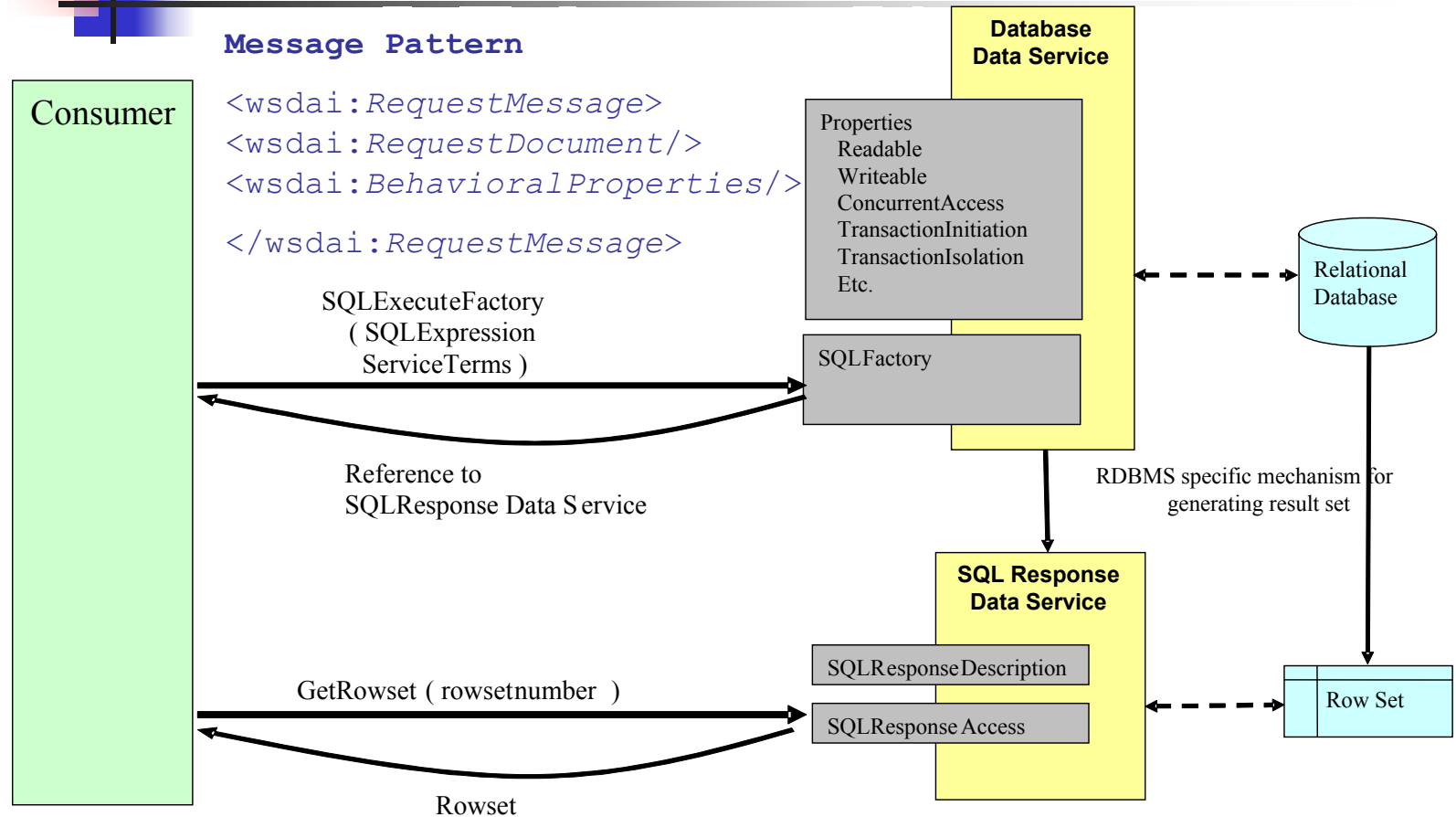SQLResponse
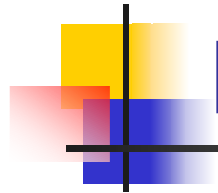
Relational Database

Data Factory Capabilities

- Data Factory collects together messages that implement the factory pattern.

- This factory pattern is used principally to support flexible handling of the results of requests.

Data Factory - 2

Message Pattern

Consumer

```
<wsdai:RequestMessage>
<wsdai:RequestDocument/>
<wsdai:BehavioralProperties/>

</wsdai:RequestMessage>
```

Database Data Service

Properties
  Readable
  Writeable
  ConcurrentAccess
  TransactionInitiation
  TransactionIsolation
  Etc.

SQLFactory

Relational Database

SQLExecuteFactory
  (SQLExpression
    ServiceTerms)

Reference to
SQLResponse Data Service

RDBMS specific mechanism for
generating result set

SQL Response Data Service

SQLResponseDescription

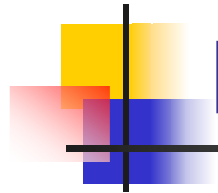SQLResponseAccess

GetRowset (rowsetnumber)

Row Set

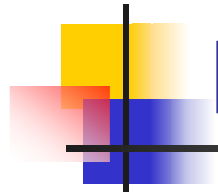Rowset

Data Management

- Data Management is not yet defined.
- Themes:
  - Managing the resource
  - Managing the service
  - Managing the service-resource relationship
- Awaiting the outcome of Web Services Distributed Management (WSDM) [WSDM] TC in OASIS
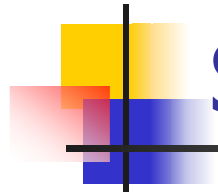
Federation Requirements

- Thorough metadata:
  - Logical, physical, …
- Flexible delivery:
  - Incremental.
  - Third party.
- Flexibility w.r.t. transparencies.
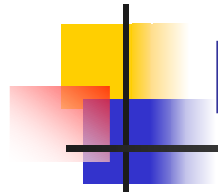 - Replication.
  - Naming.
  - …

Not in Scope/Dependencies

- Access to data in files ADF Bof
- Information dissemination INFOD WG
- Metadata modelling CGS-WG
- Data Transport
- Defining new data storage systems or access languages
- Naming schemes
- Security, Transactional management

Space Considerations

- Result set storage in 3rd party delivery
  - Volatile space may be required
- Space reservation & assignment
  - Durable and permanent space
  - Management operations on data resources
- DBMS space management
  - Database administration functions
  - Storage management solutions from vendors
    - Veritas, EMC, Net Apps, 10g ASM

DAIS - GSM Relationship

- Share models
- Metadata
- Naming
- Data Management Requirements
- Service Composability