MDSplus Remote Collaboration Support
- Internet and World Wide Web

T.W. Fredian, J.A. Stillerman

Plasma Science and Fusion Center
Massachusetts Institute of Technology
Cambridge, MA 02139

MDSplus Remote Collaboration Support - Internet and World Wide Web

T.W.Fredian and J.A.Stillerman
Massachusetts Institute of Technology
Cambridge , MA

The consolidation of fusion research into fewer major experiments has motivated the fusion community to study methods of improving the efficiency in which these experiments produce scientific results. One mechanism for increasing the scientific output of these facilities is through remote collaboration. To make the best use of existing facilities, scientists must be able to participate in the day to day operation of these facilities from anywhere in the world as easily and effectively as if they were present in the control room of the experiment. It is not sufficient to just make the scientific measurements available to more scientists. Remote collaborators must be able to provide input and communicate closely with all members of the experimental team. The rapid improvements in network technology are beginning to make this goal achievable. The increase in network bandwidth, along with platform-independent web based software technologies such as Java, are providing the infrastructure for truly networked fusion research. This paper will discuss some of the networked applications developed or employed at MIT's Plasma Science and Fusion Center to improve the research environment for remote collaborators who work as part of the Alcator C-Mod experimental team.

Introduction

In an attempt to improve the return on the investment in fusion research, the US fusion program is being consolidated into fewer major experiments. As a result, many labs and universities rich with scientific knowledge and expertise are left without local facilities to continue their studies. It is therefore becoming increasingly important that scientists be able to utilize the existing facilities for their studies regardless of the physical location of the experiment.

Fortunately, there has been an incredible growth of networking technology and infrastructure providing high speed information exchange over the Internet. This growth has been driven largely by the business sector, and there are no indications that this growth will slow in the near future. New software tools and faster networking hardware are being developed at a rapid pace.

Most of this technology may be directly applicable in helping to bring scientists closer to the experimental facilities with regard to information transfer, even though they may be physically separated by hundreds or even thousands of kilometers.

At the Plasma Science and Fusion Center at the Massachusetts Institute of Technology, tools have been developed to assist scientists to interact with the Alcator C-Mod fusion experiment. The information handling system used for the experiment is called MDSplus[1] and was developed jointly by MIT, Los Alamos National Laboratory, and Istituto Gas Ionizzati in Padova, Italy. MDSplus, designed nearly a decade ago, has networking capabilities compatible with today's technologies.

Most of the utilities provided for interacting with the experiment and analyzing data are based on the X windows graphical user interface standard[2]. Applications that use the X windows graphical user interface can be run on computers at MIT, and direct their displays to workstations anywhere on the Internet. More recently, additional applications have been implemented which utilize new technologies, such as Java, which make the applications available to remote and local users on a larger variety of computer platforms.

This paper will describe some of the tools developed as part of the MDSplus data system which enable scientists at remote locations to interact with the

Alcator C-Mod fusion experiment. It will also describe some of our experiences with remote collaboration and discuss future plans to enhance remote access capabilities.

History

In 1987, a joint project was initiated to design and implement a data acquisition system for use on the Alcator C-Mod experiment at MIT, the ZTH experiment at the Los Alamos National Laboratory, and the RFX experiment at Padova, Italy. The data system, subsequently called MDSplus, utilized the X windows graphical user interface for most of its applications. At the time when MDSplus was being designed, it was evident that the relatively new concept of desktop workstations should be considered for displaying information to the scientists. X windows was the graphical user interface of choice for these new workstations and thus became the choice for the MDSplus data system. Little did we know at the time how important remote access would become. The networking capabilities built into the X window system made it easy to run the applications from anywhere on the Internet. It was not long after the Alcator C-Mod experiment commenced operation in 1991 that it was discovered how fortuitous it was that the applications could be used by scientists located far from MIT. Given a fast enough network link, scientists from anywhere in the world could access experimental data in essentially the same manner as if they were sitting in the Alcator C-Mod control room.

In 1992, the first remote collaboration on Alcator C-Mod began. Scientists at the University of Maryland, nearly 750 kilometers from MIT, hoped to operate a spectroscopy diagnostic on Alcator C-Mod[3]. The MDSplus applications permitted the scientists at the University of Maryland to control the diagnostic and to view and analyze the data acquired immediately after each experiment cycle.

The ease with which the University of Maryland collaboration was accomplished lead scientists to wonder how far one could go with regard to remote access to the experiment. Would it even be possible to operate the experiment and major diagnostics from a remote control room even farther away than Maryland? The answer to these questions could be crucial to future world wide fusion collaborations such as the proposed ITER experiment. So in 1995, a group of scientists from MIT traveled to Lawrence Livermore National Laboratory (LLNL) in California and attempted to operate Alcator C-Mod in Massachusetts, nearly 5000 kilometers away[4,5]. The experiment made use of a remote control room at LLNL which was set up as a Remote Experimental Site (RES[6]) as part of a collaboration with General Atomics in San Diego. The remote control room contained several UNIX workstations and X terminals and video conferencing equipment. The attempt to operate the experiment remotely was highly successful. Several diagnostics, plasma shaping, and RF heating were controlled from LLNL. Operation of the experiment was slowed only slightly and real scientific results were produced.

We learned quite a lot from the remote control experiment regarding the overall experience of remote collaboration. Various performance issues were identified, and an better feel for the quantity and type of information needed to be transmitted to and from remote collaborators was obtained. This information was used to improve and add to our suite of remote access tools.

MDSplus X window based tools

Several X window based tools are available with MDSplus which allow a scientist to control diagnostics and to view raw and analyzed data. They include applications such as **traverser, dwscope,** and IDL™ [7].

The **traverser** is a tool with which the user can peruse the MDSplus database containing data acquisition and analysis setup information, acquired data and analyzed data. Using this tool, a scientist can make changes to settings such as diagnostic operating parameters and calibration coefficients. See figure 1.
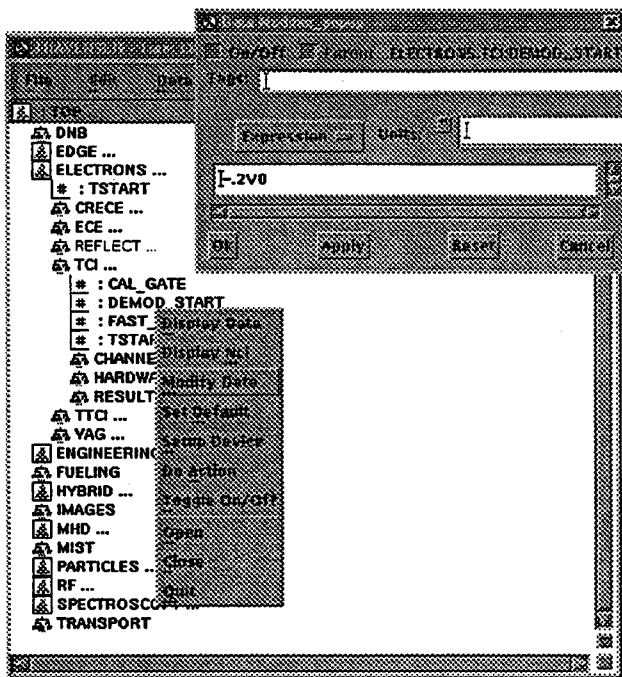
Figure 1 - Traverser

The **dwscope** application provides an efficient means of plotting X-Y traces. (See figure 2) It functions as an oscilloscope emulator, allowing the scientist to pan and zoom raw and processed data traces. This application is the preferred mechanism for viewing most of the acquired data immediately after the pulse of the tokamak. Scientists in the Alcator C-Mod control room often run 5-10 instances of this program displaying from one to sixty four traces in each.

IDL™, is an application which allows a scientist to manipulate and view arrays of data. Users can write sophisticated analysis applications using the high level data manipulation language provided in IDL™. A direct link to MDSplus was developed so IDL™ users can read and write MDSplus data.

These X window based tools are designed to be run on computers located at MIT with local access to the Alcator C-Mod experimental data. The graphical interface they present to the user can be directed to appear on any workstation, personal computer or X terminal that supports X windows and is connected to the Internet.

Remote data access library

A library of routines has been developed to enable programs running on remote computers to access MDSplus databases. Using TCPIP as the communication protocol, which runs over the Internet, the routines communicate with server processes running on computers in the Alcator C-Mod control room. These routines can be incorporated into users' analysis programs, enabling them to read and write C-Mod data. The library has been ported to many varieties of UNIX and to Microsoft Windows.
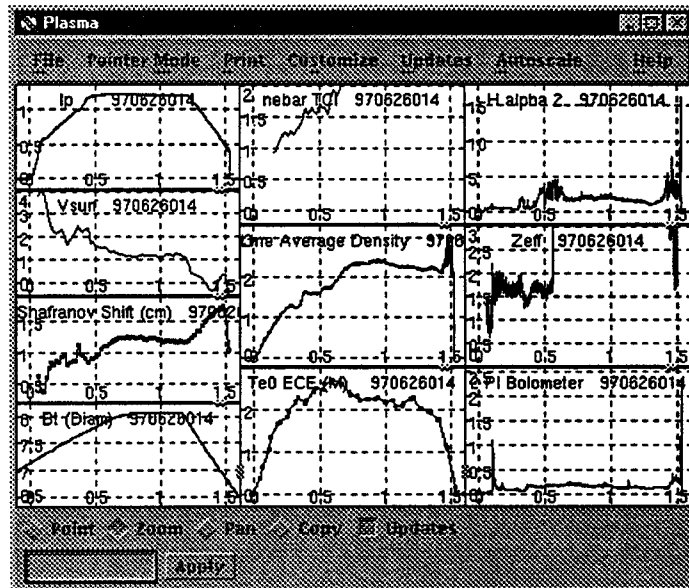

Figure 2 - Dwscope

UNIX and Windows Applications

The dwscope program, originally developed for the OpenVMS™ [8] operating system, has been ported to UNIX. The UNIX version of the program uses the remote data access library to access the data for plotting. Once the data is obtained using the remote data access routines, the user can interact with the data (i.e. pan and zoom) without any additional communication on the network.

Also an IDL™ interface has been developed which layers on top of the remote data access routines. Users on UNIX or Microsoft Windows can use IDL™ to access Alcator C-Mod data.

Web pages and News Groups

One of the things learned during the remote control room experiment was that the remote scientists did not get the feeling of being part of the C-Mod experimental group. Much more information than just data plots is exchanged among collaborators on-site at MIT. Updates to meeting schedules, run plans, and other calendar items must be accessible to remote collaborators as well. Other information such as the current status of the data acquisition hardware is also of interest. Many daily reports are generated that contain information useful to collaborators. It is also useful to have live pictures from the control room to see who is attending the run. All of this information needs to be accessible from remote sites.

To disseminate this information, we've chosen to use world wide web technology and usenet news groups. Web browsers and news readers are readily available on essentially all types of computers and operating systems. Most users are familiar with the operation of these viewers and can readily find the information desired. We have greatly expanded our web site with both static and live updating pages. These pages can be viewed at:

http://www.pfc.mit.edu/cmod/.

Java Classes for MDSplus data access

More recently, a set of Java classes has been developed which provide access to MDSplus data. Java is an object-oriented language used predominantly for world wide web applications in conjunction with a web browser. The Java classes provide the same capabilities as the remote data access libraries, and can be used to develop Java applications. These Java classes communicate using the same mechanism as the remote data access routines, and utilize the same server processes on the host system.

Java Applets

Several Java applets have been developed for displaying Alcator C-Mod data, shot cycle status information and video segments. Applets are applications written in Java which run in the context of a world wide web browser such as Netscape Navigator™ [9] or Microsoft Internet Explorer [10].

A scope applet has been developed that behaves much the same as the X window based dwscope described earlier. (See figure 3) Although this applet can be used instead of its X window based relative, it is probably more useful for publishing waveforms on the world wide web.
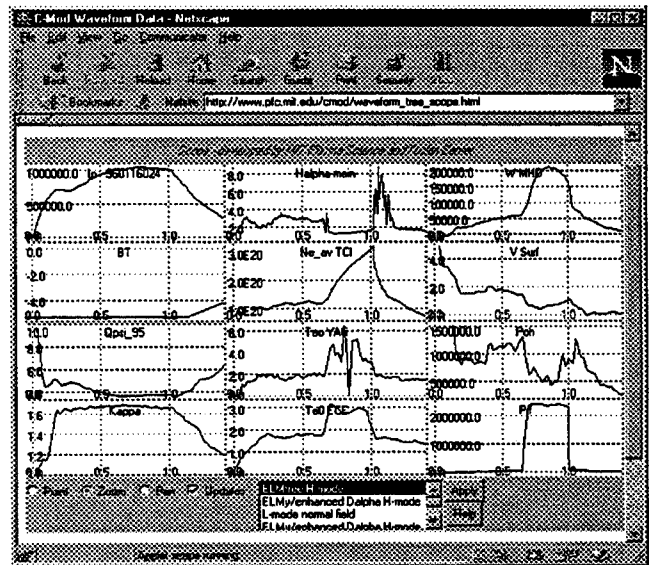

Figure 3 - Java Scope

A C-Mod state applet has been developed which tracks the shot cycle. (See figure 4) It displays the current state of the machine and provides the same audio announcements that are heard in the control room. The display is continuously updated indicating which phase of the shot cycle (i.e. initialization, check, pulse, re-cool) is currently underway.
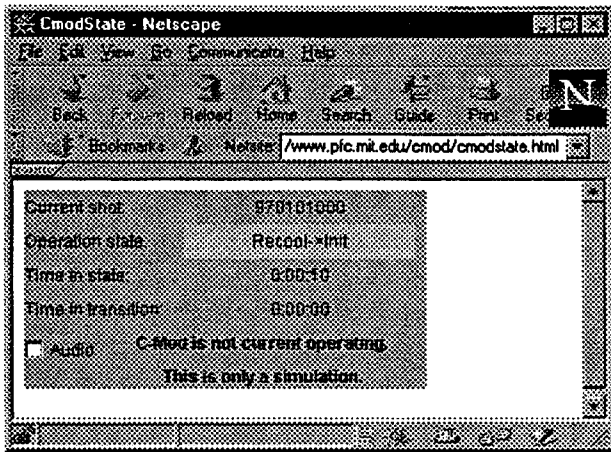
Figure 4 - C-Mod State

Several applets utilize a Java class which recognizes the occurrence of an MDSplus event. MDSplus events are a feature of the MDSplus data system which can be used to announce a significant event has occurred. An event may be generated when a particular set of data is available, for example. Events are also generated for each shot cycle state change and these events are used to drive the C-Mod state applet described above. The events can be also used to force web page updates to load new information.

Remote Access Performance Considerations

There are many performance considerations that need to be addressed when developing remote access applications. Since fusion experiments generate enormous quantities of data, efficient use of network bandwidth is critical. Some of the experience gained from the remote control room experiment and from our ongoing remote collaborations have lead to application enhancements that have changed the performance characteristics of remote access applications.

When the remote control room experiment was performed in 1995, most of the data was displayed using the dwscope program. This program could be run on computers at MIT (remotely) with the display directed to workstations at LLNL or run locally on LLNL UNIX workstations (locally) accessing the data using the remote data access routines. These two modes of operation yielded two very different performance behaviors. When run remotely, the traces were drawn much more quickly than when the program was run locally. However,

panning and zooming of the plots occurred much more slowly on the remote programs than the local programs. The reason for this behavior is that when run remotely, the program sends the x/y coordinates of the plot points to the X display server. When the plot is panned or zoomed, new sets of x/y coordinates are sent to the X display server to reflect the new scaling. On the other hand, when the dwscope program is run locally on the UNIX workstation, the full X and Y vectors are sent in floating point representation one time from the host to application via calls to the remote data access routines. The actual drawing of the plot is done locally and rescaling due to panning and zooming is computed locally. Since the data vectors are sent as floating point values (32 bits per number) instead of x/y coordinates as in the remote dwscope (16 bits), the local dwscope will display the trace initially half as fast as the remote dwscope.

Since the behavior differences in these two modes of operation were considerable, users often ran in both modes, using local dwscope programs for traces that they would likely want to manipulate frequently and remote dwscope programs for traces that they would only view and not rescale.

Since the time of the remote control room experiment, we have added vector reduction optimization to compute the minimum set of x/y coordinates necessary to generate a trace which is graphically equivalent to what would appear if all the points had been drawn. Since many of the plots are drawn in rectangles of nominally 100 X 400 pixels, one can draw a trace of 100,000 samples using fewer than 1000 x/y coordinate pairs. This optimization is reminiscent of those used when drawing vectors on graphics terminals which had slow 2400-9600 baud serial lines.

This vector reduction optimization drastically changed the performance behavior of the remote dwscope program such that the local version is rarely used.

Future development plans

The MDSplus data acquisition system is currently being ported in its entirety to UNIX and Win32 platforms. The port will enable the use of MDSplus on a larger variety of computers. If native MDSplus routines were available on the collaborator's computer, MDSplus data could be exchanged in a much more compressed format. This would greatly reduce the network load when accessing the data remotely. An good illustration of this benefit is accessing a time base for a measurement. In MDSplus, linear time bases can be represented with only a few numbers. When accessing the time base of a measurement using the current remote access library routines, the host will send back the entire array of time stamps for each point of the measurement. Depending on the measurement, the number of time stamps could reach 100,000 or more. If the code that interpreted the time base information resided on the client system, only a few numbers would need to be sent from the host and the time stamps could then be calculated on the client system.

In conjunction with the MDSplus port, we will be developing additional client/server interfaces to allow a much lower level access to the stored data. We will probably look into the use of CORBA for providing this interface. CORBA, which stands for Common Object Request Broker Architecture, is a developing standard for remote access to objects as found in the C++ programming language. It takes care of translating various data formats if the internal representation of numbers differ between the client and server. It also provides some security mechanism to restrict access to information by unauthorized users. Most Java vendors are also coming out with interfaces between Java and CORBA so Java classes could communicate directly with the CORBA compliant servers.

As the port of MDSplus progresses, we will be moving the MDSplus source code into a version control system called CVS, short for Concurrent Version System. Additional experiments, such as the NSTX experiment in Princeton, have expressed interest in using MDSplus and are willing to assist in porting and documenting the system. CVS provides a mechanism for remote sites to collaborate in the software development and maintenance.

We will continue to expand the quantity of information available from our web site. We plan to add access to Alcator C-Mod summary and logboook data now stored in relational databases. Documentation of the data system and the various utilities will be added to our web pages as well as orientation information for new members of the experiment team.

Longer range plans include the port of major portions of the MDSplus data system to Java.

Summary

We have come a long way in improving the delivery of information to remote collaborators since the remote control room demonstration in 1995 but we hope to accomplish a lot more. The number of remote collaborations on the Alcator C-Mod experiment is increasing. Diagnostics are operated by scientists at distant locations such as Princeton, New Jersey and Austin, Texas. In addition, analysis codes are being fitted with remote access capabilities to enable them to access C-Mod data directly while running on computers in locations such as Toronto, Canada.

Remote collaborations will continue to play an important role in the Alcator C-Mod fusion experiment. It is imperative that we provide them with a working environment as similar as possible to that which they would have if they were sitting at a workstation in the C-Mod control room. For these collaborations to be most successful, there must be a continuous flow of information in both directions. Collaborators must be able communicate freely with personnel at MIT and with other collaborators. Although the transfer of experimental data will likely make up the bulk of the network traffic, other information including audio and video, reports, news items, and schedule information is equally important.

In addition to the work related to Alcator C-Mod, other experiments will be using the MDSplus data system and gaining valuable experience with providing remote access to experimental data and operational information. Also, large modeling codes such as TRANSP have been fitted with MDSplus interfaces to access experimental data directly. We have also been able to provide MDSplus wrappers to other data archives such as the TFTR waveforms thus making the whole set of MDSplus tools available for retrieving and viewing the data both locally and remotely.

As we have more remote collaborations at Alcator C-Mod and other sites adopt the use of MDSplus as their data system or remote data access mechanism, we will learn much more about the needs of remote collaborators. With this knowledge, we hope to continue to make enhancements to improve the ability of scientists to interact with experiments and analysis codes from greater distances. The rapid technology advances in networking driven largely by the world wide web should make this task much easier.

It is becoming more and more apparent that large scale collaborations of scientists will be necessary to maintain the progress in fusion research. The costs associated with building and operating fusion experiments continue to rise so there is a great incentive for sharing these resources with scientists located all over the world. Hopefully the steps we are making now in providing remote information access will lead us to better tools for the future of fusion research.

[1] J.A.Stillerman, T.W.Fredian, K.A.Klare, G.Manduchi, "MDSplus Data Acquisition System", Review of Scientific Instruments, 68(1), January 1997, p 939

[2] T.W.Fredian, J.A.Stillerman, "An X-windows Based User Interface for Data Acquisition and Display", In Proceedings of the APS 18'th Topical Conference on High Temperature Plasma Diagnostics, 1990.

[3] Benjamin Welch et al. "Visible and Ultraviolet Light Measurements from Alcator C-Mod." In Bulletin of the American Physical Society, volume 38, page 1957, American Physical Society, November 1993

[4] S. Horne, et. al., "Remote Control of Alcator C-Mod from LLNL", To be published in the Journal of Fusion Technology, (1997)

[5] I.H.Hutchinson, "Fusion by Internet", IT Automation, 1996

[6] T.A.Casper and W.J.Lennon. "Remote Experimental Site: A Command and Analysis Center for 'Big Physics' experimentation." In proceedings of the IEEE/NPSS 14'th Symposium on Fusion Engineering, San Diego, CA., October 1991.

[7] IDL is a registered trademark of Research Systems, Incorporated

[8] OpenVMS is a registered trademark of the Digital Equipment Corporation

[9] Netscape and Netscape Navigator are registered trademarks of Netscape Communications Corporation.

[10] Internet Explorer is copyrighted by the Microsoft Corporation.