10th ICALEPCS Int. Conf. on Accelerator & Large Expt. Physics Control Systems. Geneva, 10 - 14 Oct 2005, TH2.2-7O (2005)

Visions for data management and remote collaboration for ITER

M. Greenwald[1] D. Schissel[2] J. R. Burruss[2] T. Fredian[1] J. Lister[3] J. Stillerman[1]

[1] *Massachusetts Institute of Technology, Cambridge, USA*
[2] *General Atomics, La Jolla, USA*
[3] *Centre de Recherches en Physique des Plasmas, EPFL, Lausanne, Switzerland*

Introduction

This paper outlines a vision for data management and remote collaboration systems for ITER, an international "burning plasma" magnetic confinement experiment and the next major step toward proving the scientific viability of controlled fusion as an energy source [1]. It will be the largest and most expensive scientific instrument ever built for fusion research. A unique feature in the operation of fusion energy experiments is the requirement to access, analyze, visualize and assimilate data, between shots in near-real-time, to support decision making during operation. This contrasts with large experiments in other fields, such as high-energy or nuclear physics which operate primarily in a batch mode. Fusion experiments put a particular premium on near real-time interactions with data and among members of the team. Enabling effective international collaboration on this scale is technically demanding, requiring powerful interactive tools and provision of a working environment that is equally productive for on-site and off-site personnel engaged in experimental operations.

Challenges

The need for efficient between-shot analysis and visualization is driven by the high cost of operating the experimental facility. ("Shots" are the basic units of fusion experiments. Today, a typical large facility might take shots at a rate of 2-4 per hour and accumulate about 2,000 shots per year.) The average cost per shot for ITER, defined here as the integrated project cost divided by the total shots estimated over the project lifetime, will approach one million US dollars. Thus, the number of shots required to optimize performance and to carry out experimental programs must be minimized. This translates into a need to carry out extensive analysis and assessment immediately after each shot. ITER shots will also be much longer than on most current machines and will generate much more data, perhaps a Terabyte per shot. The quantity of data itself, perhaps 2 PB per year, will likely not be a technical challenge at the time that ITER will be operating – about a decade from now. However long-pulse operation will require concurrent writing, reading, visualization and analysis of experimental data. More challenging is the integration across time scales. The data set will encompass more than a factor of 10^9 in significant time scales, leading to requirements for efficient browsing of very long data records and the ability to describe and locate specific events accurately from within very long time series. Not only will ITER be an expensive device, as a licensed nuclear facility and the first reactor-scale fusion experiment, security of the plant will be a paramount concern. The data systems must balance these requirements with the need to keep data access as open to the participating scientists as possible. Mechanisms and modalities for remote control must also fit into a robust security model. Further, the 10-year construction and 15+ year operating life for ITER will encompass evolutionary and revolutionary changes in hardware, software and protocols; thus the system must be based on a conceptual design that is extensible, flexible and robust enough to meet new requirements and be capable of adapting and migrating to new technologies and to new computing platforms as they arise. Backward compatibility, the ability to read old data and perform old analysis, must be maintained over the life of the experiment.

The international nature of the ITER collaboration presents a set of well-known challenges. Remote participants will require transparent and rapid data access, allowing them to participate in real-time, managing diagnostic systems and leading experimental sessions. Fortunately, the fusion community has extensive experience in this area to draw on [2-6]. We have found that the greatest challenges involve communication and engagement between remote and local researchers rather than specific technical problems. With the research team distributed geographically and communications more awkward, relationships are difficult to establish and maintain. The ITER project will cross numerous administrative boundaries, with its resources such as networks and computers not owned or controlled by a single project or program. This will add to the complexity of the deployed systems and increase stress on the security model. In such an environment, maintenance, resource management, monitoring, troubleshooting and end-to-end problem resolution are not straightforward. Finally, the higher latency intrinsic in intercontinental connections challenges network throughput and interactivity. The overall system architecture needs to optimize performance across a heterogeneous user base, with some connecting locally at low latency and some over the wide-area networks at high latency.

Architectural Outline and Implementation Strategy

Based on previous experiences [7-9], we can propose broad outlines for an architecture that could satisfy requirements for an ITER data system. As shown in Figure 1, the data would be organized into two structures. The main repository would contain all the large arrays of raw and processed data. This data would be indexed by its independent parameters, shot number, time, position, etc, and be organized to provide the highest possible access speed. It would not be conveniently searchable by the content of the data. The second data structure would be a flexible, queryable database, which would contain all metadata, configuration information, the parameters and results of high-level analysis and indexing into the repository. This database would support a data directory and diverse views into the repository, including relations and data interdependencies. Together, the two structures would provide a complete, coherent and self-descriptive structure. User interactions would be service-oriented; access would be through a single, simple, standard API and all applications would be data-driven. To ensure long life for our systems, implementations should be based on open standards wherever practical The architecture should support secure, transparent remote access to all services. Systems for remote participation should also be based on open standards, scalable to match the scope of the envisioned collaborations, interoperable to support the heterogeneous environment and extensible to meet future requirements.

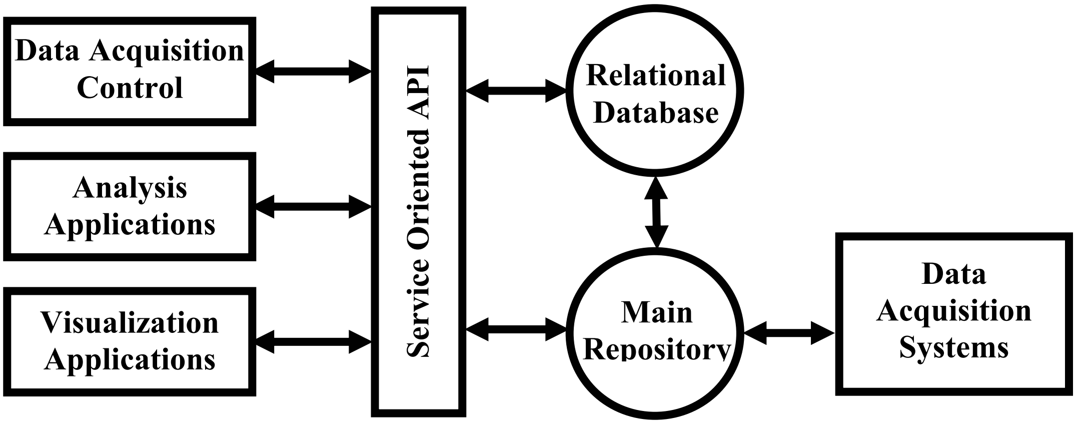

Figure 1. The proposed data system architecture would be comprised of two data structures: the main repository containing all the large multi-dimensional arrays and a queryable database containing all metadata, ancillary and high-level physics data. User interactions would be service-oriented; access would be through a single, simple, standard API and all applications would be data-driven.

While, it is too early to choose specific technologies and implementations for an experiment that will not commence operation until the middle of the next decade, some important work should begin now. First, we must reach agreement on the broad requirements, define an approach and begin building the team that will carry out the work. A set of general principles for the data system should be developed next, along with conceptual designs for the system architecture. These abstract principles must be robust with respect to the technologies that implement them. At that point, work on specific designs and implementations can begin with an emphasis on rapid prototyping and testing. Where possible, these tests should include deployments on the current generation of experiments. It also may be possible to test some of the long pulse capabilities on plasma simulation codes whose output data has many of the relevant characteristics. Systems for supporting remote participation should be incorporated into the working environment of operating facilities. Prototyping must be seen as an integral part of the concept design process with potential users exposed to the new systems under realistic operating conditions. Only then would proven implementations be expanded and elaborated to meet the full set of requirements.

We should seek commercial solutions where they are available, however, because of the uniqueness of the challenges described above, we cannot expect the software industry to have a comprehensive solution, even by the middle of the next decade. It will be left to us to innovate as required to solve these problems. Our strategy therefore should be to make use of fusion community expertise in partnership with industry and the computer science research community, to devise solutions for ITER data management.

General Features Required

We can define a set of general features that the data and remote participation systems should embody. First among these are flexibility and extensibility to meet new requirements, which will certainly arise over the life of the project. Past history suggests that the approaches taken must be scalable to allow for requirements to grow far beyond initial specifications. Hardware, software standards and protocols will evolve and our systems must be able to benefit from these changes with minimal dislocation. The generous use of modularity and layering should create systems capable of continuous improvement and modification. Given the heterogeneous nature of the ITER collaboration, we will want the system to be accessible from the widest set of computing platforms practical. (And we must anticipate significant turnover in platform types through the project life.) A security model must be built in from the beginning, supporting “single sign-on” with strong authentication. The major threat faced will be to the integrity of the data or to the facility itself; data confidentiality is not a primary driver for security. In such a large and diverse collaboration, tools to support distributed authorization and resource management at appropriate levels of granularity should be part of the system as well. The security implementation should use the best features of application and perimeter security models, and allow highly productive scientific collaboration over the wide area networks without endangering plant security in any way.

Data System Features Required

As discussed above, the data system must provide a coherent, integrated view of all data through simple interfaces and tools. This would include all raw, processed and ancillary data; calibrations, geometry, data acquisition setup analysis parameters and so forth. Metadata would be stored with every data item, allowing users to determine where the data came from, when it was written, who was responsible for it along with data types, size, shape, structure, comments, labels, creating a coherent self-descriptive structure. Logical relationships and associations among data elements would be maintained and stored as data themselves. The result is that data would be more easily interpreted and have an extended useful life. The structures could be traversed independently of reading any particular data item. A casual user would often find the data they are interested in by browsing through the repository. One likely form for representing data relationships would be multiple hierarchical structures. With the relations among data elements made explicit, generic applications could be written that would be keyed to structure and context. When additional data items were added, these applications could treat them exactly like other items that occur at the same place in a hierarchy. More generally, all applications would be data-driven, with parameters read from the repository and database, rather than imbedded in the applications or read from private data structures. For example, the names and sequence of automated processing routines would be part of the database and would drive post-shot analysis.

The system must support a rich set of primitive and composite data types with arrays of arbitrary shape and be easily extensible to support data entities not envisioned at the outset of the project. There should be intrinsic support for error estimates and data validation with multilevel, definable data quality metrics. Since analyzed data is produced from collections of raw and processed data and is itself often the source of more highly analyzed data, it is important to maintain a record of the processing “tree”. Some form of validation inheritance needs to be supported to allow users to easily verify the integrity of derived quantities. Since each data item is time-stamped, such tools can and should also verify that all derived quantities were computed and stored in the proper order. The data must be capable of being queried by content allowing for rapid location and assembly of data subsets. A capable and flexible data directory should be provided. With somewhere in the range of 10^5 – 10^6 named data items, it will be crucial that users have the tools to navigate the name space efficiently.

Overall, the system should be service rather than file oriented, with emphasis on accessibility rather than portability. Data system tasks would be carried out by a loosely coupled set of services running on a distributed set of servers. For the ITER data system, each service would refer to the same common data structure, reading data, presenting results to users and adding value to the database by writing processed data back. The interfaces would be simple and generic; applications would describe the tasks to be accomplished, but would not prescribe how they are to be done. Implementation details and complexities (like the distributed nature of the servers) should be hidden whenever possible. Service discovery would likely become an important capability in a vast collaboration like ITER. An example of a widely used computational service in fusion research is the fusion-grid deployment of TRANSP, a very large code (>million lines), used for transport analysis [10, 11, 12]. Providing this service and consolidating operation of this code to a single site benefited users by relieving them of the responsibility (and substantial effort required) to maintain up-to-date versions of the code. At the same time, it reduced the workload on the development team, allowing them to support users on a single cluster of systems, in a well characterized and uniform environment. In this instance user and service security was implemented with x.509 identity certificates and a distributed authorization system.

The ITER data system will need to support pulses in excess of 300 seconds. As noted above, the challenge lies not in the total amount of data accumulated, but in the need for concurrent reading and writing during a pulse and support for efficient access into very long records. The system should be capable of providing an integrated view across a wide range of time scales, ranging over a factor of 10^9 from sub-microsecond for electron-scale fluctuation measurements to the total discharge length approaching 1,000 seconds. With this dynamic range, support for synchronous and asynchronous burst sampling will be essential. Capabilities for data filtering and pre-analysis could help reduce the amount of data transferred to applications while data browsing could be greatly aided by intelligent pre-fetching. Support for event identification, logging and tracking will be needed by users trying to correlate physical phenomena across the large set of data signals and the wide dynamic range of the time base. The data system should support integrated and shared workspaces, such as an electronic logbook and tools to enable higher levels of information organization such as runs, tasks, campaigns and so forth. There is a need to integrate quantitative data, run and shot comments, experimental proposals, runs summaries, presentations and publications. Tools that implement these features should provide indexing into the associated data repositories which underlie them.

Remote Participation Features Required

The primary requirement in support of remote participation is secure, transparent access to all experimental data by off-site collaborators. The most demanding requirements arise for researchers working in near real-time on ongoing experiments. Data processing would also be provided within a service paradigm; we envision deployment of shared application servers, allowing maximum use of shared tools while hiding as much internal complexity as possible. Their location, close to the data archive, close to users or close to developers, would be optimized depending on whether the application was data intensive, display intensive or support intensive. Off-site researchers will also need real-time information on the machine's status, shot cycle and the status of all data acquisition and analysis applications. The ability to easily share complex visualizations and applications among remote participants must also be supported in tandem with interpersonal communications that are flexible and reliable.

Ad hoc interpersonal communications are essential activities in the control rooms of current fusion experiments. Significant effort has already been devoted to extending this environment to off-site locations [12]. For at least 15 years, fusion researchers have operated important diagnostics on major facilities remotely, over the internet [2]. Full remote operation of fusion experiments has also been demonstrated [3]. It is now routine for fusion scientists to lead experiments without leaving their home institutions. Numerous tools have been developed to create an engaging and capable off-site environment, including streaming of control room audio and video, streaming status information and shared displays and applications [4, 5]. Looking forward, we should be able to exploit the ongoing convergence of telecommunications and computing technologies. Integrated communications involving audio, video, messaging, email and streaming of data should aid in the development of a productive collaborative environment. Advanced directory services need to be deployed that allow people and data streams to be identified, located, scheduled and connected into a flexible communications fabric. The directory would be augmented by "presence" information supplied implicitly or explicitly by users. For example wireless devices can be used to determine a person's location, while shared calendars can reveal availability. The communications systems must support "roles"; that is, there is a need to locate and communicate with people who are distinguished by function rather than name or location. Examples would be the session leader of an experiment or the technical support personnel for various subsystems. While many communications will be asynchronous involving a single query or limited discussion, communications between remote control rooms may drive a requirement for persistent communication spaces. These would allow extended conversations among multiple participants to be carried out in an informal and spontaneous manner. There will also be a need for more structured interpersonal communications, such as group meetings and seminars. Adequate teleconference technologies augmented with data sharing and the capability for recording and playback are available today, though current experiences suggest that much work will be required to make all of the communications tools discussed here, easy to use, reliable and interoperable.

Summary

While ITER operation is many years in the future, given the scope of its mission and the technical challenges that it presents, work on systems for data management and remote participation should begin now. It is too early to choose specific information architectures or technologies, but requirements and the generic characteristics of its data systems should be defined so that early prototyping and testing could begin soon. The principle challenges are not due to the volume of data that the experiment will produce, but arise from the need to access, analyze and assess a massive amount of data immediately after each shot in support of decision making for the experimental run.

The architecture envisioned for ITER has all data stored in a complete, coherent, self-descriptive structure with access and analysis through a single, service oriented interface. The data structures would contain all raw, processed, ancillary data and metadata and should support tools for efficiently navigating and displaying very long times series and other large arrays. Logical relationships and associations between data elements would be represented explicitly by structural information stored in the database allowing flexible traversing and browsing. Metadata and high-level physics data would be stored in a queryable database allowing users to locate and retrieve data based on its content. Applications should be data driven, with no parameters stored in code or external data structures.

Shared network access to data and applications would emphasize accessibility over portability. The system must be extensible and flexible to meet new requirements and be capable of adapting and migrating to new technologies and new computing platforms as they arise over the long life of the experiment. All systems should provide an easy learning path, while maintaining powerful capabilities for experienced users.

To enable effective remote collaboration, secure and transparent access to the entire data system must be provided. Moreover, effective remote participation in experiments on the scale envisioned will require an immersive working environment for off-site personnel engaged in experimental operations that is as productive as what is on-site. Tools should facilitate scheduled and ad hoc interactions between researchers at geographically dispersed sites. We anticipate the convergence of physical and logical communications channels so that phone, audio, video, email, messaging, and data can be integrated into a common framework. Advanced directory services would be provided that allow people and data streams to be identified, located, scheduled and connected into a flexible communications fabric. Interpersonal communications media could be enriched via remote sharing of displays and applications among researchers. Finally, the entire implementation will depend on responsiveness to user needs, interoperability in a heterogeneous environment, a high degree of reliability and ease of use.

References

[1.] F. Perkins, D. Post, N. Uckan, M. Axumi, D. Campbell, N. Ivanov, N. Sauthoff, M. Wakatani, editors, *"ITER physics basis"*, Nuclear Fusion, **39**, 2137, 1999.
[2.] R. Fonck, G. Cospy, R. Durst, T. Gibney, M. Thompson, and S. Paul, *"Remote Operation of the TFTR BES Experiment from an Offsite Location"*, Rev. Sci. Instruments, **63**, 4803, 1992.
[3.] S. Horne, M. Greenwald, T. Fredian, I. Hutchinson, B. Labombard, J. Stillerman, Y. Takase, S. Wolfe, T. Casper, D. Butner, W. Meyer, J. Moller, *"Remote control of Alcator C-Mod from Lawrence Livermore National Laboratory"*, Fusion Technology, **32,** 152, 1997.
[4.] T. Fredian, J. Stillerman, *"MDSplus remote collaboration support-Internet and World Wide Web"*, Fusion Engineering and Design, **43**, 327, 1999.
[5.] J. R. Burruss, G. Abla, S. Flanagan, K. Keahey, T. Leggett, C. Ludesche, D. McCune, M. E. Papka, Q. Peng, L. Randerson, D. P. Schissel, *"Developments in Remote Collaboration and Computation"*, Fusion Science and Technology, **47**, 814, 2005.
[6.] D. Schissel, J.R. Burruss, A. Finkelstein, S.M. Flanagan, I.T. Foster, T.W. Fredian, M.J. Greenwald, C.R. Johnson, K. Keahey, S.A. Klasky, K. Li, D.C. McCune, M. Papka, Q. Peng, L. Randerson, A. Sanderson, J. Stillerman, R. Stevens, M.R. Thompson,6 and G. Wallace, *"Building the U.S. National Fusion Grid: Results from the National Fusion Collaboratory Project"*, Fusion Engineering and Design, **71**, 245, 2004.
[7.] J. Stillerman, T. Fredian, K. Klare, G. Manduchi, *"MDSplus data acquisition system"*, Review of Scientific Instruments, **68**, 939 1997.
[8.] J. Stillerman, T. Fredian, M. Greenwald, *"WWW interfaces for runtime relational database applications"*, Fusion Engineering and Design, **48**, 63, 2000.
[9.] T. Fredian, J. Stillerman, *"MDSplus: Current developments and future directions"*, Fusion Engineering and Design, **60**, 229 2002.
[10.] R. Hawryluk, *"An empirical approach to Tokamak transport, Physics of Plasmas Close to Thermonuclear Conditions"*, Proceedings of the International School of Plasma Physics, **1**, 19, 1981.
[11.] J. R. Burruss, T. M. Fredian, M. R. Thompson, *"Simplifying FusionGrid Security",* Proc. Challenges of Large Applications in Distributed Environments, 95-105, (2005)
[12.] D.P. Schissel, "Grid computing and collaboration technology in support of fusion energy sciences"**,** Physics of Plasmas **12**, 058104, 2005.