GA-A24805

DEVELOPMENTS IN REMOTE COLLABORATION AND COMPUTATION

by

J.R. BURRUSS, G. ABLA, S. FLANAGAN, K. KEAHEY,
T. LEGGETT, C. LUDESCHE, D. McCUNE, M.E. PAPKA,
Q. PENG, L. RANDERSON, and D.P. SCHISSEL

AUGUST 2004

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

GA-A24805

DEVELOPMENTS IN REMOTE COLLABORATION AND COMPUTATION

by

J.R. BURRUSS, G. ABLA, S. FLANAGAN, K. KEAHEY[£],
T. LEGGETT[£], C. LUDESCHE[§], D. McCUNE[§], M.E. PAPKA[£],
Q. PENG, L. RANDERSON[§], and D.P. SCHISSEL

This is a preprint of a paper to be presented at the 16th ANS Top.
Meeting on Technology of Fusion Energy, Madison, Wisconsin,
September 14-16, 2004, and to be printed in the *Proceedings*.

[£]Argonne National Laboratory, Argonne, Illinois
[§]Princeton Plasma Physics Laboratory, Princeton, New Jersey

Work supported by
the U.S. Department of Energy
under DE-FG02-01ER25455 and DE-FC02-04ER54698

GENERAL ATOMICS PROJECT 30106
AUGUST 2004

DEVELOPMENTS IN REMOTE COLLABORATION AND COMPUTATION

J.R. Burruss[1*], G. Abla[1], S. Flanagan[1], K. Keahey[2], T. Leggett[2], C. Ludesche[3],
D. McCune[3], M.E. Papka[2], Q. Peng[1], L. Randerson[3], D.P. Schissel[1]

[1]*General Atomics, PO Box 85608, San Diego, California 92186-5608*
[2]*Argonne National Laboratory, Argonne, Illinois 60439-4844*
[3]*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543*
**Corresponding author, email: burruss@fusion.gat.com*

The National Fusion Collaboratory (NFC) is creating and deploying collaborative software tools to unite magnetic fusion research in the United States. In particular, the NFC is developing and deploying a national FES "Grid" (FusionGrid) for secure sharing of computation, visualization, and data resources over the Internet. The goal of FusionGrid is to allow scientists at remote sites to participate as fully in experiments, machine design, and computational activities as if they were working on site thereby creating a unified virtual organization of the geographically dispersed U.S. fusion community.

I. INTRODUCTION TO FUSIONGRID

The National Fusion Collaboratory established the FusionGrid project to enable scientific collaboration for all aspects magnetic fusion energy research.[1] The vision of FusionGrid project is to make resources—data, software, hardware—available as network services, to be accessed transparently by fusion researchers nationwide.

II. COLLABORATIVE SOFTWARE

Collaborative software brings together geographically separated research teams. Videoconferencing, shared displays, and shared applications are examples of collaborative software. Less expensive computing resources have decreased the cost of remote collaboration to a point where smaller research facilities and even lone researchers can participate. A decade ago, researchers might collaborate remotely using SGI Indy workstations,[2] which sold for $10,500 each (inflation adjusted to 2003 dollars); today, a fast workstation with a camera and headset can be purchased for less than $3000.

II.A Access Grid (AG)

Access Grid (AG) software—developed open source by Argonne National Labs—provides for a rich communication environment that goes beyond teleconferencing and standard desktop-to-desktop videoconferencing to enable group-to-group interaction and collaboration. AG software combines multi-stream video and audio with shared applications to create an immersive collaboration between remote sites. Each AG station—referred to as a "node"—can be equipped with multiple video cameras for video transmission and multiple microphones for audio transmission. AG nodes meet in AG "venues", a metaphor used, much like the metaphor of the chat "room", to simplify the complicated software interactions that bring the separate nodes together. In the venue, participants can share files, applications, desktops, video; this sharing is made secure using GSI security. The application sharing of the AG venue, combined with audio and video, makes for a more natural videoconferencing experience.

Access Grid is very scalable. At the high end, an AG node can consist of an entire conference room, complete with sophisticated audio and echo cancellation capability, multiple video streams, and multi-screen displays. A more common AG node configuration is the mid-range Personal Interface to the Grid, or PIG. Each PIG typically has one or two video streams, a single audio input, and a two or three screen display, and is appropriate for small groups or a single user. At the low end of node configuration, AG software can be installed on a single laptop equipped with a commodity web camera and headset, but this configuration is only appropriate for a single user. Each of these AG node configurations is interoperable; thus, the lone researcher experimenting with AG software can join more established facilities in an AG meeting. Furthermore, since AG uses the Internet to send audio and video data, it is cheaper and more accessible that traditional ISDN-based videoconferencing. Bandwidth requirements vary based on the number of video streams, but the typical home cable modem is more than sufficient for receiving both audio and video.

In the fusion community, AG nodes are being applied to a number of situations. At DIII-D, AG nodes are used to broadcast the morning pre-operations meeting and the Friday science meeting. FusionGrid project meetings are held over Access Grid. AG nodes have been tested in the DIII-D control room, and will be deployed in the C-Mod control room. Outside of the control room, the most popular desktop platform for magnetic fusion scientists is Macintosh. However, there is as yet no AG client for

Macintosh; to reach the widest possible audience of fusion scientists, a Macintosh version of Access Grid must be developed. Workers at Argonne National Labs are currently developing Macintosh OS X-compatible AG software.

II.B. Case study: the deGrassie JET experiment and AG

In January 2004, fusion scientist John deGrassie remotely led the JET tokamak experiment located in England through an AG node at General Atomics in California. From San Diego, Dr. deGrassie was able to examine JET data, perform data analysis, and interact with researchers located in the JET control room.

This early experiment revealed an area of improvement in remote collaboration: researchers need to have access to familiar data analysis tools, either as part of Access Grid, or separately. When a researcher is saddled with unfamiliar tools, they fall behind in the rapid analysis process required for tokamak experiments. To make intelligent decisions and discuss data analysis with collaborating researchers, one needs to be able to rapidly analyze data. Access to familiar tools facilitates this. For this reason, the long-term vision of the AG venue is that users will have the ability to share any application.

III. Grid Computing

Grid computing has many benefits. In the traditional computing environment, software is installed and executed on the local machines of software users; this requires that the software be distributed to users, that the developers of the software create versions for each platform, and that the administrators of the various machines maintain the local copies of the software. Remote computing removes the need for multiple local software installations. Instead, using the application service provider (ASP) model, a provider enable many users to remotely access a software installation whose support and administration is carried out at a single site. This reduces the cost of computation.

Interacting in the Grid is complex: it involves bridging many security domains, interfacing with diverse local scheduling systems, and frequently requires fast communication. In order to deal with those challenges, FusionGrid uses the Globus Toolkit infrastructure which consists of a set of tools addressing these challenges.

Security is an important consideration for remote computing: a program executing on a remote site may require access to remote files or databases or may want to start other programs on remote resources. In order to avoid the necessity of presenting user credentials each time such need arises, a user needs to be able to delegate its credential to the executing program so that such requests could be automatically but securely executed on the user's behalf. In other words, the user needs the *single sign-on* capability allowing a user to "log into the Grid" once and then interact across many diverse site security systems. To provide this capability the Fusion Grid uses the Grid Security Infrastructure (GSI)[3] implemented by the Globus Toolkit (GT) which provides for reliable, secure identification of users and resources.

For example, logging into the grid enables a user to access data securely. For this purpose, a secure version of the MDSplus data acquisition and storage system has been developed to provide secure data storage and access. By integrating the fast and secure Globus XIO communication libraries, Secure MDSplus recognizes the proxy certificates used by GSI. In this way, researchers can efficiently and securely access scientific data in a common format through a single interface.

Another benefit of having obtained a credential is the ability to run computations on remote resources provided that such computation has been authorized by the resource owner. For this purpose, a scientist can use the GT Grid Resource Allocation Manager (GRAM) tool which allows a user to stage files for remote computation, start programs remotely and stream results of those executions.[4] GRAM allows a resource owner to configure authorization rules through a text file—called the gridmapfile—which not only authorizes a Grid user, but also maps a Grid entity to a specific local account in which the requested computation can be carried out. While this enables a very simple form of authorization, to build a flexible system we need more: a user may be allowed to execute only certain codes on a remote resource, those executions may have limited access to community data, and so forth. To answer this need, FusionGrid is developing an authorization system—the Resource-Oriented Authorization System (ROAM)—that will allow resource owners to formulate community-specific rules on access. GRAM can be configured to consult those community-specific rules through an authorization callout and act accordingly.

Once the computation has been started, researchers and administrators monitor the progress of FusionGrid computation through the FusionGrid Monitor (FGM). FusionGrid resources send monitoring information to this web-based monitor, which then pushes the monitoring information to client web browsers as the updates occur. In this way, monitoring is done via the familiar interface of a web browser. This has the added benefit of being very efficient, because each client web browser only refreshes when new monitoring information is available. The FGM also writes this information to a database, which is useful for statistics.[5]

In order to support experiments, a user needs the ability to reserve a resource or a set of resources for processing of high-priority codes running between the experimental cycles. In order to avoid wasting CPU cycles, the intent is that the execution of such high-priority codes would pre-empt all other executions taking place on a resource. For this reason, we are working with the Globus project to develop agreement-based services allowing for advanced reservation of resources and other time-critical facilities.[6]

In a typical scenario, before the experiment, a session leader would negotiate an advance reservation for the execution of certain codes on resources. Ideally, during the process of negotiation the session leader would receive estimates of the runtime of the involved programs in addition to information about the availability and suitability (in terms of computational power for example) of remote resources on which the programs should execute. A resource reservation made in this way is integrated with the between-pulse processing system so that, when the data becomes available, an event activates the reservation by associating it with a specific remotely run program. The session leader then has the option to either allow the program to run to completion within the required time or disassociate it from the reservation if the program does not converge for the purpose of assigning the agreement to another, more promising, computation.

III.A. Case study: TRANSP

The TRANSP transport analysis code[7] was made a FusionGrid service in 2002; the availability of TRANSP as a grid service has had a positive impact. Because TRANSP is centralized, users always have access to the most up-to-date version of the code. Each individual TRANSP run executes quickly, and many runs can be executed in parallel; as a result, researchers are more productive. Also, programmers are no longer need to maintain local copies of TRANSP at each site, saving several months of maintenance work per year.[8] TRANSP quality of support is much improved, because the service runs on machines at the institution (PPPL) where TRANSP experts reside, making trouble shooting much easier. Since TRANSP became a FusionGrid service in 2002, over 4,200 TRANSP runs have been completed. The FusionGrid TRANSP service is now the sole US TRANSP service, and is also actively used by EU researchers.

When researchers run TRANSP on FusionGrid, they first “sign on” to the grid, loads their inputs into an MDSplus database, then request a new TRANSP run. Assuming the user is authenticated and authorized, the TRANSP service then reads inputs from the MDSplus database, runs TRANSP, then writes outputs back to the MDSplus database when completed. During this process, TRANSP users monitor the progress of their TRANSP runs through the FusionGrid Monitor (FGM). Through FGM web links, run log file information can be examined. TRANSP users can even visualize TRANSP data while TRANSP is still executing. The ElVis visualization applet—developed at PPPL—is available through FGM links and is used to plot pre-selected data from a running TRANSP job, inside the user's java-enabled web browser.

IV. THE COLLABORATIVE CONTROL ROOM

The vision of the collaborative control room is to unite geographically dispersed researchers and resources in support of tokamak operations. To work effectively, a collaborative control room must enable researchers to easily share individual data analysis results with other individuals or the larger group. Remotely located researchers must be fully engaged in experimental operations though shared interactive applications and real-time video and audio. Researchers in the collaborative control room must be able to rapidly compare experimental data with simulation data. A collaborative control room must also provide for secure utilization of remote resources such as computing clusters or data storage systems.

The National Fusion Collaboratory is using the technologies of Access Grid and Grid Computing to realize the collaborative control room vision. A prototype collaborative control room was demonstrated at SuperComputing 2003 in Phoenix, Arizona. For the demonstration, offsite collaborators joined a mockup of a DIII-D experiment in San Diego, California. The offsite collaborators were able to communicate with the researchers onsite, as well as hear DIII-D announcements, and view DIII-D data. TRANSP—normally run on the timeframe of several hours—was executed instead in a few minutes as a between-shot code. The results of this demonstration and the feedback from fusion scientists are being used to fine-tune the requirements for a working collaborative control room.[6]

Fig. 1. The access grid node at SuperComputing 2003

Feedback from fusion scientists indicates that control room audio must be improved. The problem is that the control room has a lot of ambient noise. The off-site researchers pick up not just the one or two conversations they want to listen to, but side conversations and comments from off-camera researchers that happen to be picked up by microphones in the control room. From the point of view of offsite participants, this issue is easily resolved by equipping control room participants with audio headsets. However, feedback indicates that audio headsets make the control room researcher feel isolated from the events taking place in the control room, so some other solution is needed.

Feedback for shared display of data analysis results has been very positive. The approach taken by the collaboratory is to install tiled display walls. These tiled displays allow the researcher to share their individual data analysis results with the group. Rather than asking other researchers to walk over to their terminal to look at their plots, the collaborative researcher presses a button to display their plots on the tiled display wall. The collaboratory is continuing to improve upon this, looking into different ways to interact with shared images, and to share images with other individuals as opposed to the entire group.

Fig. 2. The tiled display wall in the DIII-D control room

V. FUTURE WORK & APPLICABILITY TO ITER

TRANSP will be adapted for use as a between-shot analysis code. Much of this work has been done (and demonstrated), but there still remain details—and interface questions—to be resolved.

Following the success of adapting TRANSP for use as a FusionGrid service, other codes will be made available on FusionGrid. Work has already been completed to make the GS2 microstability code a grid service, although more work needs to be done to make this test system a reliable production system. The GATO stability code, the GYRO microturbulence code, and the TORIC radio frequency code will also be made into FusionGrid services.

FusionGrid security will be enhanced though the use of a resource-oriented authorization management system. Documentation and a tutorial on understanding grid certificates will be created to better explain this important security concept to FusionGrid users.

To better support the collaborative control room concept, an agreement based resource management system will be created. This system will allow session leaders to reserve computing resources for time-critical computations prior to experimental operations. To ameliorate the problem of ambient noise in the collaborative control room, the NFC will investigate the use of Bluetooth headsets for providing individually-delivered audio. Because most researchers in the magnetic fusion community uses Mac OS X, the AG system will be ported to run on that platform. The collaboratory will also create a set of software tools called SharedD (for "Shared Display") to for more effective application sharing in the collaborative control room.

As more services are added to FusionGrid, the FusionGrid Monitor will be updated to better display monitoring information from multiple services. Further development of the ElVis collaborative visualization tool (currently used in the FusionGrid Monitor) will be pursued in order to resolve security and firewall-related performance issues.

Data management will be enhanced by extending secure MDSplus to leverage the new Globus XIO secure communication mechanism. The Globus XIO framework will allow for parallel data transfer, a useful performance enhancer for the high-bandwidth, high-latency connections between fusion institutions. These enhancements will be needed to overcome network limitations to improve data transfer performance without compromising security. If we use grid computing to generate simulation results remotely, then we will need good network data transfer performance in order to rapidly compare those simulation results with data from the experiment.

A FusionGrid software package will be created so that users can easily install the client software used by FusionGrid. The software package will stress ease-of-install, and will include documentation.

The collaborative technology being deployed is scalable to fusion research beyond the present programs, in particular to the design, construction, and operation of the KSTAR and ITER experiments that will require extensive collaboration capabilities worldwide.

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy SciDAC initiative and at General Atomics under DOE Contract No. DE-FG02-01ER25455 and Cooperative Agreement No. DE-FC02-04ER54698.

REFERENCES

[1] D. P. Schissel, et al., "Building the U.S. National, Fusion Grid: Results from the National Fusion Collaboratory Project", *Fusion Engineering and Design*, **71**, 245 (2004)

[2] T. A. Casper, et al., "Collaboratory Operations in Magnetic Fusion Scientific Research", *Interactions*, **3**, 56 (1998)

[3] R. Butler, et al., "Design and Deployment of a National-Scale Authentication Infrastructure*", IEEE Computer*, **33**, *12*, 60 (2000)

[4] K. Czajkowski, et al., "A Resource Management Architecture for Metacomputing Systems", *Proc. 4th Workshop on Job Scheduling Strategies for Parallel Processing*, Orlando, Florida, 1998, p. 62, Springer-Verlag (1998)

[5] S. M. Flanagan, et al., "A general purpose data analysis monitoring system with case studies from the National Fusion Grid and the DIII-D MDSplus between pulse analysis system", *Fusion Engineering and Design*, **71**, 263 (2004)

[6] K. Keahey, et al., "Grids for Experimental Science: the Virtual Control Room", *Challenges of Large Applications in Distributed Environments (CLADE)*, Honolulu, Hawaii, (2004)

[7] R.J. Hawryluk, "An Empirical Approach to Tokamak Transport", *Physics of Plasmas Close to Thermonuclear Conditions*, **1**, 19 (1980)

[8] J. R. Burruss, et al., "Remote Computing using the National Fusion Grid", *Fusion Engineering and Design*, **71**, 251 (2004)