

OGSA Data Movement Interface Standardization

Working Group Charter

Global Grid Forum, Data Area

1. Administrative Information

Name and Acronym:

OGSA Data Movement Interface (OGSA-DMIS)

Chairs:

William E. Allcock, ANL, USA, allcock@mcs.anl.gov

Michel Drescher, FLE, UK, Michel.Drescher@uk.fujitsu.com

Email list:

dmis-bof@ggf.org (currently)

ogsa-dmis-wg@ggf.org (presumably if we get chartered)

Web page:

No web page yet

2. Charter

2.1 Purpose

The OGSA Data working group has identified a need to define an interface that standardizes the process of invoking the movement of large amounts of data. No standardization body addresses this issue so far. OGSA ByteIO, while situated in the same area of interest, addresses a different problem. Although OGSA ByteIO could be used in the end to transfer a complete set of data from one location to another, it is explicitly designed to access and control subsets of a data set that is stored at a remote location. The OGSA DMI Working Group tackles the problems of discovering of data transport protocols available at the data's source and destination location and agreeing on one of them, and the actual invocation of the agreed data movement. This includes direct data movements and 3rd party data movements.

The OGSA DMI Working Group, though clearly associated to its roots, the OGSA Working Group, aims to define a set of interfaces that is independent of the "umbrella use case" of Grid Applications so that it can be used in scenarios and deployments other than the Grid. However, one objective of this Working Group is to produce a rendering that is compatible to the OGSA WSRF Basic Profile version 1.0 [1].

Note: the term "movement" used in this charter does not imply a semantics whereby the data is deleted from the source of the transfer. The data movement operations defined by this working group can be used to implement "move", "copy" or "replicate" semantics, or others, as appropriate.

2.2 Scope

This Working Group will focus on the processes and information exchanges involved in setting up a data movement, and executing this data movement later on.

In the first phase, the set-up phase, the participating nodes reach an agreement on important aspects of the data movement. If an agreement cannot be reached, terminating fault messages are communicated to external entities for further processing, and the ongoing negotiations are ended. In the second phase, the execution phase, the participating nodes (not necessarily the same nodes as in the set-up phase) execute the data movement using the parameters agreed upon in the set-up phase.

Prior to initiating the data movement, extra information can be provided to the DMIS in order to make prioritization decisions within the service. This includes the selection of a transport protocol, for example GridFTP, and parameters for reliability, timing, scheduling, resource usage, accounting, billing, etc. The Working Group will explore existing mechanisms to reach such agreement, e.g. WS-Agreement and use them where appropriate. It is out of scope for this Working Group to define or standardize new data transfer protocols. It will explore the drawbacks and benefits of defining a default transport protocol that all implementations must support (for example, OGSA-Bytelo specifies a default transfer mechanism, named "<http://schemas.ggf.org/bytelo/2005/10/transfer-mechanisms/simple>", that all implementations must support) allowing that there is no benefit of having a default protocol and specifying an appropriate fault instead. However, this group will reuse names where appropriate (e.g. from OGSA-Bytelo), and define new names that identify such protocols in order to reach an agreement upon the use of a protocol for actually transferring the data.

Executing a data movement includes the invocation of the transport protocol itself, and applying the previously agreed parameters where appropriate. While the data movement is executing, control and management operations on the data movement are necessary, such as "cancel", "suspend", and "resume". Progress information, including general status information, must be provided to interested parties as well.

In order to provide a specification that is useful to a most broad audience, this Working Group will define abstract documents, and renderings into two different interoperability frameworks. While both renderings will be WS-I basic profile 1.0 [3] compliant, one rendering will also be compliant to the OGSA WSRF Basic Profile 1.1 [1], and will take advantage of resource properties, notifications, etc, for exposing state, while the other rendering will be a "pure web service", i.e., state will be exposed via service methods and the client will need to poll.

To ensure maximum interoperability between implementations that are plain WS-I compliant, and implementations that are OGSA BP compliant, the Working Group will ensure that the primary goals, to agree on a transport protocol, and to actually transfer a finite amount of data, can still be achieved between two different implementations.

To support 3rd party data transfer the data source and destination may use different implementations and local naming schemes. To accomplish 3rd party data transfer, a uniform, yet abstract naming scheme for resources (data in general, files in particular) is required. This working group will provide such abstract uniform naming scheme.

This Working Group will not deal with the following issues, as they are either transport protocol specific, data specific or else out of scope:

- Data structure
- Concurrency
- Encryption and security
- User Management
- Storage Management

2.3 Goals

Documents to be produced:

Title: Data Movement Functional Specification

Abstract: This document defines operations, inputs, outputs, and the underlying semantics for a service that can provide data movement services. These definitions are in the form of an abstract description, not actual code, XML, WSDL, etc.

Type: Recommendation document

Milestone	Target Date	Completed	Completion Date
First Draft for Review	Mar 2007		
Submission for Comment	Nov 2007		
Published	Mar 2008		

Title: Naming Mechanisms and their Support in OGSA-DMIS

Abstract: Although initially the <our group name> will focus on file movement, where possible we intend to design for general data movement. A significant aspect of this will be how data is named, and represented during communication with the service. In this document we will discuss existing naming systems and specify what will be supported by <our group name>

Type: Informational Document

Milestone	Target Date	Completed	Completion Date
First Draft for Review	Mar 2007		
Submission for Comment	Nov 2007		
Published	Mar 2008		

Title: Data Movement Transport Protocols and their use in OGSA-DMIS

Abstract: The data movement service will be transport protocol agnostic. However, to be effective, it must be able negotiate an acceptable protocol between the endpoints, and provide protocol specific information in a standard, extensible way. This document will describe how this negotiation will take place, how extensibility will be achieved, and list the initial set of protocols and their interfaces.

Type: Informational Document

Milestone	Target Date	Completed	Completion Date
First Draft for Review	Mar 2007		
Submission for Comment	Nov 2007		
Published	Mar 2008		

Title: WS-I Basic Profile Rendering of Data Movement Functional Specification

Abstract: This document provides explicit WSDL, schema, etc., that defines a WS-I Basic Profile Rendering of the data movement functional specification.

Type: Recommendation Document

Milestone	Target Date	Completed	Completion Date
First Draft for Review	Sep 2007		
Submission for Comment	May 2008		
Published	Sep 2008		

Title: WS-RF Basic Profile Rendering of Data Movement Functional Specification

Abstract: This document provides explicit WSDL, schema, etc., that defines a WS-RF Basic Profile Rendering of the data movement functional specification.

Type: Recommendation Document

Milestone	Target Date	Completed	Completion Date
First Draft for Review	Sep 2007		
Submission for Comment	May 2008		
Published	Sep 2008		

Additional Planning Detail:

By When	Task
15 January 2005	Group has agreed on Charter via email
GGF17	Session 1: Review of existing WSDL
	Session 2: Discussion: What abstract / functional operations and parameters are common? <ul style="list-style-type: none"> What is within scope of the group? Get volunteer(s) to draft document and begin mailing list discussion Naming discussion <ul style="list-style-type: none"> What requirements are known for naming/ Who (both inside and outside of GGF) is working on this? Get volunteer(s) to draft requirements and brief discussion of options.
	Iterate on functional spec and naming documents.
GGF18	Session 1: Work on functional spec and naming documents Session 2: Discussion of transport protocols <ul style="list-style-type: none"> What transports are people interested in What information is required for each What might an agreement interface for this look like How can we make this extensible Get volunteer(s) to draft a document listing transports, associated parameters
	Iterate on functional spec, naming, and transport docs.
	WG Session will discuss updated WSDL
GGF19	Session 1: Review of solid draft of functional spec and naming <ul style="list-style-type: none"> Get volunteer(s) to draft WS-I rendering of functional spec Get volunteer(s) to draft WSRF rendering of functional spec Session 2: Review draft of document on transport protocols
	Iterate on functional spec, WS-I rendering, WSRF rendering. Implement proof of concept prototypes of WS-I and WSRF renderings. Incorporate transport recommendations into functional spec and prototypes.
GGF20	Session 1: Incorporate feedback from rendering efforts into functional spec and discuss rendering issues. Session 2: Incorporate feedback from rendering efforts into transport doc and discuss rendering issues.
	Incorporate rendering feedback, naming, and transport into functional spec. Continue to work on WS-I and WSRF rendering documents and prototypes.
GGF21	Session 1: Final review of functional spec and naming documents. Prepare to submit for comment. <ul style="list-style-type: none"> Find volunteers to implement a second WS-I implementation and a second WSRF implementation. Session 2: Final review of transport doc. Prepare to submit for comment.
	MILESTONE: Functional Spec, naming document, and transport documents submitted for comment.
	Continue work on WS-I and WSRF rendering documents and implementations. Second implementations begin development. Respond to comments on functional Spec, naming, and transport.
GGF22	Session 1: Discuss and respond to comments in functional specs, naming, and transport docs. Session 2: Discuss issues from second implementation, plan for interop testing.

	MILESTONE: functional spec, naming, and transport docs complete
	Conduct interop testing, update specs on knowledge gained from interop testing.
GGF22	Session 1: Work on final draft of WS-I and WSRF rendering documents.in preparation for comments. Session 2: Start work on interop document,
	Continue work on Rendering documents
GGF23	Final edits and review for WS-I and WSRF rendering documents, submit for 60 day comment period.
	Respond to comments
GGF24	Session 1: Discuss future of the WG Session 2: Final response to comments on rendering documents., continue work on interop document.
	Complete the interop document and submit for 60 day comment period
GGF25	Session 1: Respond to comments on interop document Session 2 (if necessary): further discussion of future of the WG.

2.4 Management Issues

This Working Group will have regular phone conference, at least biweekly. Also, if required or requested by WG members, it will join OGSA or OGSA Data F2F meetings. The WG will have joint review discussion with the OGSA-WG and the OGSA-D-WG before every milestone.

2.5 Evidence of commitments to carry out WG tasks

The movement of data is a critical task in many Grid applications. There are 4 existing, but non-interoperable interfaces/implementations. People are working to solve the problem, this working group will allow them to coordinate and reach consensus.

2.6 Pre-existing Document(s) (if any)

Need to add pointers about the existing implementations

2.7 Exit Strategy

Once we have shown inter-operable implementations moving file data, the working group will need to checkpoint and assess whether or not it should terminate, or continue forward with a more general version that can move any nameable data. If we do continue forward, we would follow the same pattern. Continue until we have inter-operable implementations of a more general v2.0 and then re-assess, though I suspect is is likely we would terminate at that point.

2.8 Any other relevant information

This group was generated from a discussion in the OGSA Data working group and thus will coordinate with it to ensure that it fits within the architecture.

3. References

1. OGSA WSRF Basic Profile 1.0, Foster, I., Maguire, T. and Snelling, D. Global Grid Forum, GWD-R, September 2005.
http://www.ggf.org/Public_Comment_Docs/Documents/Oct-2005/draft-ggf-ogsa-wsrf-basic-profile-v43.pdf
2. Web Services Agreement Specification (WS-Agreement), Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Nakata, T., Pruyne, J., Rofrano, J, Tuecke, S., Xu, M. Global Grid Forum, GWD-R, June 2005
<https://forge.gridforum.org/projects/graap-wg/document/WS-AgreementSpecification/en/7>
3. WS-I Basic Profile 1.1, Ballinger, K., Ehnebuske, Ferris, C., D., Gudgin,Liu, C., M., Nottingham, M., Yendluri, P. Web Services Interoperability Organization, Final Material, April 2004
<http://www.ws-i.org/Profiles/BasicProfile-1.1-2004-08-24.html>