INFOD Working Group
Global Grid Forum, Data Area

Administrative Information

Name and Acronym:
Information Dissemination, INFOD

Chairs:
Dieter Gawlick, Oracle Corp., dieter.gawlick@oracle.com
Susan Malaika, IBM Corp., malaika@us.ibm.com

Secretary(s)/Webmaster(s) (both optional):
Secretary: Steve Fisher, S.M.Fisher@rl.ac.uk
Webmaster: Chris Kantarjiev, chris.kantarjiev@oracle.com

Proposed Email list:
infod-wg@ggf.org

Proposed Web page:
http://forge.ggf.org/projects/infod-wg

Charter
Focus/Purpose

The purpose of this charter is to develop a model and define operations Information Dissemination (ID) that support asynchronous data and event distribution. Third-party data delivery and publish-subscribe in real-time environments are major use cases. Data replication is also a use case of interest.

For background information on Information Dissemination, which was previously known as 'Grid Data Distribution', please review:
https://forge.gridforum.org/docman2/ViewCategory.php?group_id=49&category_id=506.

Scope

The scope of the charter is three-fold:
- Define a model and operations to support ID
- Define typical ID scenarios and a high-level interface for these scenarios. These interfaces will be designed to simplify the use of ID for the majority of the anticipated users.
- Position the ID with respect to related standards and standards activities, especially to WS-Notification. Determine which of these evolving standards may be leveraged as infrastructure once they are ratified.

This working group is mainly concerned with the standardization of the notification mechanisms required when disseminating information, e.g., how to indicate that a portion or complete copy of a file is required at another site, but not the mechanism of distributing the data itself, e.g., through gridftp. For example describing the need to disseminate portions of High Energy Physics datasets or databases is covered, but the method of distribution of such data is not the subject of this working group.

Goals

Two documents were delivered in GGF10:
- The Functional Specifications of Information Dissemination, outline, GGF10 – Informational document
- Typical Scenarios of Information Dissemination, draft, GGF10 – Community Practice document

Deliverable/Milestone 1:
- The Functional Specifications of Information Dissemination, 'preliminary' draft, GGF12 – Recommendations document
- Informational document Typical Scenarios of Information Dissemination, draft, GGF12 – Community Practice document

Deliverable/Milestone 2:
- The Functional Specifications of Information Dissemination, draft, GGF13 – Recommendations document
- Informational document Typical Scenarios of Information Dissemination, final, GGF13 – Community Practice document

Deliverable/Milestone 3:
- The Functional Specifications of Information Dissemination, final, GGF14 – Recommendations document

Management Issues

Management issues are resolved by having dedicated co-chairs who are aware of the time commitment a GGF working group demands and deliver on that commitment. We feel that the 2 candidates we propose are perfect for these roles, as they will commit to at least 4 hours per week to run the working group effectively.

Evidence of commitments to carry out WG tasks

Susan has been deeply involved in various GGF activities; Dieter has been working consistently on ID for about a yearA list is available of at least 20 people who'd like to be observers, and 10 who'd like to be active.

Pre-existing Document(s)

As mentioned above, refer to https://forge.gridforum.org/docman2/ViewCategory.php?group_id=49&category_id=506 for some background information on 'Grid Data Distribution' which predates the new 'Information Dissemination' model. Further information can be found at: https://forge.gridforum.org/projects/infod-wg

WS-Notifications specifications can be found at the Oasis Website here
- http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsn

Exit Strategy

The exit strategy for the INFOD working group is to deliver the final version of the following document, which is the last milestone as defined above:

The Functional Specifications of Information Dissemination, final - Recommendations Document.

Any other relevant information

None.

Section 1 - Seven questions: Evaluation Criteria (from GFD-C.3)

When considering the formation of this group, the Steering Group will wish to ensure that every WG has clear and focused objectives, and has demonstrated support from the community. The Steering Group will consider the following seven issues (taken from GGF document GFD-C.3).

1.1   Is the scope of the proposed group sufficiently focused?

Is the group attempting to produce everything from beginning to end (a survey of the state-of-the-art, plus use cases, plus a requirements analysis, plus recommendations documents) or is it focused on only one or two of these areas? Is there more than one type of standard being proposed (Architecture/framework vs. information model (schema) vs. API vs. Protocol)? Is the topic area too specific or too broad (for example, overlap with other GGF WGs may indicate "too broad")? Are the milestones reasonably achievable in the proposed timeframe (1-2 years for a WG)?

The INFOD WG intent is to produce the documents mentioned in the section above (Community Practice and Recommendations documents). The group will provide a requirement analysis, compare the proposed standard to existing and major evolving standards, and define the ID model and its interfaces as well as specializations for major use cases. The milestones seem to be reasonable to the members of the working group, who have already started the work for a little less than a year.

Please refer to section 1.4 below, for a more detailed list of other working groups' overlap with the INFOD work, and how that overlap will be addressed in order to keep the original intent for the INFOD group's focus.

1.2   Are the topics that the group plans to address clear and relevant for the Grid research, development, industrial, implementation, and/or application user community?

ID has been recognized as an important element of many use cases. Major examples are replication and dissemination of query results. ID is a core element of intra- and inter-company communication. Therefore, the topics addressed by the INFOD WG are very much relevant for Grid research, industrial, implementation, development and user community.

1.3   Will the formation of the group foster (consensus–based) work that would not be done otherwise?

Does the group foster standards or practices that are greater than the work done by any single group (taking advantage of GGF to come together on neutral ground)?   How many distinct groups, institutions, and regions of the world are participating in this effort? (GGF activities typically have membership drawn from more than a single research group, institution or project).

There is no current standard that supports communication with the level of functionality or operational characteristics that will be supported by ID. The functionality and the operational characteristics are needed for mission critical applications. Specialization will simplify the use of ID for the major use cases, e.g., replication.

The challenge for this WG will be to attract and include a diverse set of contributors who belong to academia, industrial, end-user, etc. so that the standards being developed are then implemented in these varied communities.

1.4   Do the group's activities overlap inappropriately with those of another GGF group or to a group active in another organization such as IETF or W3C?

What is the nature and extent of any overlap? The proposed group may still be formed, or the GFSG may recommend that the work be done within the existing GGF (or external) group.

The INFOD WG overlap with other related working group is as follows:

- The scope of INFOD WG overlaps with the scope of the **OREP** GGF WG, as Replication is fully covered by INFOD. It is proposed that replication be taken out of OREP and moved under INFOD. A discussion as to RLS work, which is the remainder of the OREP work group work items, needs to address whether RLS is moved to INFOD, or not.
- In addition, INFOD WG addresses some of the requirements of **DAIS** GGF WG, but does not overlap with it. The distinction between synchronous and asynchronous access is clearly outlined.
- It is also necessary to resolve the issue of overlap with the proposal that was submitted in the face-face meeting of **OGSA** GGF WG on February 10, 2004 in San Diego. The co-chairs of INFOD believe that that the INFOD WG may very well consume that work.
- The scope of the INFOD WG clearly states that there is a need to position INFOD with respect to **WS-Notification** (not part of GGF). So overlap is not so much the issue as 'positioning': WS-Notification could provide the underlying infrastructure that INFOD could be based upon. So the question is, where should INFOD-related work reside, at GGF or under IEFT/W3C/OASIS? The co-chairs of INFOD believe that GGF is the right place, as INFOD provides a model that solves a Grid-specific problem: how do I disseminate the information in a distributed, asynchronous, efficient, scalable, reliable and real-time manner… all of these issues are Grid-related.

1.5   Are there sufficient interest and expertise in the group's topic, with at least several people willing to expend the effort that is likely to produce significant results over time?

How much experience do the participants collectively have in the proposed area of work? How committed are the participating individuals? An attendance list or an email subscriber list is a very weak indication of commitment; a list of people who have attended multiple teleconferences is somewhat better; a list of individuals who have committed to specific tasks, or who have made non-trivial time commitments, is much better. Additional evidence could include statements from organizations stating that they will dedicate resources (people) to participate in the group, and statements from participants expressing their personal, compelling need for the output of the group. Can you, the chair, commit to at least 4 hours per week to run this group? (Please address this question directly).

From a database perspective, the co-chairs clearly feel that the coverage is more than adequate. For others, like files, as mentioned earlier, expertise exists in the academia and industrial and user communities so the co-chairs are aware that their tasks include gathering the appropriate interest and having people dedicated to implementing solutions based on the proposed model.

From a co-chair perspective, yes, there is dedication to this work. Actually, if you look at the amount of time already spent on the 'Grid Data Distribution' work, it is clear that dedication has already been demonstrated, and continues to be.

The BoF session in Berlin should clearly the interest in this subject.

1.6   Does a base of interested consumers (e.g., application developers, Grid system implementers, industry partners, end-users) appear to exist for the planned work?

How broadly applicable will the output of the WG output be? Does the WG have true clients of its work? Such interest can be measured by the interest of industry partners, grid deployment projects, and other groups committed to implement the recommendations or adopt the results.

The success of a working group requires "buy-in" from a broad set of constituents who will use the output of the group. It is useful to indicate what is the target set of consumers in the community. While not necessarily a requirement for approval, it is essential that the organizers comment on the relationship of the work, and the level of interest, from large segments of the Grid community such as major software projects, architecture activities, etc.

Most of the clients of DAIS are clients of the INFOD-WG as data access in a Grid environment very often includes data replication in the background for Quality of Service. Also, all clients that required Third Party data delivery in DAIS are waiting for the INFOD solution.

The BoF in Berlin demonstrated the high interest in ID.

1.7   Does the GGF have a reasonable role to play in the determination of the technology?

What other organizations are working in similar areas? Is the GGF the right place for this work? Is it clear how the proposed WG will coordinate with related efforts?

The answer to this question has already been addressed in 1.4.