Information Services for Smart Decision Making: An ESI Event Theme Midterm Evaluation

Jennifer M. Schopf, June 13 2005

Theme Overview

Grid computing resources and services can advertise a large amount of data for many different use cases. These include providing data so that resource brokers can locate computing elements appropriate for a job, streaming data so application steering adjustments can be made during runtime, and notifying system administrators when changes in system load or disk space availability occur in order to identify possible performance anomalies.

We proposed to organize a series of events, public lectures, outreach to UK eScience projects, and publications around the topic of Grid information services and their use in decision making. Specifically, we will be interacting with people building monitoring systems and trying to better understand how they can be used with schedulers, by operations and system administrators, and to better understand performance faults in the context of grid applications.

About the Research Leader (RL)

Jennifer M. Schopf is a Scientist at the Distributed Systems Lab, part of the Mathematics and Computer Science Division at Argonne National Lab, and is spending the year as a researcher at the eScience Institute in Edinburgh, UK. She is a member of the Globus Alliance, and technology coordinator for the Monitoring and Discovery Service, and spent 5 years helping to establish the Global Grid Forum standards body.

In addition to her work with the Globus MDS, in the past ten years, Schopf has co-edited a book and co-authored over 30 journal papers, book chapters, and refereed conference submissions, in large part related to information services, predictions, and their use in scheduling in Grid environments. She organized two workshops on the topic of Grid performance, and has received multiple grants in support of this work.

Schopf received a BA in Computer Science and Mathematics from Vassar College, and MS and PhD degrees from the University of California, San Diego in Computer Science and Engineering.

Planned Theme Events

As a midway report for this theme, completed aspects are in regular font, planned/tentative are in italics.

<u>A set of public lectures at eSI</u>

Jennifer Schopf, "Distributed Monitoring and Information Services for the Grid", NeSC Public Lecture, Edinburgh, January 10, 2005

Additional possible speaker include:

Stephen Jarvis
Brian Tierney
Martin Humphrey

Application and research-oriented events at eSI, organized by the RL

Grid Performance Workshop, June 22-23, 2005
Grid Scheduling Workshop, Autumn 2005

Application and research-oriented events associated with the theme at other venues or organized by research colleagues

Networking for Non-Networkers, June 2005, NeSC
Others TBD

Hands-on or technology-oriented events at eSI, organized by the RL

GridFTP days, January 2005
Globus Week, April 2005

Hands-on or technology-oreiented events associated with the theme at other venues, organized by the RL

GLUE Schema 1.2 workshop, February 2005, RAL

Related monitoring and Grid talks at UK eScience Centres promoting the theme
Including visits to

- Rutherford Appleton Laboratory (January 2005)
- Warwick University (May 2005)
- Newcastle University (June 2005)
- *Leeds University (July 2005)*
- *Imperial College (Autumn 2005)*

Related monitoring and Grid talks at international venues promoting the theme
Including lectures:

- "Grid Monitoring and Information Services: Globus Toolkit MDS4 & TeraGrid Inca" LCG Operations Workshop, CERN November 2-4 2004.
- "Globus Toolkit Monitoring and Discovery System: MDS4" Technology Review Talk, Argonne Booth, SuperComputing Nov 8-11, 2004,
- "Performance Inside: Performance Monitoring and Diagnosis for NMI Software and Applications", NMI Meeting December 2004,
- "Distributed Monitoring and Information Services for the Grid", MSc in High Performance Computing Seminar, EPCC Edinburgh, January 28, 2004;
- "Distributed Monitoring and Information Services for the Grid, Master Class, University of Amsterdam, Netherlands, April 2005

Outreach to UK eScience communities

- Work with OGSA-DAI and ELDAS to define performance benchmarks
- Work with NGS to include Inca system monitoring for deployments

A set of reports and research papers associated with the theme will also be published, Including:

- Globus Week Meeting Report
- *Grid Performance Workshop Meeting Report*
- *Grid Scheduling Workshop Meeting Report*
- Benchmarking for OGSA-DAI (submitted to AHM 05)
- Analysis of MDS4 performance (submitted to SC05)
- *Possible special issue on Information Systems for Smart Decisions in the Journal of Grid Computing*

Poster/Content including

One page glossy spread
Poster for NeSC display
Poster for NeSC booth at AHM and SC '05

Coverage Metrics for Theme:

1. Span of low – high level
    - 1/3 low, 2/3 medium or high
2. Breadth of international involvement
    - Strong international involvement, slightly weighted to US over EU
3. Breadth of field involvement
    - Performance workshop and NNFN workshop should both include strong outside-CS involvement
4. Breadth of knowledgeability of participants
    - Full coverage
5. Breadth of organisational type: university, research institute/lab, government labs, R & D in industry
    - Full breadth covered by events or planned events
6. Coverage of the theory to practice pipeline
    - RL lead events are strongly pragmatic in nature, associated events are more theoretical in nature

MidTerm Evaluation Summary

Events, lectures, and talks taking place and planned offer compelling evidence that this theme will be a success. However, this theme needs closer collaboration in terms of produced materials (glossy sheet and posters) and PR to outside sources. The establishment of a theme website is needed, as is better advertisement of the theme concept. Additional visitors/events should be discussed, and clarification is needed with respect to associated workshops. As this is the first prototype theme this is to be expected.