LBNL DISTRIBUTED SYSTEMS DEPARTMENT

Supporting Distributed Science

The Distributed Systems Department at LBNL seeks to allow scientists to address complex and large-scale computing and data analysis problems beyond what is possible today. The department focuses both on bringing existing Grid and collaborative technologies to application scientists today and developing new technologies that will be essential parts of tomorrow's collaborative science.

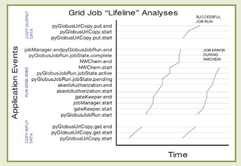
BRINGING THE POWER OF DISTRIBUTED COMPUTING TO APPLICATIONS

Higher level interfaces to Grid services

The widespread adoption of Grid middleware offers the promise of creating science applications by gluing together high-level domain-specific components. To make this vision a reality, there must be a suite of tools that will enable easy composition. In our **PyGlobus/CogKit** project we are currently developing Python tools that wrap 2.x, 3.x and WS-RF versions of the Globus Toolkit™. Python is ideally suited for use as a "glue" language, as it provides a nice combination of easy-to-read syntax, high-level constructs, support for interacting with C/C++ and Fortran, and is in common use in the scientific community.

As Grid middleware evolves from functions provided by libraries to Grid services, the user tools must also evolve. Our **pyGridware** project is building a Python framework for Grid services and providing tooling to automatically generate services and clients from WDSL specifications. In addition we are developing tools to help application scientists provide interfaces to their existing codes and command line applications.

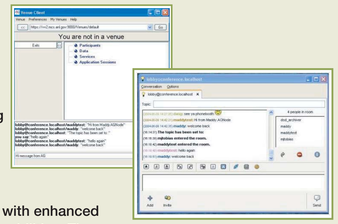
End-to-end monitoring



Grid applications are complex and often encounter failures and performance bottlenecks. The layers of middleware involved often hide the feedback needed to diagnose poor performance or failure. Our **NetLogger Toolkit** provides tools for end-to-end monitoring and troubleshooting of Grid applications. Accurate, detailed and adaptive monitoring of all distributed computing components, including the application, middleware, hardware, and network, is essential to this task. Our "end-to-end" monitoring provides the information to help track down the current status of a distributed job and locate any problems in the middleware or application.

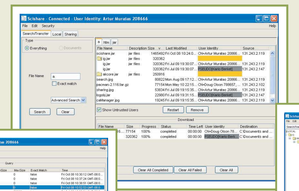
To be useful, monitoring information must be easy for application users, developers, and administrators to interpret. Invaluable for this purpose is a lifeline showing the correlation of a series of events against time. Our **NetLogger Visualization tool (NLV)** provides an interactive display of lifelines along with arbitrary additional end-to-end monitoring data. This monitoring data is invaluable for real-time analysis, anomaly identification, and fault recovery.

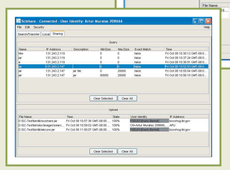
Enabling scientific collaboration

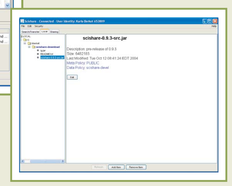


Collaboration between people involves an ongoing continuum of interaction through a variety of mediums. The **Pervasive Collaborative Computing Environment (PCCE)** is providing lightweight, non-intrusive, and flexible ways for collaborators to stay in touch and work together. PCCE presence and messaging is based on Jabber/XMPP protocols with enhanced security and personal archiving capabilities. These tools form the core of the environment and are being integrated into the next generation of the Access Grid so that they can provide a pervasive contact mechanism.

In a typical scientific collaboration, there are many different locations where data would naturally be stored. Our lightweight file-sharing system, **SciShare,** enables scientists to store and manage their files on local storage facilities and laptops while sharing them with remote participants. The underlying infrastructure that enables this is the reliable and secure group communication system composed of the **InterGroup** reliable multicast protocols and the **Secure Group Layer (SGL).** The end product is a scalable and secure peer-to-peer file-sharing system that leverages off X.509 identity certificate based authentication and authorization.

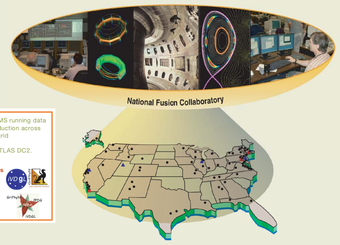


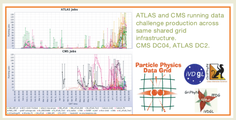




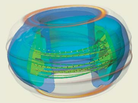
Our goal is to allow collaborators to easily begin a collaboration, locate each other, use asynchronous and synchronous messaging, share documents, share progress, share applications and hold videoconferences within a secure environment.

Customized tools and support for specific scientific collaborations



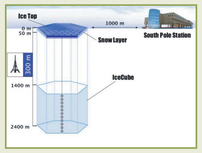


The final step of getting Grid technologies into the hands of domain scientists, requires the creators of these tools to join partnerships with the application scientists, middleware distribution and support organizations, and computing facilities to build a persistent infrastructure that uses and supports the software. At this point it is necessary to educate the scientists on how best to use the tools, fix the inevitable bugs that appear only with wider use, and help adapt the tools to the community environment. Members of our department have been working with two SciDAC pilot collaboratories: the **FusionGrid** and the **Particle Physics Data Grid.**



SCIRun visualization of NIMROD Pressure Simulation

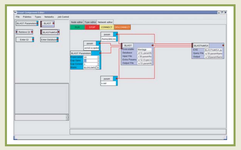
We are also leading the effort to develop the data acquisition software for the NSF-funded collaboration of scientists working on the **IceCube** neutrino detector, as well as the VDT middleware distribution and support team headquartered at the University of Wisconsin, and the Open Science Grid Consortium. These partnerships have given us valuable feedback on the difficulties of managing user credentials, executing distributed applications across firewalls, and maintaining and distributing software to a diverse user community.



NEXT GENERATION GRID TECHNOLOGIES

High-level visual programming interfaces for Grid application workflow

Currently most scientific workflows are built as one-off solutions to specific problems. Tools to make development of these workflows easier are needed. In addition, the increasing requirements of scientific projects to work with multiple offsite data stores, computational resources and collaborators, require workflow tools that can integrate standard grid services nodes into a single workflow. We are developing a visual programming environment for building and running domain science workflows with integrated data visualization and analysis. Using **ViCE,** existing legacy applications and newly developed web and Grid services can be composed by domain scientists into complex workflows. These workflows will provide high-level, standardized interfaces to low-level details, such as remote job submission, data transfer, and security. Our underlying workflow submission mechanism will be based on BPEL and use the Condor and Globus web services interfaces for job submission and tracking. ViCE will include collaboration features to allow scientists to rapidly design, debug, and execute workflows with collaborators in real time or asynchronously.



Next generation security protocols

Developing strong password-authentication technology for Grid computing involves designing a one-time password authenticated key exchange method and implementing it as a cipher suite for WS-SecureConversation. One-Time Password (OTP) authentication protects against captured passwords but not against attacks such as session hijacking — nor does it provide integrity and/or privacy of transmitted data. Secure password-based key-exchange on the other hand enables the establishment of a private and authenticated communication channel between a user's workstation and a remote server. It limits the damage due to the compromise of a server machine but does not protect against the compromise of users' workstations. In order to satisfy the security requirements of WS-RF compliant Grid services we need to combine OTP and password-based key-exchange to provide strong password authentication for Grid computing. To meet this need we are developing an implementation of the WS-SecureConversation based on OTP-authenticated key exchange **(OPKeyX).**

Incremental trust

Securing the range of collaborations from their spontaneous beginnings through multiple peers to an established virtual organization can not be accomplished by simply providing a traditional X.509 infrastructure; it requires a much richer set of capabilities. Mechanisms are needed to quickly set up secure communication based on minimal trust and to facilitate expanded sharing as trust builds throughout the life of the collaboration. We are working on an **incremental trust** authentication and authorization model that will let us explore this area. The authentication system will support multiple credentials for a user including username/password, X.509, and guest access. The authorization system will use the authentication method, trust values, and the user id to establish what permissions will be granted. In addition, the system will allow escorting of and vouching for users to temporarily elevate their privileges.

Peer-to-peer Resource Discovery Framework

Modern scientific research is conducted by large distributed multi-disciplinary teams. A difficult problem for these teams is often finding and organizing data, results, and resources. We propose to develop a flexible, powerful and extensible resource discovery and data integration framework called **FireFish.** This framework will provide a common resource discovery framework for a wide variety of scientific applications.

The work completed on this project will include the research, design, development, evaluation and benchmarking of a plug-in based reusable software infrastructure for Peer-to-Peer resource discovery, thereby enabling a range of innovative research directions building on it. Such an infrastructure is critical to achieve a functioning adaptive Grid and to enable geographically separated scientists to more effectively work together as a team.

TEID 9851