

Workflow Management Research Group Proposal

Status of This Memo

This memo provides information to the Grid community regarding GGF working group formation. Distribution is unlimited.

Copyright Notice

Copyright © Global Grid Forum (2003). All Rights Reserved.

Abstract

This document is provided in support of the formation of the Workflow Management Research Group. The document include the charter for the group as well as the answer to the “Seven Questions”.

Contents

Abstract	1
1. The Charter	2
Workflow Management Research Group	2
Global Grid Forum, Scheduling and Resource Management Area	2
2. Seven questions: Evaluation Criteria (from GFD-C.3)	3
1. Is the scope of the proposed group sufficiently focused?	3
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?	3
3. Will the formation of the group foster (consensus-based) work that would not be done otherwise?	4
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?	4
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?	4
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?	4
7. Does the GGF have a reasonable role to play in the determination of the technology?	5
Editor Information.....	5
Intellectual Property Statement.....	5
Full Copyright Notice	6

1. The Charter

Workflow Management Research Group

Global Grid Forum, Scheduling and Resource Management Area

Chairs:

Ewa Deelman, deelman@isi.edu

Jim Blythe, blythe@isi.edu

Miron Livny, miron@cs.wisc.edu

Secretary(s):

Tbd

Email list: wfm-rg@ggf.org, to subscribe email majordomo@gridforum.org with the following command in the body of your email message: `subscribe wfm-rg`

Web page: (temporary) <http://www.isi.edu/~deelman/wfm-rg>

Focus/Purpose: The purpose of this group is to explore, evaluate and propose workflow representation and mapping techniques that enable the high-level description of application workflows and their execution in the Grid environment. An area of emphasis will be on scientific workflows, which describe the behavior of complex applications. We will investigate workflow issues as they pertain to the entire workflow lifecycle, starting from the reuse of generic idioms for composing groups of tasks, to the tracking of the progress of a complex task or workflow from the initial formulation of desired end tasks or data products through to execution.

We will investigate workflow issues as they pertain to the entire workflow lifecycle, from initial formation, for example by the reuse of generic idioms for composing groups of tasks to achieve desired end tasks or data products, to tracking the progress of a complex workflow through to execution, including refinement and repair.

The group will aim to work closely with other areas and groups within the GGF such as: Distributed Resource Management Application API (DRMAA-WG), Data Access and Integration Services (DAIS-WG), Open Grid Services Architecture (OGSA-WG), Semantic Grid (SEM-RG), and Advanced Programming Models (APM-RG).

Additionally the group will follow the developments in other industry and research bodies, such as the W3C.

Scope: Workflows today play a very important role in the scientific and business communities. Applications today are not monolithic entities, rather complex workflows that need to be realized in a given environment. The Grid, as an workflow execution environment is very complex because of its dynamic and heterogeneous nature. The workflow management research group will focus on a variety of aspects of workflow management in the Grid environment. We will identify existing technologies as well as open issues that pertain to the entire workflow lifecycle.

Although one focus of the group is on scientific workflows, business workflows are also within the scope.

Where appropriate, the research group may spin off working groups.

Proposed tasks:

1. Provide a forum to discuss and share best practice in projects that use workflows to represent complex applications.
2. Track workflow activities within W3C and inform the Grid community on what tools and ideas are relevant in the Grid environment.
3. To evaluate the applicability of web services coordination to grid services.
4. To explore various existing technologies applicable to workflow management.
5. To identify the open issues in the area.

Goals:

Produce the following documents:

Open Issues in Workflow Management in Grids (GGF 10)

Overview of Business Workflows and their relevance in the Grid Environment. (GGF11)

Workflow Specification Languages (GGF11)

A roadmap for representing and reasoning about workflows. (GGF12)

Overview of Workflow Management Research in Web Services and Applicability to Grid Services (GGF 13)

Overview of Workflow Execution Systems in Grids (GGF 14)

Workflow Specification Languages (GGF 15)

Exit Strategy:

Given that workflow management is just beginning to be addressed in Grids, this research group may extend beyond the 2 year limit.

2. Seven questions: Evaluation Criteria (from GFD-C.3)**1. Is the scope of the proposed group sufficiently focused?**

This group is focused primarily providing a forum for workflow management researchers to share ideas as well as on conducting surveys of the state-of-the art in the area and identifying open issues in the research area. The group identified a specific list of documents that will be produced by this group. There is some overlap with other research group such as the GCE and Sem-Grid research groups. However, this group has discussed the overlaps with these groups and brought up the issue during the BOF at GGF9 and discovered that there is a need for a group that will be solely focused on the workflow management issues: workflow composition, refinement and execution.

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?

The chairs of the group have been working with many application scientists over the years, exploring issues of composition and execution complex applications in the Grid environment. Specifically, the chairs have been working with physicists, astronomers and biologists to understand the requirements for workflow management. It is clear that these communities would

greatly benefit from the surveys conducted by the research group in formulating their future directions. Additionally the grid community will benefit from the identification of the important open issues in the area in order to target those in their research. It is also expected that there is a need to quantify the differences and similarities between business and scientific workflows in the hopes of leveraging the technologies developed within the business community.

3. Will the formation of the group foster (consensus-based) work that would not be done otherwise?

The attendance at the GGF9 BOF for the group was very well attended (approximately 75 participants) and brought together people from various groups and organizations. Many people signed up to the mailing list, including people from various research groups in the US and Europe. GGF is the only forum that could bring these participants together in order to conduct the research.

Many separate groups are now coming to look at workflows in their research, e.g. GriPhyN, SCEC. However, only by acting within this research group can researchers looking at workflow find the critical mass and the combined breadth of vision required to produce the solid foundations needed for reasoning about workflows on the Grid.

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?

There are three main groups within GGF that overlap with the proposed workflow management research group: the Grid Computing Environments (GCE), the Semantic Grid (Sem-Grid) and the Lifesciences research groups. The focus of GCE is primarily on the role of workflows within the portal environment whereas the focus of Sem-Grid is on workflow composition at the metadata/semantic level and finally the focus of Lifesciences is on generating the requirements of the lifesciences community. Although there is some overlap with these research group, the proposed workflow management research group focuses on additional and different aspects of the workflow management, not only workflow composition but also workflow refinement, scheduling and execution. The group will also conduct cross-cutting surveys that are proposed by the other groups.

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?

The chairs are ready to commit the time necessary to make this group a productive and successful group within ggf. As such the chairs commit to spend at least the requisite 4 hours per week on running the group.

The chairs believe that there is evidence of interest in working within the group's scope. During the BOF, at least 3 different group participants (not including the chairs) have committed to work on 3 different documents (as indicated in the minutes of the BOF).

The chairs have also worked in the area of workflow management for a long time and thus have enough experience in the area of the proposed work. The chairs believe that group participants have the necessary experience as well.

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?

The participants in the group are involved in many projects both research and deployment, such as the NSF funded GriPhyN, NVO, SCEC, iVDGL projects and others. The participants also release workflow management software as part of these and other projects. As such the work conducted within the proposed group will be directly transitioned to the "clients" of these projects, namely the physics, astronomy and seismology communities as well as others.

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

There are efforts within W3C to standardize on languages for workflow composition of web services. Obviously, there are some commonalities between web service and grid service composition and the proposed group will follow closely the developments within W3C. However, web service workflow specifications typically do not focus on allocating tasks to resources, because a service is often inseparable from its host. Moreover, the dynamic nature of the grid environments brings new challenges to workflow composition (discovery of workflow components), workflow refinement (discovery of appropriate resources for example) and workflow execution (dealing with failures). The GGF is the only place where these issues can be addressed. In terms of working with other GGF groups on the topic of workflow management, the proposed group has formed ties with these group and is collaborating with them closely, for example in running a workshop at GGF10 related to workflow issues.

Editor Information

Ewa Deelman
Information Sciences Institute
University of Southern California
4676 Admiralty Way, suite 1001
Marina Del Rey, CA 90292

James Blythe
Information Sciences Institute
University of Southern California
4676 Admiralty Way, suite 1001
Marina Del Rey, CA 90292

Miron Livny
Computer Science Department
University of Wisconsin Madison
1210 W. Dayton St.
Madison, WI 53706-1685

Intellectual Property Statement

The GGF takes no position regarding the validity or scope of any intellectual property or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; neither does it represent that it has made any effort to identify any such rights. Copies of claims of rights made available for publication and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of

such proprietary rights by implementers or users of this specification can be obtained from the GGF Secretariat.

The GGF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights which may cover technology that may be required to practice this recommendation. Please address the information to the GGF Executive Director.

Full Copyright Notice

Copyright (C) Global Grid Forum (2003). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the GGF or other organizations, except as needed for the purpose of developing Grid Recommendations in which case the procedures for copyrights defined in the GGF Document process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the GGF or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE GLOBAL GRID FORUM DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE."