

# Putting the Globus Toolkit in Its Place: The Grid “Ecosystem”

Lee Liming  
Argonne National Laboratory  
The Globus Alliance

## Abstract

The Globus Alliance aims to provide solutions to the most persistent and vexing problems that come up in Grid projects and applications. Our solutions to date are collected in the Globus Toolkit and these solutions are used in many Grid applications and systems.

While the Globus Toolkit makes it easier to conduct Grid-based projects, the challenges are still far from easy and the Globus Toolkit does not provide a “turnkey” solution. Success in a Grid project depends on a clear vision of the problem(s) to be solved, awareness of relevant tools (both within and beyond the Globus Toolkit), and a strategy for applying the technology.

This half-day tutorial provides answers to critical questions for Grid project planners and product developers, including:

- What types of problems is the Grid intended to address? How far does the Globus Toolkit go toward solving these problems? How does the Globus Toolkit fit into a Grid project or product strategy?
- What do you need besides the Globus Toolkit to have a useful solution to your problem?
- What additional tools are available from the Grid community, and what are some examples of how these tools have been used in successful projects?

The Globus Toolkit will be put into context, and examples and roadmaps for the most common uses of the Globus Toolkit will be provided.

## Goals

Attendees will learn answers to the questions posed in the abstract above. In summary, attendees will learn how to account for the Grid (broadly speaking) and the Globus Toolkit (more narrowly) when planning useful products, applications, and/or Grid projects. This in turn will provide attendees with more tools and materials for producing better project and product development plans.

## Content Level

60% beginner / 30% intermediate / 10% advanced

## Targeted Audience

- Project leaders, team leaders, managers, IT executives, product managers, principal investigators, researchers, program managers, scientists, engineers
- Anyone who has recently or soon will accept a position of responsibility in a current or potential Grid project or product development activity

Attendees should be familiar with the basic principals of information technology. For example: general computer and network architecture, general software engineering processes, client/server systems, databases, current types of commercial IT products, basic internet concepts.

## Relevance to Attendees

Grid technologies offer solutions to problems in the following areas: computational science, programming tools, scientific data management, visualization, and high performance computing. So far, these technologies fall short of delivering “turnkey” solutions to most problems. Experience has unequivocally shown that in order to use Grid technology to solve a problem, one must first understand how the many Grid tools are meant to work together, select the appropriate components, and integrate them with commodity and application-specific components. Many tutorial attendees are present specifically to learn about these technologies and to find out how they can use them to make their work easier.

## General Description of Tutorial Content

This half-day tutorial will be presentation-based with no hands-on component.

As indicated in the notes above and the attached outline, the tutorial provides a context for Grid technology and the specific components provided by the Globus Toolkit using examples from scientific and commercial projects. This context and these examples are then used to ground the subsequent discussion.

The central material of the tutorial is an overview of the capabilities provided by the Globus Toolkit and related Grid technologies, using specific examples to explain their relevance to problems that the audience may be familiar with. Particular emphasis will be placed on how to apply these technologies in specific circumstances to achieve specific

goals. The tutorial will provide several prototypical “roadmaps” for building applications and systems drawn from experiences gained in previous and ongoing Grid projects.

Throughout the tutorial, attention will be drawn to the central “integration” challenge that is faced by all current Grid projects: the need to pull together a combination of commodity/off-the-shelf components, Grid community components, Globus Toolkit components, and application-specific components in order to build a complete application or system.

## Reuse of Materials

The materials produced for this tutorial are licensed for reuse by anyone under the terms of the Globus Toolkit Public License (<http://www.globus.org/toolkit/license.html>).

## Detailed Outline

- I. What types of problems is the Grid intended to address?
  - A. Computation intensive (specific examples)
  - B. Data intensive (specific examples)
  - C. Distributed collaboration (specific examples)
  - D. Review of the problems common to all three areas
- II. How far does the Globus Toolkit go toward solving these problems?
  - A. Three ways of viewing the world (end user, application developer, system administrator)
  - B. Grid Architecture
    1. The Grid “hourglass” model and the Grid “layered architecture”
    2. Use of Web Services framework
    3. Critical role of standards
    4. The Grid’s “Resource Layer” (vs. other layers)
  - C. How current Grid projects use the Globus Toolkit
    1. Grid3
    2. NEESgrid
    3. SDSS
- III. How does one install and configure the Globus Toolkit, and then what does one do with it? What do you need besides the Globus Toolkit to have a useful solution to your problem?
  - A. What to Download
  - B. Installation
  - C. Configuration
  - D. The “missing pieces” (what *isn't* included)
    1. Commodity (“off the shelf”) components
    2. Application-specific components
    3. Collective (“aggregation”) components
- IV. What other Grid software works well with the Globus Toolkit, and what can one do if those are added?
  - A. Security tools
  - B. Monitoring/discovery tools
  - C. Computing/processing tools

- D. Data movement tools
  - E. Data management tools
  - F. System management tools
  - G. System Packaging/Distribution tools
- V. Roadmaps
- A. Computation-intensive application: Grid3
  - B. Data-intensive application: Earth System Grid
  - C. Distributed collaboration application: NEESgrid

## Lee Liming

Lee Liming is the Manager of the Distributed Systems Laboratory at Argonne National Laboratory, the "Chicago home" of the Globus Alliance. The Distributed Systems Laboratory hosts computer scientists, software architects, and software developers working together to develop solutions to the challenging problems associated with distributed collaboration in science and engineering. Lee has worked for fifteen years on distributed systems issues in both academia and industry, with experience ranging from system administration and software development to product and project management.

Lee became a member of the Globus Alliance in 1999, and his roles in the Alliance have included project leader, technology developer, and technical manager. Early Grid communities in which he participated include the NASA Information Power Grid, the ASCI DisCom program, the National Computational Science Alliance, and NEESgrid. He currently has prominent roles in the NSF-sponsored NEESgrid and NSF Middleware Initiative/GRIDS Center projects.

The software development team whose work Lee directs at Argonne National Laboratory and the University of Chicago support the DSL's involvement in a wide range of Grid-related projects, from infrastructure-focused projects like TeraGrid and Grid3 to science-focused projects like the Earth Systems Grid and FusionGrid. This team also, of course, contributes a sizable portion of the code in the Globus Toolkit.

Lee has presented tutorials on behalf of the Globus Alliance at project meetings and conferences from Tsukuba, Japan to Krakow, Poland. Most recently, he presented at the Air Force Research Laboratory's 2003 Cyberinfrastructure Seminar Series and at the GGF-11 meeting in Hawaii. While most of Lee's writing takes the form of project reports and white papers, his most recently-published article was *The Role of Standards in the Grid*, co-authored with Ian Foster, appearing in the Spring 2004 *DOD Software Tech News*.