

## Grid Computing

---

### Globus Toolkit Future Directions & Related Software

Paul A. Farrell  
Fall 2006

Based on

**Ian Foster, Globus Toolkit version 4**  
Tutorial, 1st Intl. Conf. on e-Science and Grid Computing,  
Melbourne, Australia, December 12, 2005

---

Paul A. Farrell 2006 KENT STATE Grid Computing 1

## The Globus Commitment to Open Source

---

- Globus was first established as an open source project in 1996
- The Globus Toolkit is open source to:
  - allow for inspection
    - for consideration in standardization processes
  - encourage adoption
    - in pursuit of ubiquity and interoperability
  - encourage contributions
    - harness the expertise of the community
- The Globus Toolkit is distributed under the (BSD-style) Apache License version 2

---

Paul A. Farrell 2006 KENT STATE Grid Computing 2

## The Future: Structure

---

- NSF Community Driven Improvement of Globus Software (CDIGS) project
  - 5 years of funding for GT enhancement
  - Regular Globus **roadmaps** outlining plans
- GlobDev <http://dev.globus.org>
  - Apache-like community development site
  - Community governance of components
  - “Globus Toolkit” & other related software
  - Open for business early 2006
  - “Globus Alliance” = “GlobDev committers”

---

Paul A. Farrell 2006 KENT STATE Grid Computing 3

## GlobDev

---

- The current set of Globus components will be organized into several “Globus Projects”
  - Projects release products
- Each project will have its own group of “Committers”
  - committers are responsible for governance on matters relating to their products
- The “Globus Management Committee” will
  - provide overall guidance and conflict resolution
  - approve the creation of new Globus Projects

---

Paul A. Farrell 2006 KENT STATE Grid Computing 4

## The Future: Content

---

- We now have a solid and extremely powerful Web services base
- Next, we will build an expanded open source Grid infrastructure
  - Virtualization
  - New services for provisioning, data management, security, VO management
  - End-user tools for application development
  - Etc., etc.
- And of course responding to user requests for other short-term needs

---

Paul A. Farrell 2006 KENT STATE Grid Computing 5

## Short-Term Priorities: Security

---

- Improve GSI error reporting & diagnostics
- Secure password, one-time password, Kerberos support for initial log on
- Trust roots, use of GridLogon
- Identity/attribute assertions in GT auth. callouts (e.g., Shib, PERMIS, VOMS, SAML)
- Extend CAS admin & policy support
- Security logging with management control for audit purposes

---

Paul A. Farrell 2006 KENT STATE Grid Computing 6

## Short-Term Priorities: Data Management

---

- Space & bandwidth management in GridFTP
- Concurrency in globus-url-copy
- Priorities in RFT
- Data replication service
- Enhance policy support in data services
- Physical file name creation service
- Scalable & distributed metadata manager

---

Paul A. Farrell 2006 KENT STATE Grid Computing 7

## Short-Term Priorities: Execution Management

---

- Implement GGF JSDL once finalized
- Advance reservation support
- Policy-driven restart of "persistent" jobs
- Improved information collection for jobs
- Improved management of job collections
- Credential refresh
- Development of workspace service
- Integration of virtual machines (Xen, VMware) and associated services
- Windows port of WS GRAM

---

Paul A. Farrell 2006 KENT STATE Grid Computing 8

### Short-Term Priorities: Information Services

---

- Many more information sources, including gateways to other systems
- Automated configuration of monitoring
- Specialized monitoring displays
- Performance optimization of registry
- Archiver service
- Helper tools to streamline integration of new information sources

---

Paul A. Farrell 2006 KENT STATE Grid Computing 9

### Short-Term Priorities: WS Core

---

- Streamlined container configuration
- Remote management interface
- Dynamic service deployment
- Service isolation: multiple service instances
- WS-Notification, subscription performance
- Full functionality in C WS Core
- Optimized WS-ServiceGroup support
- WS-SecureConversation support

---

Paul A. Farrell 2006 KENT STATE Grid Computing 10

### The Globus Ecosystem – Related Tools

---

- Globus **components** address core issues relating to resource access, monitoring, discovery, security, data movement, etc.
  - GT4 being the latest version
- A larger Globus **ecosystem** of open source and proprietary components provide complementary components
  - A growing list of components
- These components can be combined to produce **solutions** to Grid problems
  - We're building a list of such solutions

---

Paul A. Farrell 2006 KENT STATE Grid Computing 12

### Many Tools Build on, or Can Contribute to, GT4-Based Grids

---

- Condor-G, DAGman
- MPICH-G2
- GRMS
- Nimrod-G
- Ninf-G
- Open Grid Computing Env.
- Commodity Grid Toolkit
- GriPhyN Virtual Data System
- Virtual Data Toolkit
- GridXpert Synergy
- Platform Globus Toolkit
- VOMS
- PERMIS
- GT4IDE
- Sun Grid Engine
- PBS scheduler
- LSF scheduler
- GridBus
- TeraGrid CTSS
- NEES
- IBM Grid Toolbox
- ...

---

Paul A. Farrell 2006 KENT STATE Grid Computing 13

Grid Research Integration Deployment and Support Center

**GRIDS**

## Documenting The Grid Ecosystem

**The Grid Ecosystem: Software Components for Grid Systems And Applications**

Software is a vital part of Grid projects. The open source community has produced a wide variety of Grid software components. Understanding the capabilities of each component, the strengths and weaknesses of each, and the types of problems that are currently covered (or not covered) by these components can be a bewildering challenge for new Grid projects. Moreover, each new Grid project has its own set of requirements and constraints.

**The Grid Ecosystem: Software Components for Grid Systems And Applications**

software plays in Grid projects and descriptions of many of the software tools that GRIDS Center members and our partners have used successfully in ambitious Grid applications. You may choose from the following areas, which are reflected in the navigational submenu to the left:

- An Ecosystem of Grid Components** describes the GRIDS Center's general approach to using software components in Grid projects and applications.
- Security** describes a number of useful software tools for meeting the security requirements in Grid systems.
- Monitoring and Discovery** describes software components that can provide monitoring and discovery features in Grid systems.
- Computation** describes software tools that can be used to manage computational tasks in Grid applications.
- Data** describes software tools that can be used to manage data and datasets in data-intensive applications.
- Collaboration** describes software for facilitating and encouraging collaboration in distributed projects.
- Packaging and Distribution** describes tools for helping to create integrated software distributions for use in Grid projects.

[www.grids-center.org](http://www.grids-center.org)

Grid Computing 14

## Example Solutions

- Portal-based User Reg. System (PURSE)
- VO Management Registration Service
- Service Monitoring Service
- TeraGrid TGCP Tool
- Lightweight Data Replicator
- GriPhyN Virtual Data System

---

Paul A. Farrell 2006 KENT STATE Grid Computing 15

## Condor-G

- The Condor Project @ U Wisconsin Madison develops software for high-throughput computing on collections of distributed compute resources
- Condor-G is an interface to GRAM created by the Condor team that allows users to submit jobs to GRAM servers

---

Paul A. Farrell 2006 KENT STATE Grid Computing 16

## GridShib

- Allows the use of Shibboleth-transported attributes for authorization in GT4 deployments
  - And, more generally, SAML support
- 2 year project started December 1, 2004
- Participants
  - Von Welch, UIUC/NCSA (PI)
  - Kate Keahey, UChicago/Argonne (PI)
  - Frank Siebenlist, Argonne
  - Tom Barton, UChicago
- Beta software released September 16, 2005

---

Paul A. Farrell 2006 KENT STATE Grid Computing 17

## Handle System

---

- The Handle System from CNRI (<http://www.handle.net>) is a general-purpose global name service enabling secure name resolution over the internet
- The Handle System-GT Integration Project leverages the Handle System for identifier and resolution services through tight integration with GT4's Web services protocols

---

Paul A. Farrell 2006 KENT STATE Grid Computing 18

## MPICH-G2

---

- MPICH-G2, developed at Northern Illinois University and Argonne National Lab, is a grid-enabled implementation of the MPI v1.1 standard
- MPICH-G2 is implemented using the pre-WS GRAM component in GT4; integration with GT4 WS GRAM is expected in the near future

---

Paul A. Farrell 2006 KENT STATE Grid Computing 19

## Nimrod/G

---

- Nimrod is a specialized parametric modeling system from Monash University
- Nimrod/G uses a simple declarative parametric modeling language to express parameter sweep experiments. Based on GT4 WS services, Nimrod/G enables the formulation, execution and monitoring of multiple individual parametric experiments

---

Paul A. Farrell 2006 KENT STATE Grid Computing 20

## Ninf-G4

---

- Ninf-G4, from AIST, is a reference implementation of the GGF standard GridRPC API
- Ninf-G4 provides higher-level programming APIs for the development and execution of parallel applications on the Grid

---

Paul A. Farrell 2006 KENT STATE Grid Computing 21

## PERMIS

---

- PERMIS is an EU-funded Privilege Management service that implements Role-Based Access Control
- Thanks to the work of the UK Grid Engineering Task Force, services running in a Java WS Core container can use PERMIS via GT4's SAML authorization callouts

---

Paul A. Farrell 2006 KENT STATE Grid Computing 22

## SRB

---

- SRB is a package from SDSC providing a uniform interface for connecting to network-based heterogeneous data resources
- GT4's GridFTP includes an interface to SRB data sources, and vice versa

---

Paul A. Farrell 2006 KENT STATE Grid Computing 23

## Sun Grid Engine

---

- Sun Grid Engine is an open source distributed resource management system from Sun Microsystems
- In a collaboration between the London e-Science Centre, Gridwise and MCNC, the Sun Grid Engine has been integrated with GT4

---

Paul A. Farrell 2006 KENT STATE Grid Computing 24

## Grid Tools & Solutions

---

- We list links to related projects on the "Related Software" of the Globus Toolkit web [www.globus.org/toolkit/tools/](http://www.globus.org/toolkit/tools/)
- "Solutions" are documented on the Globus web [www.globus.org/solutions/](http://www.globus.org/solutions/)

---

Paul A. Farrell 2006 KENT STATE Grid Computing 25