NOVEMBER/DECEMBER 2006
WWW.EOSJ.COM

ENTERPRISE Open Source JOURNAL

Focusing on Open Source Strategies in the Enterprise

Integrating Business Rules & Processes for the Mass Market

Do-It-Yourself Linux?

10 Considerations for Implementing SOAs

Global Grid Computing

Open Source Project and Portfolio Management Applications

Ready for Prime Time?

A tci PUBLICATION

Global Grid Computing
Where Are We Today?

By Krishna Nadiminti and Rajkumar Buyya

Grids—large-scale distributed, heterogeneous, resource-sharing collaborative environments—are seeing increased adoption beyond their traditional territory of academic and scientific communities.

Grids can offer various services such as:

- **Compute services:** CPU cycles by pooling computational power
- **Data services:** collaborative sharing of data generated by people, processes, and devices such as sensors and scientific instruments
- **Application services:** access to remote software services/libraries and license management
- **Interaction services:** e-learning, virtual tables, group communication, and gaming
- **Knowledge services:** data mining and knowledge acquisition, processing, and management.

The business and commercial sectors increasingly are beginning to use different aspects of Grids to achieve their objectives. In our previous *EOSJ* article, “Enterprise Grid Computing: State-of-the-Art,” (March/April 2006), we looked at the state of enterprise Grids and the benefits and challenges they pose for a business looking to leverage Grids to improve the efficiency of their IT processes. In this article, we go a step further and look at global Grids, the challenges involved, the current status, and barriers to adoption. We also look briefly at an open source Grid toolkit, called the Gridbus middleware toolkit, and the future directions for global Grids.

Global Grids

A global Grid is a worldwide (or large-scale and geographically distributed) computing environment, offering a diverse range of heterogeneous services managed and owned by self-interested parties. Global Grids offer various kinds of services, including those listed previously, and support distributed supercomputing, high-throughput computing, content-sharing, data-intensive computing, on-demand or real-time computing, collaborative computing, and the like. Many of these services are provided and managed by self-interested parties and no single organization has control over the entire Grid. Figure 1 shows a variety of applications global Grids can support; Figure 2 shows an example scenario of global Grid computing. This is a typical distributed, high-throughput computing environment where users have large-scale or specialized computing and data analysis requirements and access to shared resources located at various sites around the world, each owned by different institutions.

There’s considerable complexity involved in getting everything working from both the users’ perspective and resource-owners’ perspective. On the resource-owner side, the concept of “Virtual Organizations” (VOs) is an increasingly popular approach to overcoming the administrative challenges. People using resources, applications, data and processes from different institutions form a VO to achieve a common goal. Using a Grid resource broker can simplify things on the users’ side. The broker:

- Mediates access to distributed resources by discovering suitable computation and data resources for a given analysis scenario
- Optimally maps analysis application tasks to resources
- Deploys and monitors task execution on selected resources
- Accesses data from local or remote data sources during task execution
- Collates and presents results to the user.

Global Grid Challenges and Status

Grid environments have various sources of complexity, including:

- Size (large number of nodes, providers, consumers)
- Heterogeneity of resources (PCs, workstations, clusters, and supercomputers)
- Heterogeneity of fabric management systems (single system image operating system, queuing systems, etc.)
- Heterogeneity of fabric management polices
- Heterogeneity of applications (scientific, engineering, and commerce)

COMPUTATION-INTENSIVE

- Interactive simulation (climate modeling)
- Large-scale simulation and analysis (galaxy formation, gravity waves, event simulation)
- Engineering (parameter studies, linked models).

DATA-INTENSIVE

- Experimental data analysis (e.g., physics)
- Image and sensor analysis (astronomy, climate).

DISTRIBUTED COLLABORATION

- Online instrumentation (microscopes, X-Ray)
- Remote visualization (climate studies, biology)
- Engineering (large-scale structural testing).

Figure 1: Examples of Grid Applications

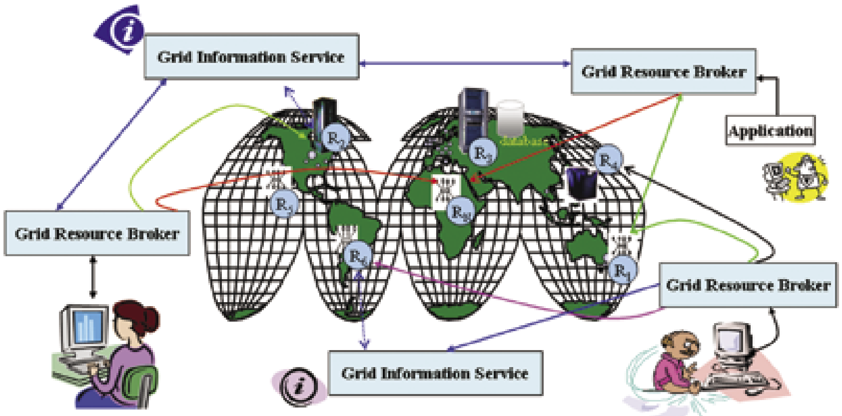

Figure 2: Typical Global Grid Scenario

- Heterogeneity of application requirements (CPU, I/O, memory, and/or network-intensive)
- Heterogeneity in demand patterns
- Geographic distribution and different time zones
- Differing goals (producers and consumers have different objectives and strategies)
- Insecure and unreliable environment.

When operating at a global scale, though, most of these issues are magnified manifold (see R. Buyya et al., "An Architecture for a Resource Management and Scheduling System in a Global Computational Grid," The 4th International Conference on High Performance Computing in Asia-Pacific Region, Beijing, China. IEEE Computer Society Press, USA, 2000). Traditional centralized policies for distributed computing aren't scalable and aren't practically applicable to global Grids, since ownership, control and management of resources are inherently distributed, and complete state information isn't available to any single entity in the Grid. Also, it becomes difficult to define standard global performance metrics and a common management policy for heterogeneous resources.

Creating a global Grid that tackles these challenges requires separating the various concerns and delegating the management of each aspect to different entities that work in unison. We could use a generalized layered architecture (see Figure 3). At the lowest level, we have an integration of individual software and hardware components into a combined networked source (such as a single system image cluster). Low-level middleware is built on this Grid fabric to provide a secure, uniform access to services provided by different resources. User-level middleware uses the services of the low-level middleware to support application development and aggregation of distributed resources. The construction of distributed applications then becomes easier; the underlying complexity is largely hidden. In such a scenario, the various components are developed, deployed and managed by different participating entities and organizations in the way they seem fit. The interfaces through which the components communicate or interact are standardized, achieving uniformity and interoperability.

Various approaches address the challenges of creating global-scale distributed systems. Many open source middleware projects have emerged as Grid computing gained momentum. These include:

- Object-oriented systems (e.g., Legion, http://legion.virginia.edu)
- General purpose Grid middleware (e.g., Globus, www.globus.org)
- Internet-based, peer-to-peer-style systems (e.g., SETI@Home, http://setiathome.berkeley.edu)
- Network-enabled solvers (e.g., NetSolve, http://icl.cs.utk.edu/netsolve)
- Service-oriented computing (e.g., Gridbus, www.gridbus.org).

Figure 4 shows which parts of the layered architecture some of these projects implement.

Let's look at the Grid tools available from two open source Grid projects—Globus toolkit and Gridbus.

The Globus Toolkit

The Globus toolkit is a collection of ser-

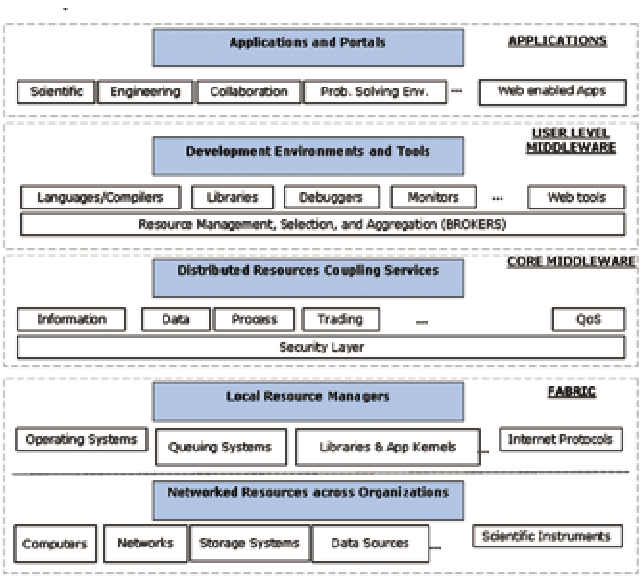

Figure 3: Layered Grid Architecture

vices and tools addressing key technical problems in building Grids.

The Globus Alliance includes the Argonne National Laboratory, Chicago; USC Information Sciences Institute; Univa Corp.; University of Edinburgh; the National Centre for Supercomputing applications, USA; and others.

The toolkit includes general infrastructure tools (i.e., middleware) that can be applied to many application domains to solve issues related to security, information infrastructure, resource and data management, communication, fault detection, and portability.

Version 4 of the Globus Toolkit (GT4), released in April 2005, represents a significant advance relative to the previous implementations of the toolkit services and functionality in terms of the range of components provided, functionality, standards conformance, usability, and quality of documentation. GT4 includes an implementation of the Web Services Resource Framework (WSRF) specification for stateful Web services. The main components of GT4 are:

- Common run-time components and hosting environment for WSRF services
- Security, including WS authentication and authorization, Community Authorization Service (CAS), delegation service, Grid Security Infrastructure (GSI), and proxy management
- Data management, including Reliable File Transfer (RFT) service, GridFTP, and Replica Location Service (RLS)
- Information services, including Monitoring and Discovery Services (WS-MDS)
- Execution management, including Remote Execution Management (WS-GRAM).

The toolkit is designed to support applications in which sets of services interact via standard protocols. The software includes both complete services and libraries implementing useful protocols. Developers can use these services and libraries, plus other related software, to relatively quickly build both simple and complex systems.

The Gridbus Middleware

Gridbus, the flagship project from the GRIDS Lab, at the University of Melbourne, Australia, is an open source R&D project focused on Grid economy, utility Grids, and service-oriented computing. It's engaged in the design and development of Grid middleware technologies to support e-science and e-business applications. These include:

- **Grid service broker** is a competitive, user-centric, economy-capable Grid scheduler for compute- and data-intensive applications that mediates access to distributed resources by (a) discovering suitable data sources for a given analysis scenario, (b) suitable computational resources, (c) optimally mapping analysis jobs to resources, (d) deploying and monitoring job execution on selected resources, (e) accessing data from local or remote data source during job execution, and (f) collating and presenting results. The broker supports a declarative and dynamic parametric programming model for creating Grid applications.
- **Alchemi** is a Microsoft .NET-based enterprise Grid computing framework for the Windows operating system that lets you painlessly aggregate the computing power of networked machines into a virtual supercomputer (desktop Grid) and to develop applications to run on the Grid.
- **Grid Market Directory (GMD)** provides directory and information services for utility Grids and enables Grid service providers to publish their services and related costs to the public, so consumers can browse the information, using a Web browser, or query the directory programmatically, using Web services and SOAP to find a suitable service meeting their requirements.
- **Grid Bank** is an accounting and resource usage management system with support for computational economy that can lead to self-regulated accountability in Grid computing.
- **Workflow management engine** is an economy-capable adaptive Grid application workflow scheduling and monitoring system that facilitates users to link standalone applications and execute their workflow applications on Grids. It provides an XML-based workflow language for users to define tasks and dependencies, and supports a Just-In-Time (JIT) scheduling system, allowing the resource allocation decision to be made at the time of task execution and hence adapt to changing Grid environments.
- **GridSim toolkit** is an advanced toolkit for modeling and simulation of entities in global Grids: users, applications, resources, and resource brokers (schedulers) for designing and evaluating scheduling algorithms. It provides a wide set of features for creating different classes of heterogeneous resources that can be aggregated using resource brokers for solving compute- and data-intensive applications.
- **Libra** is an economy-based resource management and scheduling system for clusters. It supports resource allocation strategies reflecting Service Level Agreements (SLAs) that let resource owners manage competing demands from different users and also lets resource consumers specify their Quality of Service (QoS) requirement parameters.
- **Gridscape** is a tool for the creation of interactive and dynamic Grid test bed Web portals for resource and Grid application monitoring based on Google Maps.

Used in combination, these sets of tools support the integration of both Windows and Unix-class resources for enterprise and global Grid computing and cover most aspects of the layered architecture previously described.

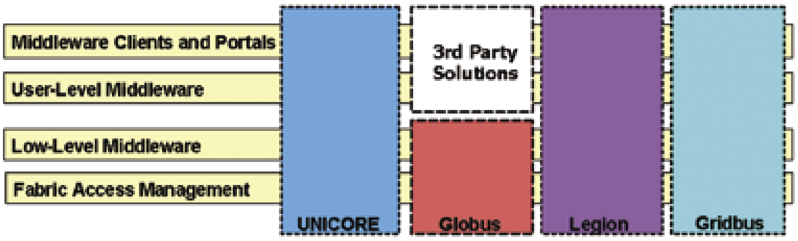

Figure 4: Projects Implementing Part of the Grid

Grid Adoption

The Globus project is one of the most widely used Grid-middleware systems and has become a defacto standard in the area of Grids. The Open Grid Forum (OGF, www.ogf.org/) is a standards body that aims to accelerate Grid adoption by providing an open forum for Grid innovation and developing open standards for Grid software interoperability. They have various working groups in areas such as resource allocation, distributed resource and data management, applications, infrastructure, information services, security and more. The OGF enjoys strong support from the industry, including all major companies such as Microsoft, IBM, HP, Sun, Oracle, and others. In spite of all these efforts, the current industry adoption is limited to intra-enterprise Grids. Large, national scale Grids are in use today, but mostly by scientific research communities in various countries. Some examples include:

- Australia: www.grid.apac.edu.au/
- China: www.chinagrid.edu.cn/chinagrid/index.jsp
- Europe: http://eu-datagrid.web.cern.ch/eu-datagrid/
- India: www.garudaindia.in/
- Japan: www.naregi.org/index_e.html
- Korea: www.gridcenter.or.kr/
- Singapore: www.ngp.org.sg/
- U.S.: www.teragrid.org/.

Various Grid applications have been run on these and other community Grid infrastructures set up for research purposes. For instance:

- The U.K. e-science program facilitates large-scale scientific research projects such as AstroGrid (enabling a virtual observatory), BioSimGrid (bimolecular simulations), Comb-e-chem (exploiting combinatorial chemistry), GridPP (Grid for particle physics), and many others.
- The EU Datagrid program supports applications in high-energy physics, environmental science, and bioinformatics.
- The Earth System Grid in the U.S. is focused on delivery and analysis of large climate model data sets for the climate research community.
- The National Aeronautics and Space Administration (NASA) Information Power Grid (IPG) uses a production Grid for aero-sciences and other NASA missions.
- The Japanese BioGrid project focuses on life sciences and biological simulations.
- The APACGrid program in Australia provides common base Grid infrastructure for a range of projects, including KidneyGrid, Australian BioGrid, Australian Virtual Observatory, Belle Data Grid for High Energy Physics, Oceans and Climate digital library, and more.

Many commercial companies are setting up their own Grids and there are several software vendors offering Grid solutions. Most of these efforts are currently focused on improving the use of enterprisewide computing and data resources. However, many of the current Grid applications are driven by the line-of-business for enterprises in the financial, pharmaceutical, and engineering sectors.

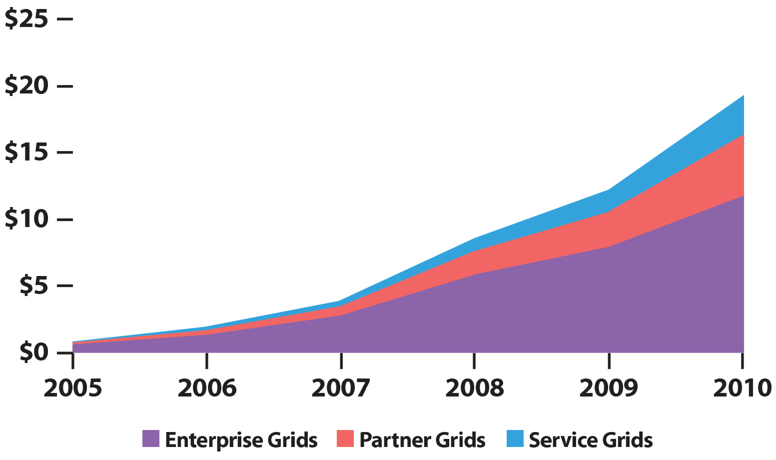

Figure 5: Grid Adoption in Industry (Source: Insight Research Group)

The true potential of Grids is far from being realized and adoption is still at a nascent stage. Worldwide investments in Grids are expected to grow rapidly beyond 2007 (see Figure 5). Security, trust, license management, administration and management policies, pricing models, usage and metering, QoS guarantees, and resource allocation are issues still pending satisfactory resolution even as Grid standards continue to evolve.

The Utility Grid

The term "Grid" is analogous to the electrical power Grid, which is the basic infrastructure that enables electricity to be distributed and made available as a utility to anyone who needs it. The long-term vision for Grid computing is to make computing available as a true utility, with consumers not having to worry about generation, distribution, or management of computing resources. Compute power, data, information, and knowledge would all be available on demand in a ubiquitous, reliable manner.

Just as the other utilities we have today—water, electricity, telephone, gas, and the Internet—computing will be delivered as a commodity with pay-per-use pricing. It will take time to realize this vision, but advances in networking technologies, growth in the power of the desktop PC, improvements in storage management and retrieval systems, and maturity of Grid and resource virtualization software will eventually propel our society toward the utility Grid. ●

Krishna Nadiminti is a research programmer at the Grid Computing and Distributed Systems (GRIDS) Laboratory in the Department of Computer Science and Software Engineering at the University of Melbourne, Australia. As a member of the Gridbus Project, he contributed toward the development of various Grid technologies, including the Grid Service Broker and Alchemi Enterprise Grid platform.
Email: kna@csse.unimelb.edu.au

Rajkumar Buyya is director of the GRIDS Lab. He also serves as chair of the IEEE Technical Committee for Scalable Computing. He has pioneered an economic paradigm for service-oriented Grid computing and demonstrated its utility through his contribution to the conception, design and development of cluster and Grid technologies, such as Gridbus, that power emerging e-science and e-business applications.
Email: raj@csse.unimelb.edu.au