A Scalable and Secure Peer-to-Peer Information Sharing Tool

Deb A. Agarwal, LBNL
PI - Karlo Berket, LBNL
Olivier Chevassut, LBNL
Abdelilah Essiari, LBNL
Artur Muratas, LBNL

Summary

We are developing a peer-to-peer system to support secure, location independent information sharing in the scientific community. Once complete, this system will allow seamless and secure sharing of information between multiple collaborators. The owners of information will be able to control how the information is stored, managed, and shared. In addition, users will have faster access to information updates within a collaboration.

Groups collaborating on scientific experiments have a need to share information and data. This information and data is often represented in the form of files and database entries. In a typical scientific collaboration, there are many different locations where data would naturally be stored. This makes it difficult for collaborators to find and access the information they need. Our goal is to create a lightweight file-sharing system that makes it easy for collaborators to find and use the data they need. This system must be easy-to-use, easy-to-administer, and secure.

Current collaboration tools and environments provide a set of persistent services to users. However, they often rely on a centralized infrastructure. This makes the tools useless when a specific resource or server is unavailable. Ideally, the collaboration environment should not depend on any specific resource or server; instead, the presence of resources and servers should add value to the system. A collaboration environment should be structured to support informal, spontaneous collaborations as well as highly structured environments. Using on-line tools, it should be easy to begin collaborating, and incrementally add users and services as needed.

Our information-sharing tool is designed from the beginning for such an environment. It uses group communication, in particular the InterGroup protocols, to reliably deliver each query to all of the current participants in a scalable manner, without having to discover all of their identities. We will use the Secure Group Layer (SGL) and Akenti to provide security to the participants of our environment. SGL will provide confidentiality, integrity, authenticity, and authorization enforcement for the InterGroup protocols and Akenti will provide access control to other resources.

In our design we are building a file-sharing specific application specific layer and an underlying layer containing the implementation of a generic discovery model. We are currently developing implementations of both.

We have developed the Resource Discovery Messaging Framework (RDMF) as a representation of a generic discovery model. The RDMF provides a messaging infrastructure for resource discovery that is

easily configurable to a specific task. For example, the information-sharing application provides a configuration file to the RDMF that specifies how queries should be sent and how the responses should be sent back. In our initial prototype, we have specified that queries and responses should be sent using the HTTP protocol. Once the InterGroup protocols are added as a transport to the RDMF implementation, a configuration file change is all that is needed to allow the queries to be sent using InterGroup. In fact, an important advantage to using RDMF is the ease with which it allows systems using a centralized resource discovery scheme to transition to a distributed one.

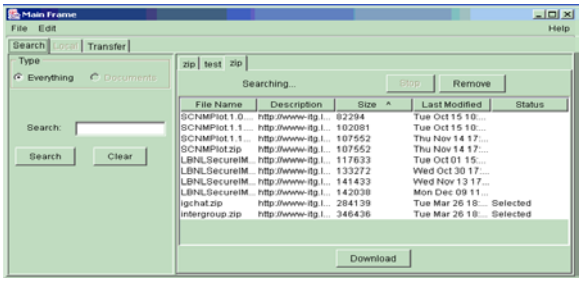

Figure 1. Performing a search in the scishare application.

The initial prototype implementation of the file-sharing application, dubbed *scishare*, is a graphical Java application. It provides the user the ability to search for remote files and transfer those files to the local machine (see Figures 1 and 2).

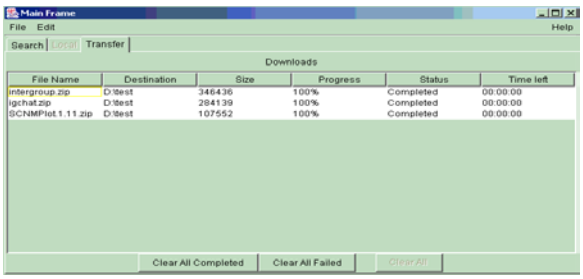

Figure 2. Transferring files in the scishare application.

The initial prototype uses the RDMF prototype to query and receive responses from the remote machine. The responses describe the files, as well as the service required to retrieve them. The application uses the information in the response to transfer the file to the local machine.

We are planning an initial release of the RDMF and *scishare* for June 2003. The released RDMF implementation will include the ability to use the InterGroup protocols as a transport for messages. The released *scishare* implementation will include the ability to share files with other users, and the ability to save and restore the state of the application.

In the next couple of years we will enhance the system by improving functionality and adding security mechanisms. We expect these enhancements to be driven by the needs of our users and collaborators.

We collaborate closely with the Reliable and Secure Group Communication and Distributed Security Architectures projects in realizing the group communication and security aspects of this project. We have recently started working with the Distributed Monitoring Framework (DMF) project. They plan to use RDMF to allow distributed discovery of DMF producers and the events they are generating. The Pervasive Collaborative Computing Environment project is planning to incorporate the *scishare* application into its environment. We have also had some initial discussions with the Middleware to Support Group to Group Communication project about incorporating the *scishare* application into the Access Grid environment.

For further information on this subject contact:
Karlo Berket, Scientist
Lawrence Berkeley National Laboratory
Phone: 510-486-4807
KBerket@lbl.gov