Possible CGS and DAIS Collaboration

[CGS Session at GGF9]

[Wed 8 October 2003]

Susan Malaika and Amy Krause

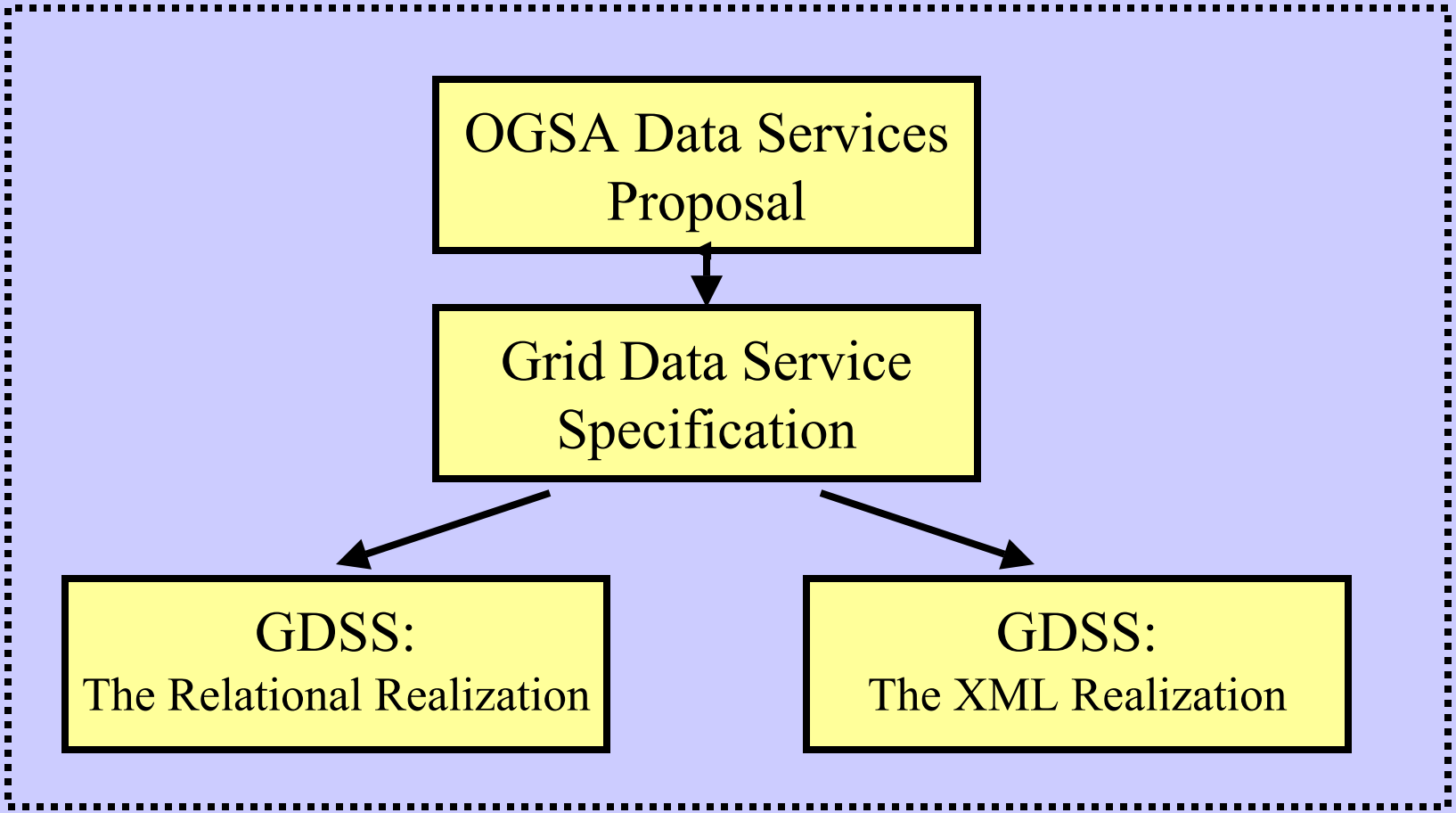
Agenda

- DAIS-WG Introduction (for CGS folk)
  - Data Access and Integration Services Working Group
- CGS-WG Background (for DAIS folk)
  - DMTF CIM-based Grid Schema Working Group
    - Distributed Management Task Force
    - Common Information Model
- CGS-DAIS Collaboration Motivation
  - Grid Data Service Service Data (GGF8 and GGF9)
  - Motivating Scenarios
- Possible Actions

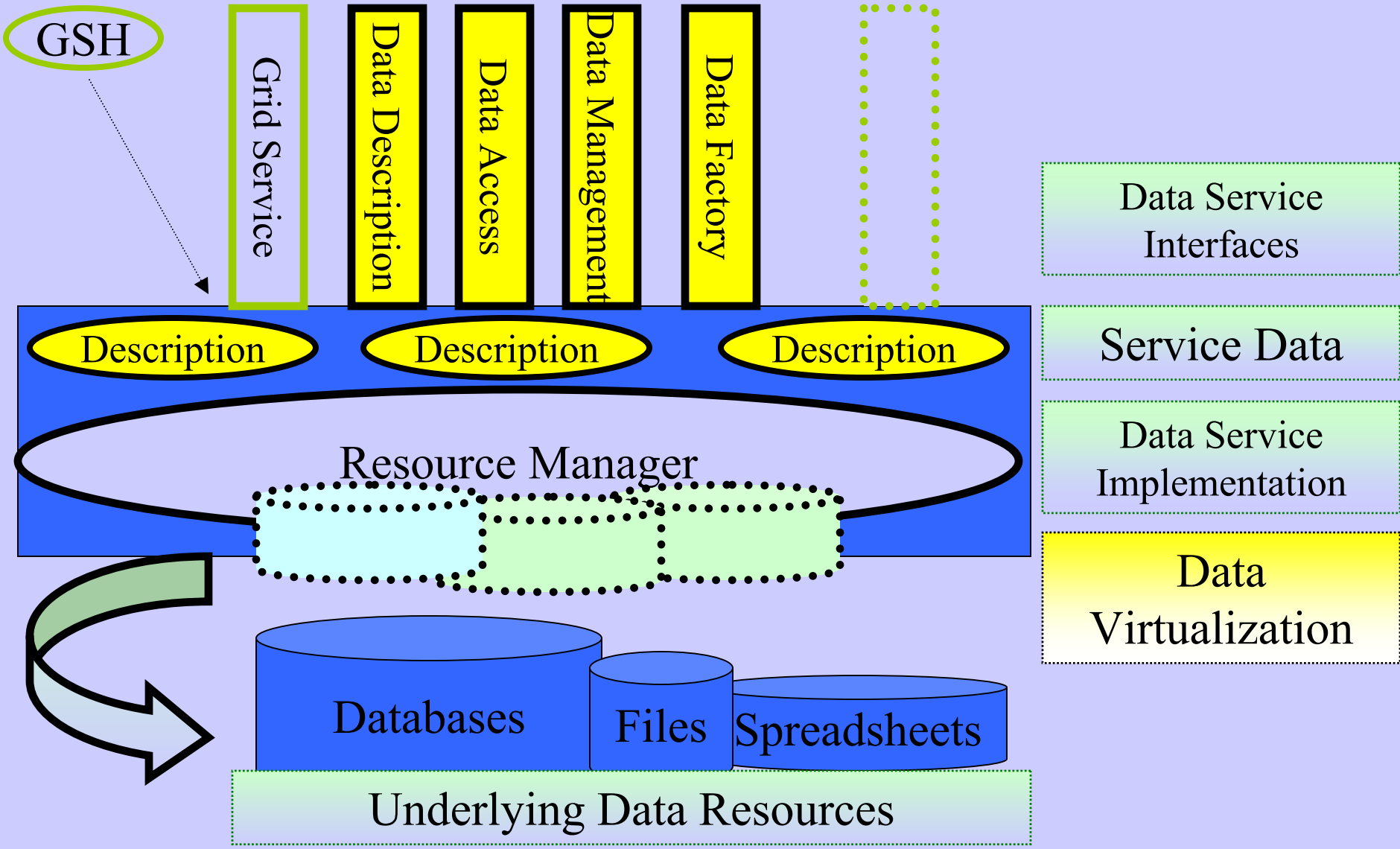
DAIS Introduction

- DAIS-WG is part of the Data Area at GGF
- The DAIS specifications describe how relational and XML databases can be accessed through Grid Data Services
  - Goal: to define general purpose interfaces that can be extended easily to access files etc
- The Grid Data Service Interfaces are categorized into
  - data description, e.g., for getting access to a relational table definition
  - data access, e.g., for issuing a SQL or XQuery request
  - data service creation, e.g., for creating a data service to issue SQL requests
  - data management, e.g., for creating a relational database or view
- The data description area and service data for the various portTypes are of most interest to the CGS group
  - Data management interfaces may turn out to be relevant too

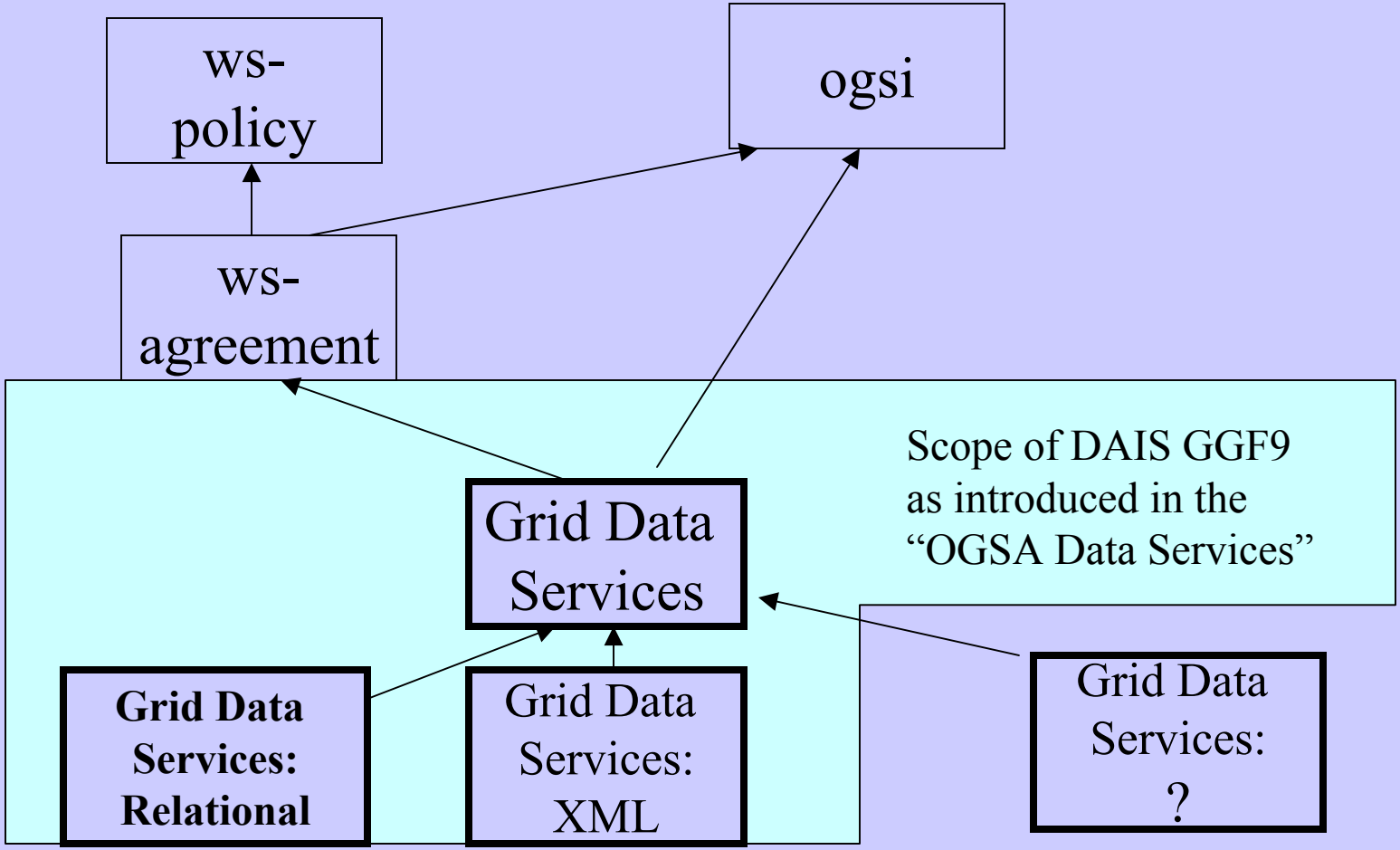
Current DAIS Document Structure



OGSA Data Services (provide a general context)



Grid Data Service Structure



DAIS-WG Specifications

Focus areas for the DAIS specification drafts

- GGF6:
  - Defining Key Data Delivery Styles, e.g., asynchronous, third party
- GGF7:
  - Aligning with OGSI 1.0
- GGF8:
  - Expressing the Underlying Service Model
- GGF9:
  - Aligning with OGSA Data Services, Modeling sessions and Agreement

CGS Background

- DMTF CIM-based Grid Schema Working Group
  - Distributed Management Task Force
  - CIM:
    - A model for describing overall management information in a network enterprise environment
- Produced JSIM
  - Job Submission Information Model

CGS DAIS Collaboration Motivation

- Some data descriptions in the DAIS-WG effort are common to all software components and not just to data components.
- The descriptions are sometimes common across domains such as systems management and application access.
- The overlapping constructs seemed similar to items described in DMTF.
  - The CGS-WG is collaborating with CIM and DMTF
- Note: in GGF9 we have reduced the general purpose data descriptions in the DAIS specifications. Nevertheless these descriptions are needed, e.g., for provisioning a system.

Service Data

GGF8 Grid Data Service Service Data

- In GGF8 Grid Data Service included these items
  - ***ExternalResourceManagerType***: product name, version and vendor name)
  - ***PhysicalPropertiesType***: physical constructs; includes space used, buffers
  - ***SoftwareCapabilitiesType:*** capabilities of the software; includes software
  - ***FeaturesInstalledType***: description of installed features; includes product enhancements, features, extra functionalities
  - ***FeaturesActivatedType***: lists the items from featuresInstalled that are activated in the installation.
  - ***SecurityCapabilitiesType***: lists the security capabilities are provided by the resource manager..
  - ***TransactionCapabilitiesType***: lists the transactional capabilities provided by the resource manager.

GGF9 Grid Data Service DataDescription portType

- ***name:***
  - a name associated with the data represented by the Data Service.
- ***structure:***
  - a description of the structure of the Data represented by a Data Service. The mechanism by which the structure is described must be extensible as it is specific to the data model.
- ***size:***
  - the size, in bytes, of the Data represented by a Data Service.

GGF9 Grid Data Service DataAccess portType Service Data

- ***status:*** status of the Data Service with respect to data access. An enumeration with the values:
  - Ready – The Data Service is ready to be accessed.
  - Initializing – The Data Service is not ready to be accessed.
  - Error – An error has occurred leaving the Data Service in an error state.
- ***dataFormat:*** valid formats of data passed out of the get() operation and into the put() operation.

GGF9 Grid Data Service DataFactory portType Service Data

- ***proposedAgreement:*** WS-Agreement documents during the process of requesting the creation of a new Data Service.

GGF9 Relational DataDescription portTypes

- **RelationalDescription**:
  - *relationalSchema*
  - *indexes*
  - *storedProcedures*
  - *userDefinedTypes*
  - *userDefinedFunctions*
  - *trigger*
- **RowsetDescription:**
  - *rowSchema*
  - *noOfRows*

Motivating Scenarios

Scenario 1: Discovering a Database

- Party 1:
  - Publish database information in a registry (use information from CIM model)
- Party 2:
  - Search the registry
- Party 2:
  - Locate the database service with suitable characteristics
- Party 2:
  - Access the database service

Scenario 2: Provisioning a Complete System

- Party 3:
  - Install the operating system
- Party 3:
  - Install the database management system (use information from CIM model to configure)
- Party 3:
  - Define the databases and tables
- Party 3:
  - Load the tables
- Party 3:
  - Install the applications

Scenario 3: Provisioning a Database in an Existing Database System

- Party 4:
  - Publish database management system information in a registry (use information from CIM model to configure)
- Party 5:
  - Search the registry
- Party 5:
  - Locate suitable database system with sufficient storage
- Party 6:
  - Define the tables
- Party 6:
  - Load the tables

Scenario 4: Creating a replica in an Existing Database System

- Party 4:
  - Publish database management system information in a registry (use information from CIM model to configure)
- Party 5:
  - Search the registry
- Party 5:
  - Locate suitable database system with sufficient storage for a replica
- Party 6:
  - Define the replica
- Party 6:
  - Initialize the replica
- Party 7:
  - Keep the replica in sync

Scenarios 4 & 5: Starting up and Shutting down a complete system

- Start the computer
- Start the operating system
- Start the database system
- Run the applications
- Run the database backup utility
- Stop the applications
- Stop the database system
- Stop the computer

Possible Actions

- Examine the DMTF models to determine if they can support grid scenarios. DMTF models of particular interest include:
    - database, storage, application (install, deployment and related capabilities), and security
    - CIM_Application28 and CIM_Database28
        - http://www.dmtf.org/standards/cim/cim_schema_v28
- Alternative: Start work on modeling areas that have not been attempted
    - data provenance: Or is this the domain of the metadata management group BOF?

Possible Actions

- Define a CIM schema for data
  - that meets grid requirements that encompass database, files and storage
  - That includes the characteristics of installing and maintaining the installed product, its deployment model, and related capabilities such as security.
- Identify gaps in CIM and model the missing entities in UML and CIM MOF.
  - The resulting UML and CIM MOF will be returned to DMTF, targeted for preliminary release of CIM 2.9 or CIM 2.10
  - Given the work agreement that is in place between GGF and DMTF, it is expected that there will be DMTF participation in the CGS working group.

References

- CGS-WG
  - http://www.isi.edu/~flon/cgs-wg/
- DMTF CIM
  - http://www.dmtf.org/standards/cim
- DAIS-WG
  - http://www.gridforum.org/6_DATA/dais.htm