Fusion Energy Sciences Networking: A Discussion of Future Requirements

D.P. Schissel, General Atomics Fusion Group
M.J. Greenwald, MIT Plasma Science and Fusion Center
W.E. Johnston, Lawrence Berkeley National Laboratory

1. Introduction

The long–term goal of magnetic fusion research is to develop a reliable energy system that is environmentally and economically sustainable. To achieve this goal, it has been necessary to develop the science of plasma physics, a field with close links to fluid mechanics, electromagnetism, and nonequilibrium statistical mechanics. Nuclear fusion, the power source of the stars, has been the subject of international research since its worldwide declassification in the late 1950s [1]. The highly collaborative nature of the Fusion Energy Sciences (FES) combined with a few experimental facilities and a computationally intensive theoretical program are creating new and unique challenges for computer networks.

FES research is a worldwide effort conducted at some 90 sites in the United States, 60 in Europe, 40 in Japan, and several each in South America, China, and Australia. Fusion experiments have a dual role: providing data for the advancement of plasma science and increasing plasma parameters (densities, temperatures, etc.) toward the levels that would be needed in a power production reactor. In the United States, experimental magnetic fusion research is centered at three large facilities (Alcator C–Mod, DIII–D, NSTX [2,3,4]) with a present day replacement value of over $1B; clearly too expensive to duplicate. As these experiments have increased in size and complexity, there has been concurrent growth in the number and importance of collaborations between large groups at the experimental sites and smaller groups located at universities, industry sites, and national laboratories.

Teaming with the experimental community is a theoretical and simulation community whose efforts range from the very applied analysis of experimental data too much more fundamental theory like the creation of realistic non–linear 3D plasma models. There are presently 5 major high–end computational fusion projects funded by the SciDAC initiative [5]. These computational fusion physicists make up one of the largest user groups at the National Energy Research Scientific Computing Center (NERSC) [6].

The SciDAC initiative is also funding the National Fusion Collaboratory Project [7] that is creating and deploying collaborative software tools. This work includes creating a fusion computational and data Grid as well as new and innovative collaborative visualization capabilities. The overall objective is to allow scientists at remote sites to participate as fully in experiments and computational activities as if they were working at a common site. An implicit assumption of the Collaboratory Project is that adequate networks of quantifiable high performance are available on–demand for priority tasks.

2. Network Usage Scenarios

The demands placed on computer networks by FES research are illustrated below in two usage scenarios. These scenarios cover remote experimental operations, remote code development and invocation, online code and data access, planning and coordination tools, and look 3 to 5 years out in the requirements estimates.

2a. Experimental Science

The three main magnetic fusion experimental sites in the United States operate in a similar manner. The gross tokamak machine hardware parameters are configured before the start of the experimental day. Magnetic fusion

experiments operate in a pulsed mode producing plasmas of up to 10 seconds duration every 10 to 20 minutes, with 25–35 pulses per day. Throughout the experimental session, hardware/software plasma control adjustments are made as required by the experimental science. These adjustments are debated and discussed amongst the experimental team (typically 20–40 people) with most working on site in the control room (Fig. 1) but with many participating from remote locations. Decisions for changes to the next plasma pulse are informed by data analysis conducted within the roughly 15 minute between-pulse interval. This mode of operation places a large premium on rapid data analysis that can be assimilated in near–real–time by a geographically dispersed research team.

For each plasma pulse up to 10,000 separate measurements versus time are acquired at sample rates from kHz to MHz, representing hundreds of Megabytes of data. Improvements in plasma diagnostic techniques have made it increasingly feasible to demonstrate excellent correlations between experimental results and theoretical models. In particular, the development of diagnostic instruments that can make high–resolution measurements of electric and magnetic fields and cross-sectional measurements of turbulent fluctuations have greatly improved the basic understanding of the mechanisms controlling plasma confinement.

Figure 1 *Pictured is a section of the control room of the Alcator C–Mod tokamak, a high–field magnetic confinement experiment at MIT. Using the MDSplus data system combined with components from the Globus toolkit, remote collaborators can exchange data securely with researchers working on the Alcator C–Mod, DIII–D, or NSTX tokamaks.*

The highly collaborative environment of the three experimental facilities motivated the creation and adoption of a common data storage mechanism MDSplus. Presently each of the three experimental facilities is operating their own MDSplus data server. This client server system allows data access via the Internet from any application run by a user who has been given the authorization to access the specific data. Commercial relational database technology has been adopted to aid the exploration of these large experimental MDSplus data repositories. Additionally, an Electronic Logbook has been created for simultaneously sharing information among hundreds of geographically diverse researchers. This Logbook has been built upon relational database technology for rapid searching of the many hundred thousand text based entries.

The computational emphasis in the experimental science area is to perform more, and more complex, data analysis between plasma pulses. For example, today a complete time–history of the plasma magnetic structure is available between pulses by using parallel processing on Linux Beowulf clusters. Five years ago, only selected times were analyzed between pulses with the entire time–history completed overnight. Five years from now, analysis that is today performed overnight will be completed between pulses. Such enhanced between pulse data analysis will include more advanced simulations. The ability to more accurately compare experiment and theory between pulses will greatly enhancing the value of experimental operations. Today, these comparisons are done after experimental operations have concluded when it is too late to adjust experimental conditions.

It is anticipated that the data available between pulses will exceed the 1 GByte level within the next 5 years. Overall data transfer rates must be fast enough to allow time for detailed analysis and subsequent examination by the scientific staff within the 20 minutes between plasmas. Peak network rates on the order of 500 Mbits/sec are required if a third of a minute is allowed to transfer the entire repository in order to allow the remaining intra–pulse time to be used for computational and human analysis. This peak rate is required intermittently over the course of a year. Typically, experimental operation on one of the three main machines is 8 hours or 30 pulses a day, 5 days a week for approximately 20 weeks a year (two or more machines can be operating at the same time). During an experimental day anywhere from 5 to 10 remote sites can be participating. Although the entire repository is not

transferred to each remote site, subsets of the data are transferred for visualization and analysis with results being written back intermittently into the main MDSplus data storage system over days, weeks, or years. It is the dynamic nature of the data repository combined with the large number of distributed users that makes replication at remote sites more than just a simple task.

With the creation of more data between pulses there exists an increasing burden to assimilate all of the data. Enhanced visualization tools are presently being developed that will allow this order of magnitude increase to be effectively used for decision making by the experimental team. Clearly, the movement of this quantity of data in a 15–20 minute time window to computational clusters, to data servers, and to visualization tools used by an experimental team distributed across the United States and the sharing of remote visualizations back into the control room will place a severe burden on present day network technology.

2b. Theoretical and Simulation Science

Although the fundamental laws that determine the behavior of fusion plasmas are well known, obtaining their solution under realistic conditions is a computational science problem of enormous complexity. This is due in large part to the strong non–linearity of the problem and the enormous range of temporal and spatial scales and is therefore a challenge that can only be met with advanced scientific computing. Because of this complexity, researchers have a long history of productive use of advanced computation and modeling. Such work is typically broken down into large (macro) and small–scale (micro) disturbances that relax the plasma to a lower energy state.

Integrated simulation of magnetic fusion systems involves the simultaneous modeling of the core plasma, the edge plasma, and the plasma–wall interactions. Presently the SciDAC initiative is funding five separate programs with the long–term goal of creating an integrated fusion plasma simulation. The five SciDAC areas of study are: 1) Plasma Microturbulence, 2) Plasma Magnetohydrodynamic (MHD), 3) Wave–plasma Interactions, 4) Edge Plasma Physics, and 5) Edge Atomic Physics. Each project is a multi–institution collaboration with researchers spread across the United States.

Datasets generated by these simulation codes will approach the 1 TB level per run within the next three to five years. These datasets will be analyzed like experimental plasmas are analyzed to extract further information. Therefore, the data repository for simulations will be dynamically evolving rather than write–once. Network rates of 500 Mbits/sec to 1000 Mbits/sec are required, similar to Section 2a, if we assume that subsets of a simulation code run (~ 1 GB) are to be visualized and analyzed interactively (~ 20 s). Several of the SciDAC fusion projects have begun to use the MDSplus data system for data management and in the next five years the remainder will most likely do the same. These large datasets will most likely be dispersed across the United States and will be made available using the MDSplus client/server interface. To facilitate efficient data transfer, parallel network I/O will need to be made routine between computational computers, data repositories, and visualization systems.

Simulation data sharing will include new complex visualization capability that is presently being developed. One such example is shared visualizations between tiled display walls (Fig. 2). Such visualization sharing can be used for working meetings and for communication with the experimental control room. As stated in the section above, the desire to perform more complex simulation between experimental plasmas is strong and it will be desirable for data from these SciDAC simulation codes to be available in a 15–20 minute time window. Eventually, there will be an integrated simulation of the plasma that can be compared to the plasma itself all between plasma pulses. Such large–scale simulations using computer resources and data repositories shared across the United States, combined with the eventual compression of this analysis into a 20 minute time window will place a severe strain on existing network capability.

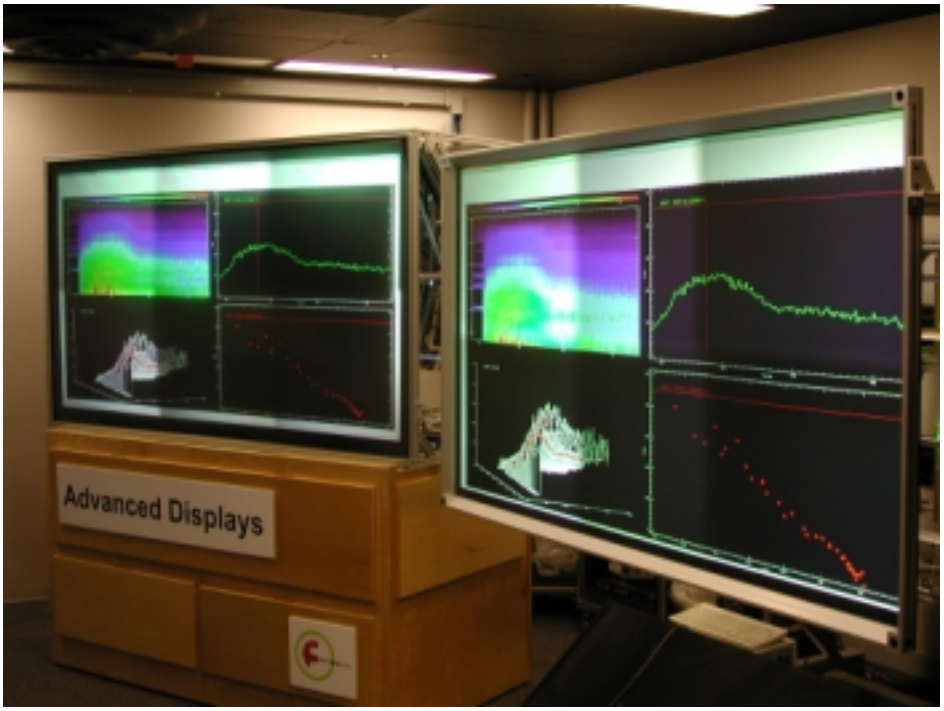

Figure 2 *Shared visualizations between tiled display walls will offer a collaborative visualization environment that is presently not available for magnetic fusion scientists.*

3. Network Services Common to both Scenarios

In addition to the network bandwidth requirements described above, the nature of FES research also leads to requirements for advanced network services. As in other sciences, valuable resources such as computers, data, instruments and people are distributed geographically and must be shared for successful collaboration. In fusion, the need for real-time interactions among large experimental teams and the requirement for interactive visualization and processing of very large simulation data sets are particularly challenging. Shared tools and solutions are especially valuable – reducing problems with n x m or n! interactions to more tractable scales.

The apparently conflicting requirements for transparency and security in a widely distributed environment point up the need for efficient and effective services in this area. Central management of PKI or equivalent technologies using "best practices" and providing 24x7 support is essential. Further, it is essential that the user authentication framework and operational environments are such that common policy may be negotiated among international collaborators in order to enable collaborations to span international boundaries and between application development and site security groups. Development of mutually agreed upon tools and protocols for resource authorization is equally important.

As fusion collaboratory activities grow, the needs for global directory and naming services will expand as well. A hierarchical infrastructure with well–managed "roots" can provide the necessary glue for many collaborative activities. Analogous to the Internet's domain name services, this infrastructure would give local resource managers needed flexibility while maintaining global connectivity and persistence. A global name service could even solve the longstanding problem in the field of variable name translation between codes or experiments. Distributed computing services for queuing and monitoring are also needed. These must be easy to configure and deploy and robust in operation.

The fusion applications described above will also require network quality of service (QoS) in order to provide guaranteed bandwidth at particular times or with particular characteristics. Such QoS will be required to tailor the

network to match the time dependent demand requirements rather than maintaining sustained bandwidth. For experimental collaborations, low network latency with minimum jitter and very low packet loss are essential if instruments and experiments are to be controlled remotely. It is anticipated that the next generation of fusion experiments will be routinely operated by remote teams. Relatively small quantities of monitoring data must be passed back reliably and quickly. Moderately large quantities of intermediate data are produced by simulations in burst mode before a code's conclusion and should be made available as quickly as possible or even during the code run for computational steering. End to end performance is crucial and must include partners at universities, private companies or international sites. Real-time network performance monitoring and problem resolution tools are essential.

Finally, both scenarios discussed above will generate such an increasing quantity of data that new and improved ways to efficiently mine the MDSplus repositories to extract knowledge will need to be developed. Such new data mining capabilities will be required to formulate, analyze, and implement basic induction processes that facilitate the extraction of meaningful information from unstructured data. These capabilities will allow the semi–automatic discovery of knowledge in the form of patterns, changes, associations, anomalies, rules, and statistically significant structures and events from the large MDSplus repositories. This type of efficient means of sifting through large quantities of data will reduce the amount of information that must be transferred to the scientists thereby reducing the strain on the network.

4. Conclusions

The fusion community anticipates a computing–for–data–analysis model that involves moving data at data rates of approximately 500 Mbit/sec between 40 sites. This network traffic is periodic with large bursts (10–30 seconds out of 20 minutes, 8 hours a day, 5 days a week, 20 weeks a year). When the traffic appears on the network it requires guaranteed bandwidth, however the network could be used for other purposes during the remaining ~99% of the time. This traffic will flow between the sites of the major experimental participants and their many collaborating institutions.

The distributed nature of the Fusion Community and its paradigm of widely shared data require several types of persistent middleware that is operated by a reliable, long–lived third–party. This middleware includes

- The PKI Certificate Authorities that enable strong authentication of the community members and the use of Grid security tools and services.
- Directory services that can be used to provide the naming root and high level (community-wide) indexing of shared, persistent data that transforms into community information and knowledge base as it matures through analysis and comparison with simulations
- Management functions for network Quality–of–Service that provides the request and access mechanisms for the experiment run time, periodic traffic noted above.
- Efficient means to sift through large data repositories to extract meaningful information from unstructured data.

References

[1] http://www.iaea.org/

[2] http://www.psfc.mit.edu/cmod/

[3] http://web.gat.com/diii-d/

[4] http://nstx.pppl.gov/

[5] http://www.osti.gov/scidac/

[6] http://www.nersc.gov/

[7] http://www.fusiongrid.org/