Data Services: Issues & The Way Forward

June 8, 2004

Allen Luniewski
Grid/On-Demand Strategy
IBM Data Management, SWG
luniew@us.ibm.com

Agenda

- **Review open issues**
- **Raise new issues**
- **Discuss issues**
- **Plan the path to GGF 12**

Unfinished Business from Current Draft

- **Common ResourceProperties elements to be defined**
- **Common DataAccess operations to be defined**
- **Common DataFactory operations to be defined**
- **Common DataManagement operations to be defined**

Open Issues in GridForge Tracker

1. Move design principles from base to data services

2. Using EPRs to denote replicas

3. DS Placement Mgr vs. Other Placement Services

4. Naming of Data Services?

5. Third Party Data Delivery vs. Workflow Services

6. Clarification of Management Interface

7. Data Description Interface: Needed in WSRF?

8. Factory Interface - Should not be an interface

9. Relationship of Data Services to Other Grid Services

10. Clarify: Not All Data Related Services are Data Services

547: Move design principles from base to data services

- **The base Grid Data Service Specification describes a set of design principles which are our interpretation of how to apply Data Services to the OGSI environment. Some of the information may be more appropriate in the OGSA Data Services document.**

846: Using EPRs to denote replicas

- **The Data Services specification implies that EPRs (endpoint references) can be used to denote particular replicas of data for, e.g., performance reasons. The Data Services specification should discuss means for doing this (e.g., how to exploit the Reference Properties part of the EPR).**

- **The Data Services specification leads to the creation of a placement service within a data service to manage the creation, destruction and placement of caches and replicas of data within the Data Service implementation. There are also higher level data placement services (e.g., at the job services level). The Data Service specification should properly position its placement service w.r.t those other services.**

848: Naming of Data Services?

- **Should the data services specification specify how it relates to a global naming service?**

849: Third Party Data Delivery vs. Workflow Services

- **The data services specification permits the notion of third party delivery of data (& this is developed further by other DAIS specifications). How does this ability relate to higher level workflow services?**

850: Clarification of Management Interface

- **The purpose and bounds of the Management interface needs to be clarified. It clearly allows management of the virtualization. But, does it support, or permit, management of the data resources "backing" a particular data service.**

851: Data Description Interface: Needed in WSRF?

- **The OGSI 1.0 version of Data Services had a Data Description interface to capture SDEs and operations to access the SDEs. In the WSRF framework, SDEs have disappeared, having been subsumed by the Resource Properties document of the WSRF resource. WSRF also provides operations against the Resource Properties document in its "base" interface. The current version of the Data Services specification has removed the Data Description interface under the belief that it has no content. Is there a need for that interface?**

852: Factory Interface - Should not be an interface

- **There is no notion of a generic Factory interface in WSRF. The Data Services specification should be cleaned up to remove the notion of a Factory interface, replacing it with something more akin to a category of interface items.**

853: Relationship of Data Services to Other Grid Services

- **The relationship of a Data Service to other grid services should be clarified. This clarification needs to be done in the context of the "environment' that WSRF specifies. In addition, this clarification needs to be done compatibly, and not redundantly, with the work in the OGSA Architecture document to specify the data architecture.**

854: Clarify: Not All Data Related Services are Data Services

- **The Data Services document describes a set of services providing a certain class of services. Although this class is (intentionally) broad, there is no reason to believe that it is either all inclusive nor that it provides the optimal way to describe all data related services. The specification should make this clear in its preamble.**

New Issues

The Path to GGF 12

- **Goal: for GGF-12 to have a document that:**
  - ▶ **All "future" work completed**
  - ▶ **All known open issues resolved**
  - ▶ **Compatible with work of OGSA Data Architecture design team**
- **Mechanisms:**
  - ▶ **GridForge Tracker**
  - ▶ **E-mail**
  - ▶ **Teleconferences**
  - ▶ **Leverage OGSA Data Arch. F2F**
  - ▶ **???**