What is the Globus Toolkit?

by Gregg Zepp

What is Globus? "The Globus Alliance is an international association dedicated to developing fundamental technologies needed to build grid computing infrastructures." The Globus Alliance was officially established in September 2003, however it was created out of the previous Globus Project that had been established in 1995. [1] From the home page of their website, it states, "The Globus Alliance is a community of organizations and individuals developing fundamental technologies behind the "Grid," which lets people share computing power, databases, instruments, and other on-line tools securely across corporate, institutional, and geographic boundaries without sacrificing local autonomy." [2]

> What distinguishes grid computing from conventional high performance computing systems such as cluster computing is that grids tend to be more loosely coupled, heterogeneous, and geographically dispersed. Although a grid can be dedicated to a specialized application, it is more common that a single grid will be used for a variety of different purposes. Grids are often constructed with the aid of general-purpose grid software libraries known as middleware. [3]

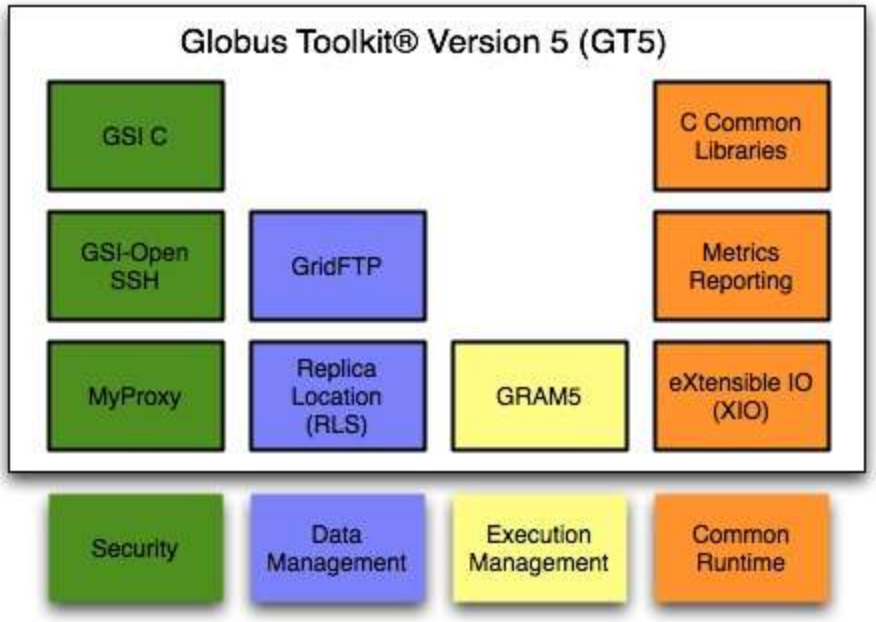
Begun in 1996, the Globus Project was initially based at Argonne, ISI, and the University of Chicago (U of C). What is now called the Globus Alliance has expanded to include the University of Edinburgh, the Royal Institute of Technology in Sweden, the National Center for Supercomputing Applications, and Univa Corporation. Project participants conduct fundamental research and development related to the Grid. Sponsors include federal agencies such as DOE, NSF, DARPA, and NASA, along with commercial partners such as IBM and Microsoft. [4]

What is the Globus Toolkit? The Globus website it states, "The Globus Toolkit is an open source software toolkit used for building Grid systems (aka, Grids) and applications. It is being developed by the Globus Alliance and many others all over the world. A growing number of projects and companies are using the Globus Toolkit to unlock the potential of grids for their cause." [5] It's main page is: http://www.globus.org/toolkit/ It's current release is v5.0.2. [6] Its release manuals are located here: http://www.globus.org/toolkit/docs/5.0/5.0.2/

---

[1] Wikipedia: http://en.wikipedia.org/wiki/Globus_Alliance
[2] Globus: http://www.globus.org/
[3] Wikpedia: http://en.wikipedia.org/wiki/Grid_computing
[4] Globus Toolkit: http://www.globus.org/toolkit/about.html
[5] Globus: http://www.globus.org/
[6] Globus Downloads: http://www.globus.org/toolkit/downloads/

© 2010. Gregg Zepp



Its Beginnings In late 1994, Rick Stevens, director of the mathematics and computer science division at Argonne National Laboratory, and Tom DeFanti, director of the Electronic Visualization Laboratory at the University of Illinois at Chicago, proposed establishing temporary links among 11 high-speed research networks to create a national grid (the "I-WAY") for two weeks before and during the Supercomputing '95 conference. A small team led by Ian Foster at Argonne created new protocols that allowed I-WAY users to run applications on computers across the country. This successful experiment led to funding from the Defense Advanced Research Projects Agency (DARPA, and 1997 saw the first version of the **Globus Toolkit**, which was soon deployed across 80 sites worldwide. The U.S. Department of Energy (DOE) pioneered the application of grids to science research, the National Science Foundation (NSF) funded creation of the National Technology Grid to connect university scientists with high-end computers, and NASA started similar work on its Information Power Grid. [7]

The project has spurred a revolution in the way science is conducted. High-energy physicists designing the Large Hadron Collider at CERN are developing Globus-based technologies through the European Data Grid, and the U.S. efforts like the Grid Physics Network (GriPhyN) and Particle Physics Data Grid. Other large-scale, science projects relying on the Globus Toolkit include:

- The Network for Earthquake Engineering and Simulation (NEES)
- FusionGrid
- The Earth System Grid (ESG)

[7] Globus Toolkit: http://www.globus.org/toolkit/about.html

- The NSF Middleware Initiative (and its GRIDS Center)
- The National Virtual Observatory

In addition, many universities have deployed campus Grids, and deployments in industry are growing rapidly. Since 2000, companies like Avaki, DataSynapse, Entropia, Fujitsu, Hewlett-Packard, IBM, NEC, Oracle, Platform, Sun and United Devices have pursued Grid strategies based on the Globus Toolkit. This widespread industry adoption has brought a new set of objectives, with the cardinal purpose being to preserve the open-source, non-profit community in which the Globus Project has thrived, while seeding commercial grids based on open standards. 2004 saw the formation of Univa Corporation, a company devoted to providing commercial support for Globus software, and 2005 the creation of the Globus Consortium by a group of companies with an interest in supporting Globus Toolkit enhancements for enterprise use.

How to get started using it? The release manuals are the best source of information. There's are installation and quickstart guides and guides for end users and developers. This page is the central location for the various versions' documentation: http://www.globus.org/toolkit/docs/ Global Toolkit 5.0.2 release manuals are divided into the following 4 areas [8]:

- Data Management Components

  Data management tools are concerned with the location, transfer, and management of distributed data. GT5 provides various basic tools, including GridFTP for high-performance and reliable data transport, and RLS for maintaining location information for replicated files.

- Execution (Jobs) Management Components

  Execution management refers to the initiation, monitoring, management, scheduling, and/or coordination of remote computations. GT5 supports the Grid Resource Allocation and Management (GRAM5) interface as a basic mechanism for these purposes.

- Security Components

  These components establish the identity of users or services (authentication), protect communications, and determine who is allowed to perform what actions (authorization), as well as manage user credentials. GT5 provides distinct WS and non-WS authentication and authorization capabilities. Both build on the same base, namely the standard X.509 end-entity and proxy certificates, which are used to identify

[8] Globus Toolkit 5.0.2 Release Manuals: http://www.globus.org/toolkit/docs/5.0/5.0.2/

persistent entities such as users and servers and to support the temporary delegation of privileges to other entities.

- Common Runtime Components

  These components include a set of C Common libraries needed for building grid infrastructure and XIO, a sophisticated and extensible I/O library suitable for the dynamic needs of grid applications.