NATIONAL FUSION COLLABORATORY: EXECUTIVE SUMMARY

The National Fusion Collaboratory will advance scientific understanding and innovation in magnetic fusion research by enabling more efficient use of existing experimental facilities and more effective integration of experiment, theory, and modeling. Specifically, this project will create and deploy collaborative software tools throughout the national magnetic fusion research community comprised of over one thousand researchers from over forty institutions. Built on a foundation of established computer science toolkits, successful deployment of the Collaboratory will nevertheless require significant computer science research to extend the toolkits beyond their present capabilities. The National Fusion Collaboratory will enable networked real–time data analysis and instantaneous communication amongst geographically dispersed teams of experimentalists and theoreticians. This represents a fundamental paradigm shift for the fusion community where data, analysis and simulation codes, and visualization tools will be thought of as network services. In this new paradigm, access to resources (data, codes, visualization tools) is separated from their implementation, freeing the researcher from needing to know about software implementation details and allowing a sharper focus on the physics.

Magnetic Fusion Research. Fusion, the power source of the stars, has been the subject of international research since the late 1950's. Experimental magnetic fusion energy research in the United States is centered at three large facilities with a present day replacement value of over $1B. Teaming with this experimental community is a theoretical and simulation community that concentrates on the creation of realistic 3D plasma models. As the capabilities of wide area networks have increased, more researchers have begun to collaborate with the experimental institutions from their home laboratory rather than travelling for experiments. While the community has made significant progress in accommodating this new pattern of use, much remains to be done to take full advantage of newly emerging technologies.

Goals of the Collaboratory. The aim of the Collaboratory is to:

- Create transparent and secure access to local/remote computation, visualization, and data servers.
- Develop collaborative visualization that allows interactive sharing of graphical images among control room display devices, meeting room displays, and with offices over a wide area network.
- Enable real–time access to high–powered remote computational services allowing such capabilities as between pulse analysis of experimental data and advanced scientific simulations.

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. For each pulse up to 10,000 separate measurements versus time are acquired representing hundreds of Megabytes of data. Decisions for changes to the next plasma pulse are made by data analysis conducted within the roughly 15 minute inter–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.

Benefits to Fusion. The National Fusion Collaboratory will increase physics productivity by

- Enabling more efficient utilization of experimental time on the three large facilities through more powerful between pulse data analysis resulting in a greater number of experiments at less cost.
- Allowing more transparent access to analysis and simulation codes, data, and visualization tools resulting in more researchers having access to more resources.
- Creating a standard tool set for remote data access, security, and visualization allowing more researchers to build these services into their own tools.
- Facilitating the comparison of theory and experiment.
- Facilitating multi–institution collaborations.

The Collaboratory will also increase the productivity of code and tool developers by

- Supporting SciDAC computational initiatives under the collaboratory framework.
- Supporting more users with fewer installations at reduced cost.
- Facilitating shared code development projects resulting in more rapid code creation.
- Creating a standard tool set for remote data access, security, and visualization allowing these services to be easily built into new tools.

Computer Science Research. To accomplish these goals, fusion scientists with expertise in large experiments and simulation code development have joined computer scientists with expertise in security, distributed computing, and visualization to form a closely coordinated team. This team, leveraging existing computer science technology where possible, will deploy a collaboratory prototype. For requirements not met by current capabilities, new technologies will be developed. The variety of users, resources, applications, and policies encountered will serve as an excellent proving ground for new technologies that prepares the way for their use in other scientific disciplines.

Benefits to Computer Science Toolkits. The Collaboratory will enhance existing toolkits by

- Enabling automatic propagation of security credentials in multi–server contexts.
- Allowing complex use policy implementation and remote monitoring of computational resources.
- Extending security architecture to encompass commercial databases.
- Creating pre–emptive scheduling capability and advance reservation of computational resources.
- Extending the Access Grid to large tiled display walls.
- Developing a quantitative visualization capability allowing data comparison with uncertainties.
- Supporting numerous SciDAC computer science projects that cross–cut with the Collaboratory.

The computer science research necessary to create the Collaboratory is centered on three main activities: security, remote and distributed computing, and scientific visualization.

Security. The sharing of data, codes, and visualization tools as network services requires a system for protecting these valuable resources against unauthorized use. The Collaboratory will exploit state-of–the–art authentication, authorization, and encryption technologies provided by the Globus Security Infrastructure and the Akenti authorization service. Existing fusion community codes will be modified to use this infrastructure for remote execution and data access. To meet the needs of the Collaboratory, the current version of these middleware tools will be extended. In particular, it will be necessary to incorporate rules for fair use of shared resources into the policy enforced by the security model and to enhance tools that enable valid credentials to propagate automatically from resource to resource.

Distributed Computing. The remote and distributed computing requirements of the Collaboratory will utilize the Globus facilities including remote job scheduling, monitoring, exception handling, and accounting. This will enable researchers and their institutions to share the community's computational resources. The components of the Globus toolkit that can be immediately deployed to create the foundation of the Collaboratory are Grid Information Services, Grid Security Infrastructure, and Globus Resource Allocation Manager. Research components that are required to fully meet the needs of the Collaboratory and that will create new functionality for Globus include managing batch versus preemptive job priorities, providing status display and accountability to users, monitoring the adherence of resources to policies, and providing advance reservations. Fusion community codes to be adapted to the Collaboratory include serial and parallel MHD stability codes and serial and parallel transport codes.

Scientific Visualization. The demand placed on visualization tools by the Collaboratory is intense due to both the highly collaborative nature of fusion research and the dramatic increase in data resulting from the enhanced computing capabilities. The visualization component of the Collaboratory will focus on the development of a collaborative control room, collaborative meeting room, and enhanced visualization tools. The collaborative visualization requirements will utilize the Access Grid that enables distributed meetings and collaborative teamwork sessions. Extensions to the Access Grid software include a more closely integrated shared experience with the researcher's current work environment and the support of large tiled displays that will provide collaborative capabilities to large-format remote visualizations. New software will be tested on display walls that already exist within the proposal team. Extensions to display wall software include the ability to have visualizations not tied to an individual projector allowing the size of the visualization to vary depending on the researcher's need. Extensions to the visualizaton toolkit will be the ability to quantitatively compare theory to experiment with uncertainties.

Project Scope. The three–year project has just begun and is a collaboration between experts in the magnetic fusion and the computer science and enabling technology communities funded by the USDOE SciDAC program. Further information on the Collaboratory is available at http://www.fusiongrid.org/.