Grid Computing

Globus Toolkit Evolution

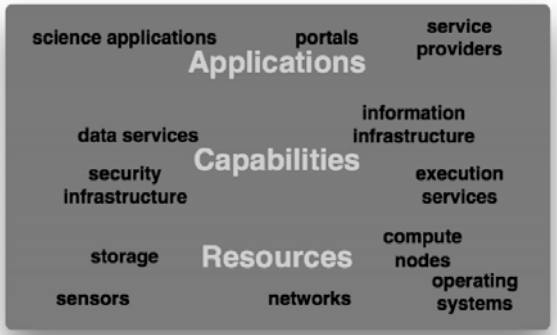
Paul A. Farrell
Fall 2006

The Grid: Core Technologies
Maozhen Li, Mark Baker
John Wiley & Sons; 2005, ISBN 0-470-09417-6

Lisa Childers
Argonne National Laboratory
University of Chicago

International Summer School on Grid Computing 2006
Ischia (Naples) July 9-21, 2006

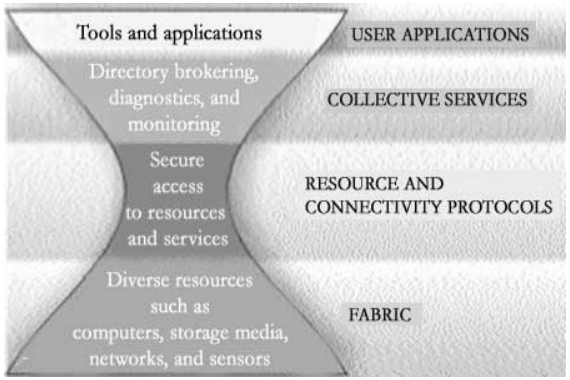
Grid Infrastructure: a Conceptual View



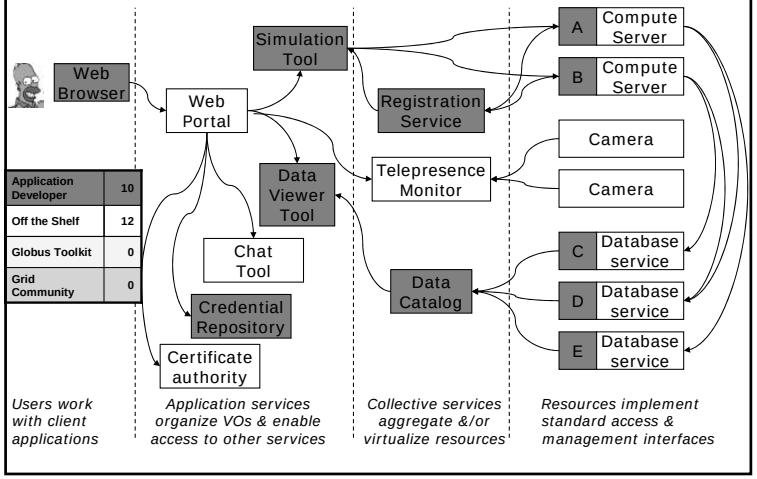
The Role of the Globus Toolkit

- The Globus Toolkit is a collection of solutions to problems commonly found in collaborative distributed applications
- Heterogeneity
  - A focus, in particular, on wrapping heterogeneity for application developers
- Abstractions
  - Supporting general-case patterns and interactions, not specific to a particular application domain
- Standards
  - Capitalize on and encourage use of existing standards (IETF, W3C, OASIS, GGF)
  - GT also includes reference implementations of new/proposed standards in these organizations

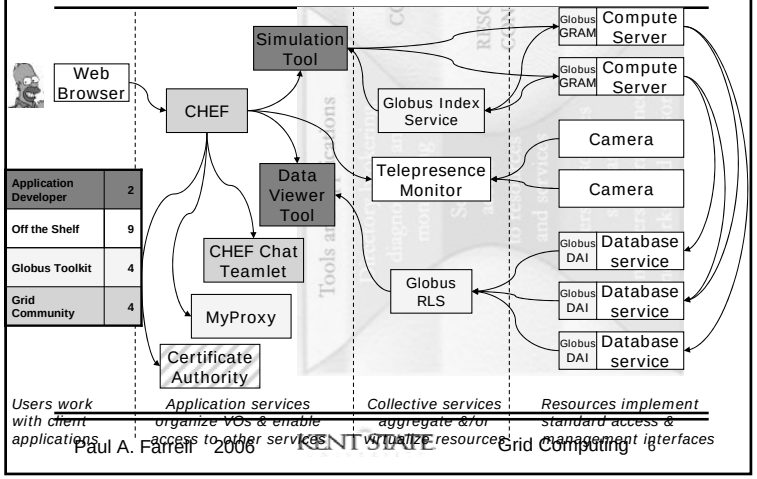
Layers in the Grid



Without the Globus Toolkit



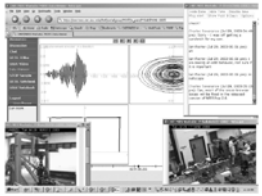
With the Globus Toolkit



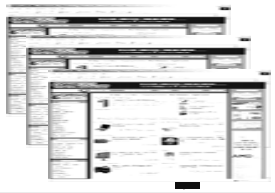
The Globus Toolkit: "Standard Plumbing" for the Grid

- Today the majority of the GT public interfaces are usable by application developers and system integrators
  - Relatively few end-user interfaces
  - In general, not intended for direct use by end users (scientists, engineers, marketing specialists)
- *Not* turnkey solutions, but *building blocks & tools* for application developers & system integrators
  - Some components (e.g., file transfer) go farther than others (e.g., remote job submission) toward end-user relevance
- Better to reuse than reinvent!
  - Plenty of interesting unsolved problems to work on
  - Compatibility with other Grid systems comes for free

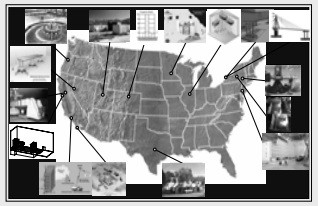
The Application-Infrastructure Gap

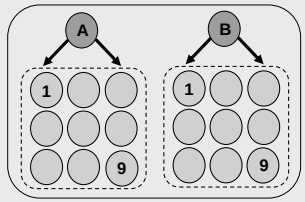


Dynamic and/or Distributed Applications



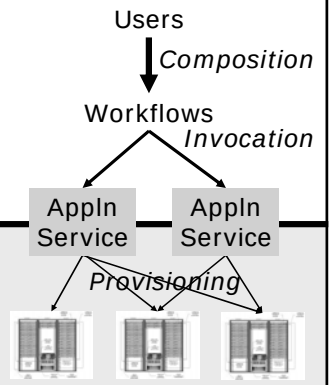
Shared Distributed Infrastructure





Bridging the Gap: Grid Infrastructure

- Service-oriented **applications**
  - Wrap applications as services
  - Compose applications into workflows
- Service-oriented **Grid infrastructure**
  - Provision physical resources to support application workloads



The Globus Commitment to Open Source

- To allow for inspection
  - for consideration in standardization processes
- To encourage adoption
  - in pursuit of ubiquity and interoperability
- To encourage contributions
  - harness the expertise of the community

http://dev.globus.org

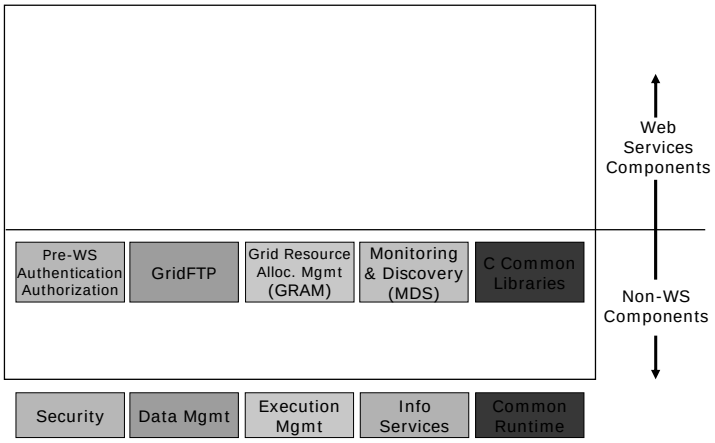
Globus Toolkit by Domain Areas

- Core runtime
  - Infrastructure for building new services
- Security
  - Apply uniform policy across distinct systems
- Execution management
  - Provision, deploy, & manage services
- Data management
  - Discover, transfer, & access large data
- Information services
  - Discover & monitor dynamic services

Globus Toolkit by Protocol Type

- Web service protocols
  - WSDL, SOAP
  - WS Addressing, WSRF, WSN
  - WS Security, SAML, XACML
  - WS-Interoperability profile
- Non Web service protocols
  - Standards-based, such as GridFTP
  - Custom

Globus Toolkit version 2 (**GT2**)



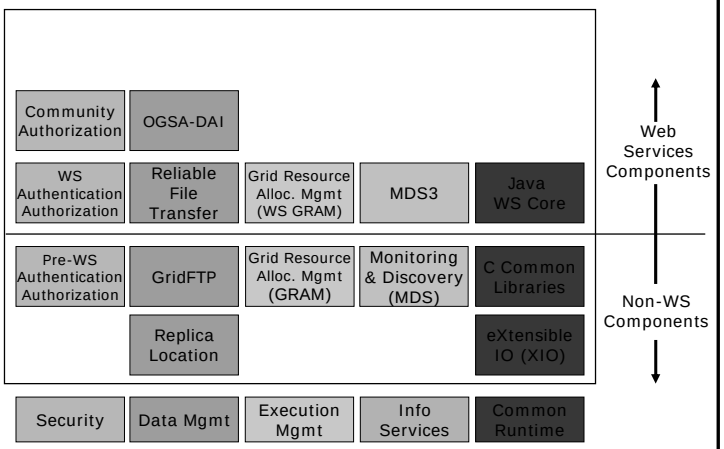
GT2 Key Protocols

- The Globus Toolkit v2 (GT2) centers around four key protocols
  - Connectivity layer:
    - *Security*: Grid Security Infrastructure (GSI)
  - Resource layer:
    - *Resource Management*: Grid Resource Allocation Management (GRAM)
    - *Information Services*: Grid Resource Information Protocol (GRIP)
    - *Data Transfer*: Grid File Transfer Protocol (GridFTP)
- Also key collective layer protocols
  - Info Services, Replica Management, etc.

OGSA and the Globus Toolkit

- Technically, OGSA enables
  - Refactoring of protocols (GRAM, MDS, GridFTP), *while preserving all GT concepts/features*!
  - Integration with hosting environments: simplifying components, distribution, etc.
  - Greatly expanded standard service set
- Pragmatically, Globus proceeded as follows
  - Develop open source OGSA implementation
    - Globus Toolkit 3.0; supports Globus Toolkit 2.0 APIs
  - Partnerships for service development
  - Also expect commercial value-adds

Globus Toolkit version 3 (**GT3**)



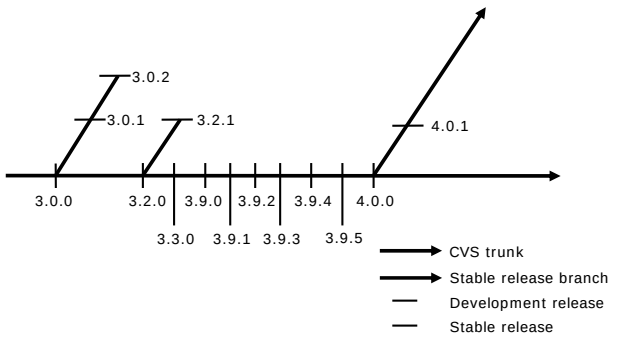
GT2 Evolution To GT3

- What happened to the GT2 key protocols?
  - Security: Adapted X.509 proxy certs to integrate with emerging WS standards
  - GRIP/LDAP: Abstractions integrated into OGSI as serviceData
  - GRAM: ManagedJobFactory and related service definitions
  - GridFTP: Unchanged in 3.0, but will evolve into OGSI-compliant service in 2004
- Also rendering collective services in terms of OGSI: RFT, RLS, etc.

GT Timeline

- GT 1.0: *1998*
  - GRAM, MDS
- GT 2.0: *2001*
  - GridFTP, packaging, reliability
- GT3 Technology Preview: *Apr-Dec 2002*
  - Tracking OGSI definition
- GT3.0 Alpha: *Jan 2003*
  - OGSI Base, GT2 functionality
- GT3.0 Production: *June 2003*
  - Tested, documented, etc.
- GT3.2.1 Production: *July 2004*
- GT3.9.2 Development: Aug 2004
- GT4 Scheduled beta development Dec 2004
- GT 3.9.5 (the GT4 beta release) Feb 2005
- GT 4.0 Mar 2005

4.0 is not a typical ".0" release, but the culmination of months of testing



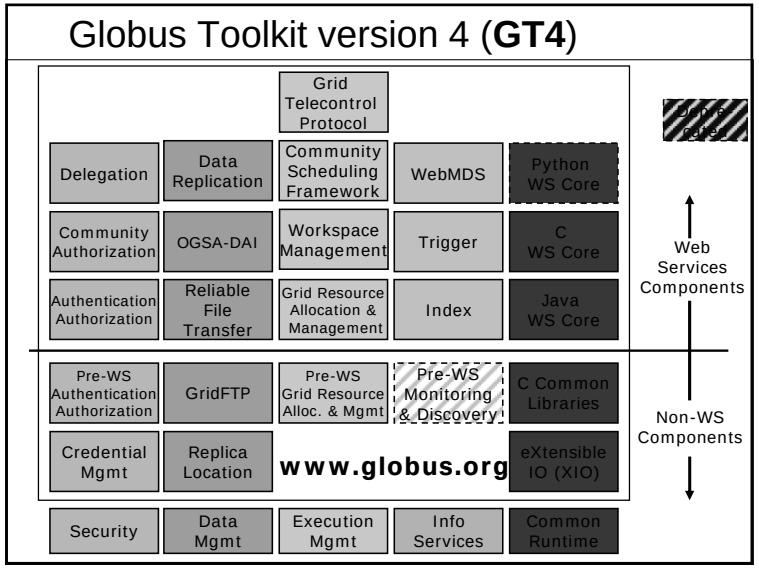
GT Timeline

- GT 4.0 Mar 2005
- GT 4.0.1 Aug 2005
  - Adoption of Version 2 of the Apache License, without modification
  - Improved support for Apache Tomcat
  - Improved support for message-level security in C WS Core
  - An updated version of OGSA-DAI, containing the complete set of public interfaces
  - The addition of a new tech preview component, an early release of the SweGrid Accounting System (SGAS)
- **Nov 2005**
  - Rocks 4.1 Cluster Software Released, with Support for GT 4.0.1
  - Weka4WS: GT4-based distributed data mining

GT Timeline

- April 2006 **GT 4.0.2**
  - Support for Tomcat 5.5.x
  - Improved WSDL parsing and serialization of complex types in C WS Core
  - Improved performance and a new programmer's tutorial for Python WS Core
  - Beta-quality integration between the SweGrid Accounting System (SGAS) tech preview and WS GRAM
  - Improved ODBC support in the Replica Location Service (RLS)
  - Improved resource management in OGSA-DAI WSRF
  - Scalability improvements to the Data Replication Service (DRS)
  - Job throughput improvements in WS GRAM

GT Timeline

- Apr 2006 WSRF 1.2 approved as OASIS standard
- July 2006 New stable release of the GridWay Meta-scheduler
- Aug 2006 GT 4.0.3
  - Updated version of the CoG JGlobus library
  - Updated version of the Apache Axis library
  - Improved credential renewal support in MyProxy
  - Improved error reporting for WS GRAM multijobs

Globus Toolkit version 4 (**GT4**)