Recommendations for ITER Experimental Operation, U.S. Team Formation and Participation

M. Greenwald, D. Hillis, A. Hubbard, J. Hughes, S. Kaye, R. Maingi (coordinator), G. McKee, D. Thomas, M. Van Zeeland, M. Walker

U.S. Burning Plasma Organization working group: Modes of Participation in ITER

Disclaimer: This is a report of a subcommittee of the U.S. BPO and does not necessarily reflect the views of the U.S. ITER Project Office or Department of Energy

April 24, 2015

Executive Summary

The U.S. Burning Plasma Organization undertook an activity to identify modes of US Participation in ITER, to contribute to and benefit from the science and technology of ITER. The group considered A) recommendations for ITER Research Program Planning procedures, based on US device experience, and B) formation of the U.S. team for optimal participation in ITER. Each of these two areas is summarized below.

Chapter 1: Recommendations for ITER Research Program Planning and Experiment Execution

Discussions within the BPO group and with the ITER Organization (IO) were conducted on ITER experiment and program planning. Here, the resulting recommendations for modes of operation and analysis in ITER are presented, as are strategic recommendations for ITER experimental operation procedures and support systems by the IO. This area is under the responsibility of the ITER Directorate of Plasma Operation, and the work was triggered by discussions between US BPO subcommittee leaders and IO personnel.

Specifically, we make the following recommendations in this area:

1. Research planning should be developed on two different time scales. Long-term (~5 years) planning should involve developing high-level strategic goals, while short-term (1 to 2 years) planning should focus more on the experimental campaign specifics.
2. Long-term goals are best addressed through “Topical Groups” (TGs) that are based on general research areas, while short-term goals should be targeted by “Task Forces” (TFs).
3. Development of experimental proposals should begin with an open planning process and culminate with specific proposals being written, reviewed and allocated run time within the TF/TG framework.
4. Allocation of run time should be guided by experimental priorities.
5. After experiment execution, the data management system must provide consistent, complete, up-to-date and timely views of the raw and analyzed data.
6. Infrastructure needs to be in place that supports data analysis requests, and that documents data provenance through the entire data acquisition, analysis, review and dissemination chain.
7. The initial determination of analysis and publication “rights” (i.e., a publication plan) should be made at the time of the experimental proposal, and should ensure adequate representation among scientists directly involved in the experiment. However, the Experiment Team should be open to analysis results of those outside the Experiment Team, and should be flexible enough to incorporate these into the analysis/publication plan.
8. Periodic assessments of program progress are necessary for long-term success. Such assessments would involve reviewers both inside and outside ITER.

Moreover, a set of general principles for program planning is proposed, drawing on the experience of research conduct in major US facilities:

- Decision-making should emphasize broad participation, openness and traceability.
- Access to experimental data should be open to all members of the world ITER team.
- Policies should be promulgated that ensure efforts by team members are rewarded with recognition, e.g. in priority for publication or conference presentation.
- Opportunity should exist for full group participation in the review of experimental proposals, presentations and publications.
- Opportunities should exist for graduate student participation in diagnostic development, experiment development and execution, and results analysis.
- Flexibility should be built into the Program structure to adapt to changing priorities and to pursue new findings.

Chapter 2: U.S. Team formation and participation in ITER

The set of questions considered in this section were: A) What is the process by which the ITER Team should be formed? B) How should the ITER Team be managed?, and C) How should the research part of the program be accomplished, including data storage and access issues?

A. ITER Team Formation

A successful long-term collaboration by the U.S. on ITER will require a skill mix that includes plasma, materials, fusion technology and control physics expertise, as well as expertise in engineering. It is envisioned that the U.S. ITER Team will consist of core scientists and engineers, who reside on site in Cadarache and who can become part of the ITER Team leadership, as well as those who spend varying amounts of time on-site but who are based in the U.S. and have other U.S. responsibilities. Our overall recommendations to facilitate ITER Team participation are:

- All participants from the U.S. should be considered as part of the ITER Team and party to and with rights from the ITER implementing agreement.
- All participants, including diagnosticians, machine operators, etc., should be able to join ITER Task Forces and get involved in proposal writing and analysis.
- A strong cadre of post-docs that will serve to establish scientific continuity will be essential to the long-term health of the U.S. and ITER Projects.
- Long-term participation of engineers should be enabled where appropriate.
- An approach to intellectual property rights that is consistent with the U.S. government regulations and which is acceptable to all parties should be developed well before the start of the ITER research program.
- There is a need for a thorough assessment within the ITER project, in discussion with DAs, of the number of people and skill mix needed to operate ITER. This will aid in resource planning for all partners.

It is expected that the U.S. Team will consist of members from national labs, universities and industry, and there is a need for recognizing the different resource and potential participation constraints of each type of organization. Furthermore, it is essential to create the mechanism for a transparent and fair process for members from all three types of organizations to become members of the U.S. ITER Team.

The appropriate size of the ITER Team can be resolved only by a "bottoms-up" assessment of the operational needs of ITER, including the need for scientific, engineering, technical, software, control and diagnostic expertise. Once this is decided, FES should ultimately be responsible for the selection and funding of the U.S. research proposals through peer review, with consideration of how well aligned the proposals are with the research goals of both the U.S. and ITER. Contingency funding should be kept in reserve to address research needs not previously identified, and a process for addressing these needs in a timely and efficient manner needs to be developed.

B. ITER Team Management

A single domestic entity should provide oversight of the U.S. ITER participation for both scientific and technical tasks. The role of this entity would be to:

- Coordinate planning activities with the ITER IO including development of research thrusts within the various Task Forces.
- Disseminate these goals and thrusts to the U.S. fusion community and facilitate U.S. team building for participation in the Task Forces.
- Mediate disputes within the U.S. ITER Team to ensure that research interests are represented, most notably with the smaller groups.
- Review progress, level of participation and deliverables, and make appropriate funding and personnel recommendations to FES.

There should be an on-site coordination of the U.S. ITER participation to ensure that U.S. interests are adequately represented and managed. This includes both physics and technical participation.

C. Execution of research program and data access/storage issues

Since not all U.S. ITER Team members will be residing on-site in Cadarache, effective mechanisms for remote participation and timely access to data will be critical for the success of U.S. research on ITER. They will allow ITER Team members to work effectively no matter where they are located. The overarching guiding principles are that:

- All raw and processed data should be made available to all members of the ITER Team.
- The computer and communications architecture should support all modes of participation to the fullest extent possible.

The following specific recommendations are made regarding data access:

- The U.S. government and community should develop a consistent position supporting remote data access and data sharing on ITER.
- The U.S. should develop an architecture for distributed data caching.
- The U.S. should develop a position supporting remote control of some ITER functions, particularly diagnostics.
- The U.S. IPO should ensure adequate computing resources at the Cadarache site for analysis of ITER data by U.S. researchers.
- The U.S. should ensure consistency between rules and procedures for data/code sharing and access that are emerging from the U.S. government and the ITER organization.

The remainder of this document expands on our committee's perspectives on how the major activities should be carried out. We list, where appropriate, issues that still need to be resolved. A glossary of roles we will be using throughout this document is given in Appendix A.

Chapter 1: Recommendations for ITER Research Program Planning and Experiment Execution

0. Introduction

Discussions within the BPO group and with the ITER Organization (IO) were conducted on ITER Experimental and Program Planning. In this chapter, the resulting recommendations for modes of operation and analysis in ITER are presented, as are strategic recommendations for ITER experimental operation procedures and support systems by the IO. This area is under the responsibility of the ITER Directorate of Plasma Operation, and the work was triggered by discussions between US BPO subcommittee leaders and IO personnel. This document summarizes the view and subsequent recommendations of the US BPO subcommittee, and may be used by the IO in finalizing the ITER experimentation strategy, but does not necessarily reflect the view of the ITER Organization.

To support this work, documents previously provided to the IO included data analysis workflows for specific examples of high-level between-shot analyses, one each from C-Mod, DIII-D, and NSTX, and a description of the experimentation procedure workflow for each of the C-Mod, DIII-D, and NSTX devices [1]. We extrapolate from that work to make recommendations for ITER Research Program Planning and Experiment Execution, taking into account the expected constraints on ITER operation. In making these recommendations, we propose a set of general principles:

- Decision-making should emphasize broad participation, openness and traceability
- Access to experimental data should be as open as possible, with a broad definition of the ITER Team
- Policies should be promulgated that ensure efforts by team members are rewarded with recognition, e.g. in priority for publication or conference presentation
- Opportunity should exist for full group participation in the review of experimental proposals, presentations and publications
- Opportunities should exist for graduate student participation in diagnostic development, experiment development and execution, and results analysis
- Flexibility should be built into the Program structure to adapt to changing priorities and to pursue new findings

Table 1 provides an overview summary of our recommendations, which are presented below. Column 1 represents the major activities of ITER experimentation, which includes everything involved in planning and executing the research program. Column 2 lists the major decision points associated with each activity. Column 3 identifies the major parties ultimately responsible for making or coordinating the making of those decisions. The activities listed are not in strict chronological order. Information systems will be needed to support and document these decision points in a set of integrated views.

The remainder of this document expands on our committee's perspectives on how the Major Activities should be carried out. We list, where appropriate, issues that still need to be resolved. The order of this discussion will follow that of the topics in Table I. Also given, in Appendix A, is a glossary of roles we will be using throughout this document.

DISCLAIMER: We note that this document represents the consensus approach advocated by the U.S. Burning Plasma Organization, and it is not an official position of the U.S. ITER Project Office or the U.S. Department of Energy.

1. Long- and Short-Term Planning

We recommend separate but overlapping activities for planning to occur on two different timescales: long-term (on the order of 5 years) and short-term (1 to 2 years). Planning activities on both time scales will form the basis for the ITER experimental campaigns, which are envisioned to be on the order of 1-2 years as described in the draft ITER Research Plan [2]. We use the word 'campaign' in this document to mean a period of approximately continuous experimental operation for which experiments are scheduled and planned in some detail.

Long-term planning will involve developing the high-level focus and strategic goals for the ITER Physics and Technology Programs. This planning should be consistent with the current or latest version of the ITER Research Plan document [2]. It will also develop rationales and timelines for facility enhancements and evolution. These high-level research and development goals should ultimately be decided upon by the 'ITER management' (the group empowered to make high-level project-wide decisions - see Appendix A), but in the spirit of broad participation, should be done in an iterative fashion with consultation and recommendations from the individual Domestic sponsoring Agencies (DAs), and including input from the full ITER Team. Shorter-term priorities, which will support the long-term goals and which will more directly impact the specific experimental campaigns and Team structure that will support the campaigns, should be developed similarly, but with more emphasis on input from the Research Team members, those who will be carrying out the experiments and analysis.

In this document, we use the term "ITER management" to refer to project level governance structures and leadership. The nature and structure of that management during the exploitation phase will need to be determined. We anticipate that it will include distinct entities to carry out both operational and oversight functions but have no detailed recommendations at this time for that structure. We believe that the governance structure that is developed should be based on a set of carefully thought-out and well-documented principles.

The ITER Team will include all researchers participating in the ITER project. Each DA should identify members of the Team. We advocate that the DAs consider all interested and qualified fusion researchers to be eligible, without a set number of members per party. This should include Team members who receive full-time support for ITER research, as well as part-time researchers who will receive partial support from the ITER project as well as other projects. We consider it very important to include university faculty, scientific staff, post-doctoral researchers and graduate students in ITER research.

ITER research should be organized in a manner that best supports both the long- and short-term goals. For long-term goals, we recommend a Topical Group (TG) structure that is based on more general research areas (e.g., Transport and Turbulence, Edge Physics, etc.). The TGs will develop strategies to explore the physics of certain aspects of the plasma or accomplishing a more specific goal. These groups should exist for at least as long as the long-term planning period. More specific, short-term goals should be targeted by Task Forces (TFs), whose focus is to achieve those goals within the designated period. An example TF might be "Qualification of Neutron Measurements." Taken collectively, the TG membership should comprise the entire ITER Team. TF members should be drawn from the appropriate TGs as needed.

Leaders of both the longer-lived Topical Groups and the shorter-term Task Forces should be chosen by ITER management with recommendations from and in consultation with the DAs. The criteria for selecting leadership of the TGs and TFs should emphasize scientific expertise and demonstrated leadership skills, and, especially for the TGs, should rotate. The TG and TF leaders should be part of the group ultimately responsible for defining the long- and short-term research objectives and priorities.

2. Experiment Campaign Planning

The development and execution of experiments should be accomplished within either the TGs or TFs. For each experimental campaign, the research priorities and high priority areas of specific experimental focus should reflect the long- and short-term research priorities, which are decided upon by project management (ITER management, TG and TF leaders and representatives from the DAs) with input from the full ITER Team. Other areas of experimental focus can be determined through discussions within the TF and TGs.

The overall operations schedule for each campaign should be made by ITER management. This includes number of run weeks and machine operation period, scheduled maintenance weeks, etc., based on machine and supporting technology availability and capability. This information and information regarding the current capabilities and limits of the experimental facility should be widely disseminated at the earliest possible time to aid in specific experiment planning. Specific recommendations or examples include posting current operational status to the ITER public website, and the eventual development of an ITER operational calendar, to which Team members may subscribe.

Our recommendations for allocation of actual experimental run time are given below. We note that Steps 1 and 2 should be an iterative process.

1. ITER management makes an initial run time allocation (i.e., number of weeks) to each TG and TF based on research priorities.
2. Topical Groups and Task Forces define experimental priorities and propose run time priorities to address the research priorities from discussions within their respective groups using the initial run time allocation as guidance. The plans

developed within the TGs and TFs regarding relative priority and necessary run time for experiments will be presented to ITER management, who may then modify the initial allocation.

3. Ultimately, allocation of specific run time to approved experiments will be made within the Topical Groups and Task Forces

Following current practice, we recommend that a significant amount of run time (~20%) in each experimental campaign be initially withheld. This portion of the available run time should be reserved for contingency, to be allocated later in the campaign based on unmet experiment needs (e.g., due to hardware failures on the day of an experiment) and evidence of promising results seen during the experiment. Contingency time during each week of operation should also be assigned. As far as possible, the operations and maintenance schedules should be maintained so that off-site Team members can make travel plans.

3. Development, Review, and Scheduling of Experimental Proposals

a) Planning via a Research Forum

Selection of experiments to be executed during a campaign should begin with an open planning process that occurs well in advance of the campaign. Idea generation is best accomplished using an open, science-focused approach, in which any member of the ITER Team or community-at-large is free to propose ideas for experiments. Solicitation of specific types of proposals or proposals in specified areas should be encouraged as a method to get experiments that support program goals. To this end, we recommend a Research Forum in which specific experimental proposals will be presented. All experimental proposal information should be archived and available online.

Carrying out such a broad-based Research Forum in the international context of the full ITER team will be challenging. Ideally, the Research Forum should be accessible in real-time to the entire community through on-site or remote participation. However, since the community spans time zones around the world, this may be difficult to accomplish. It might be preferable to have multiple sessions such that each national team has access to at least one real-time interaction with onsite personnel. In addition, the number of experimental proposals submitted during this process is likely to be very large, based on the anticipated large size of the ITER team and the large numbers of proposals typically submitted even by the much smaller teams associated with present devices. Some compromises will need to be made for practicality. Some type of pre-filtering of experimental ideas will likely be required to reduce to a manageable number the experimental proposals requiring discussion in an open forum and later evaluation for assignment of experimental time. This process should validate the need for this research to be done on ITER as opposed to alternative devices. A standard format for submission should be developed that simplifies the filtering process. Video recording of presentations should be used to provide a practical alternative for ITER team members unable to attend remotely.

There must be a process for consolidating ideas developed for the Research Forum and promoting those with scientific merit and programmatic importance. Since experience has shown that separate individuals can have similar ideas or similar requirements for experiments, this process can also be used to combine separately proposed experimental ideas, to build on other experiments or determine which experiments can be effectively executed together with another experiment.

We recommend that TG and TF leaders perform this consolidation function using a process that ensures transparency and, ideally, participation by all individuals/parties who have a stake in the results ("stakeholders"). Since combined experiments would have multiple authors, the TG and TF leaders should also identify an "Experiment Leader" with primary responsibility for experimental proposal writing and creating the team that is responsible for executing the experiment, ensuring adequate representation among participants. While consideration should be given to having equitable representation of Experiment Leaders among the various parties, the primary criterion for selecting experiments, and the Experiment Leader to write the detailed proposal, should be the scientific quality of the idea presented.

Here we note a couple of assumptions about the process for ITER pulse programming:

- An ITER experiment will consist of an assigned number of time-limited segments within one or more ITER pulses. We use the term "segment schedule" to mean the requested plasma and system behavior in one of those segments as specified by time-dependent reference signals and parameters.
- We assume that before segment schedules for an approved experiment can be scheduled, those segments must be combined with segments from other approved experiments and the resulting full pulse schedule run through ITER validation.

b) Prioritization

Allocation of run time and scheduling of experiments should be guided by clear definitions of experimental priorities, which need to be provided early in the planning process by ITER management with input from the TG and TF leaders. These priorities should be considered as one of several metrics for approval and prioritization of proposals, and ultimate allocation of run time. A proposed set of criteria is listed below:

- Relevance to research goals and program milestones
- Scientific value and motivation
- Technical feasibility and likelihood for success (in particular, must be executable with current capabilities without jeopardizing the facility)
- Potential for developing new capabilities and/or operational regimes, i.e. scientific novelty, balanced against assessment of facility risk
- Fair representation among all partners

We recognize that the balance of prioritization criteria will shift in different phases of the project. For example, facility milestones may largely drive the program during commissioning of new capabilities, and scientific goals be more important in other campaigns.

c) Preparation and Review

A detailed experimental proposal must be written by the Experiment Leader, with input from that experiment's team. Proposal content should be standardized to include at a minimum the following information:

- Required actuators, and their settings
- Required diagnostics, and their settings
- Preliminary list of team members
- Discharge segment schedules and requirements
- Detailed plan and sequence of execution of those segments
- Expected between-shot analyses and software needed to support it
- Tentative analysis and publication plan

The Experiment Leader should be responsible for ensuring that the necessary preliminary segment schedules are prepared, submitted, and reviewed; the actual task will be performed by a qualified "Segment Developer" (see Appendix A).

The draft experiment proposal should be posted on the web and reviewed initially by the members of the relevant TG or TF, with a subset of members with appropriate expertise chosen to be a "review committee", by device operators and by any interested Team Member. This review process should be an open one, with interaction between the proposers and reviewers. An offsite Team Member should have the ability to document comments and concerns on the draft proposal through the Web or through remote connections to the initial review. These concerns should be addressed by the proposers. The review process is expected to strengthen the experimental proposals and ensure that the experiment can be run with available device capabilities and without risks to the integrity of the device.

Once the input from the initial TG or TF reviews has been incorporated, the now near-final proposal should be submitted to an on-line database. This database should connect the detailed proposal with all other data related to this experiment to be generated later. The segment schedules should be run through ITER validation software and the results also stored in the online database.[1] This validation software is expected to be composed of a suite of validated codes developed within the present fusion community that will be implemented and run on the IO computer system. This software will enable development of discharge scenarios and expected performance and will ensure the integrity and safety of the full plasma pulse.

The detailed proposal should then be evaluated and approved by an Operations Management group[2], after which the segment schedules can be combined with segment

---

[1] A large number of proposals, all requiring validated segment schedules, would require a large amount of resources. Therefore, this should be done only for the "pre-filtered" experimental proposals that are given high priority in the respective TGs and TFs.

[2] Exact membership to be determined, but should include TF/TG leaders, and operators responsible for responsible for device, facility, and physics.

schedules from other approved experiments to form full pulse schedules; this role is termed "Physics Operator". When the resulting full pulse schedules are validated, they can then be scheduled for execution.

d) Scheduling

After experimental proposals are approved and prioritized, specific run time allocations should be made by the TG and TF leaders, in consultation with Experiment Leaders. Then with time allocated, each experiment must be scheduled for a time period that is consistent with the constraints of the facility and availability of key personnel involved in the experiment. To optimize the utilization of the machine when available, we suggest the development of backup or alternative experiment teams in case of personnel availability or equipment complications. This would ensure a suitable team is always ready to take advantage of scheduled run time.

4. Experiment preparation and execution

Preparations for executing the experiment include assigning lead operational responsibilities, e.g. for diagnostics coordinator, analysis coordinator, etc., and also for facility specific tasks, e.g. Physics Operator and Lead Engineer (see Appendix 1A). The pulse segment specific responsibilities should be assigned by the Experiment Leader, while the facility specific roles should be assigned by Operations Managers. Run-day assignments of personnel for other experiment support functions such as diagnostic coverage and run-time analysis should requested and verified by the Experiment Leader. Operations managers in conjunction with Experiment Leaders should review any issues not explicit in experimental proposal. This should be done two months prior to the scheduled time for the experiment and re-reviewed several days prior to the actual execution of the experiment.

The Experiment Leader ideally should be on-site for a defined period of time before, during and after the experiment is scheduled to run. In extenuating circumstances, when the Experiment Leader cannot be on-site, a representative appointed by the Experiment Leader, with approval by the TG or TF leader, should run the experiment. The on-site team should make shot plan option or recommendations for modification, with input from remote participants as practical. To ensure that on-site presence for the session leader does not overly constrain grants or contracts, we recommend the formation of a modest mobility fund to separately support travel for Experiment Leaders.

Because we anticipate that there will be multiple Experiment Leaders for a single pulse, an extra person ("Session Leader") will be needed to interface between Experiment Leaders and the Engineering Operator, who has the responsibility to load pulses into the ITER plasma control system. Ideally, the Session Leader should have personally validated the entire pulses, or have reviewed the validation done by a Physics Operator.

The concept of multiple experimental segments with likely differing goals underscores the need for pulse-wide coordination of diagnostics and other systems. While most diagnostics are required to be available on every shot, many have settings that must be selected on a shot-to-shot or even run-to-run basis, for example selection of frequencies, wavelengths or viewing chords. Similarly, heating and other auxiliary systems will have settings that may not be quickly varied (eg, ECH antennas, ICRH frequency, NBI energy). Compatibility of settings across pulse segments could become a key consideration in determining which segments can be combined to form pulses.

In present devices, it is typical to have some deviation from a written plan due to device conditions and results of prior discharges. We recommend allowing the Experiment Leader the flexibility to modify segment schedules used during a particular pulse based on results from prior pulses. The type of modification allowed between pulses should depend on the level of risk engendered, which roughly correlates with the energy content of the plasma and coils system. Changes to segments in "high-energy" pulses should be limited to a pre-validated envelope of values programmed into the segment schedule. Changes to segments in "low-energy" pulses should be far more flexible in order to support, for example, the processes of plasma startup and control development. Metrics need to be developed to evaluate the level of risk associated with a segment change to inform policies regarding what changes are allowed between pulses. Procedures to enforce these policies and thereby minimize risk to the device should be put in place. Any modifications to the segment schedules should have been approved by the Operations Management group defined above, and conveyed to the Engineering Operator via the Session Leader. Some clarification of any nuclear regulations that affect experiment design should be provided as a reference for this planning process.

To support this kind of between-pulse decision-making, and to generally understand the performance of the pulses, tools for timely access, analysis and display of selected diagnostics and derived physics quantities must be provided. Requirements will vary depending on the goals of the experiment. Since some participants will be off-site, remote access to the selected subset of data must also be sufficiently fast to allow between-pulse visualization. An “events” system that allows users to know when raw and processed data is available should be implemented. Timely analysis may require reserved cycles on High-Performance Computing (HPC) platforms co-located with the data archive. Between shot analysis results should be archived with the raw data. We note that while the time between ITER pulses is not yet known, all parties can expect in principle to contribute to next pulse decisions, with the degree of decision-making affected by the cycle time between pulses.

We recommend that ITER provide an online method for making comments about the results of individual segments shortly before and after execution (“Electronic Logbook”). Key information would include basic discharge and segment details, how well the discharge ran, any missing data, any interesting observations or analysis results, etc. This capability can be used to support the decision making for the limited adaptability described above and for later analyses of an experiment, when it is valuable for identifying key experiment segments to analyze. This online system should also provide a means of entering one or more written summaries of the experiment, e.g., its objectives and results, for later reference. Such summaries should be required of all experiments within a very short time after experiment completion. All of the operations and Experiment Leaders/sub-leaders are expected to enter comments as necessary. All members of the ITER Team should be able to read the comments and add their own comments and annotation. We note that current Logbook implementations have the comments tagged with the commenter's role, name, date and time.

Short presentations to the experimental team are expected shortly after the day of the experiment, perhaps at a Weekly Team meeting that is remotely accessible. Given that time differences may prevent some interested team members from participating regularly, we recommend that these presentations are also made available online. This should ideally include video recordings to follow the discussions, and a forum to send other comments or questions. These capabilities will be generally useful for ITER team communication.

5. Experimental Data Analysis and Results Dissemination

After the experiment is completed, the process of analyzing the experimental data begins. Under the overall direction of the Experiment Leader, it is envisioned that different researchers will analyze different portions or characteristics of the experimental data, based on their area of expertise. Since the results of these analyses will typically be shared and combined with data from other members of the experiment/analysis team, methods for data sharing (storage and access) for intermediate analysis results must be provided along with the final results. These methods must be flexible enough to allow for archiving of varying quantities of analyzed data, some of which may have size or format that has never been archived before. Since analysis methods, calibrations or estimation of errors may evolve over time, the archive should support versioning for analyzed data. Details on modifications to data must be documented and readily available to all Team Members.

Although preliminary experiment analysis can be performed rather quickly or be automated, more extensive data analysis may continue nearly indefinitely. The data management system must support this process by providing consistent, complete and up-to-date views of data at all stages of analysis. The Experiment Leader (or a designated analysis coordinator) would submit requests for data analysis tasks and data quality review, which would be carried out by diagnosticians and data analysts. The data analysis and review should be performed in a timely fashion, with reasonable dates for providing validated data that are agreed and adhered to. The results of all data analysis or review should be open and readily accessible to Team Members whether they are on-site or remote, though the use of the data for publication, presentation or sharing with third parties would be subject to guidelines discussed below. To support all these requirements, infrastructure will need to be provided that supports data analysis requests and that documents data provenance through the entire data acquisition, analysis, review and dissemination chain.

The initial determination of analysis and publication "rights" should be made at time of experimental proposal. Given the large size and distributed nature of the ITER team, a more formal method than is used on present devices will be needed for planning and managing scientific publications. Beginning at around the time that proposals are approved for run time, a publication plan should be initiated and tracked as the experiments and subsequent analyses progress, with modifications made to the analysis/publication plan as necessary. This plan should define an allocation of expected publications and presentations, ensuring adequate representation among scientists directly participating in the preparation, execution, and data analysis of individual experiments. Topics that are clearly within the scope of a TG or TF can be managed by the leaders of those groups, while cross-cutting issues will need coordination between those leaders and ITER management.

It is expected that the Experiment Leader (or his/her designee who was deeply involved in the experiment) has first priority on the major results of the experiment, but others who participate or support an experiment are expected to write papers describing certain details of the experiment and their analysis. For conference presentations and publications to journals, all who have contributed in a significant way to the results should be able to lead publications or be included as co-authors on publications and given ample opportunity for input. Co-authorship and the ordering of authors on publications will be governed by an explicit set of guidelines put in place by ITER Management that has been agreed upon by all ITER parties. This agreement can be modified and updated as detailed analysis reveals new behavior that may warrant additional publication(s). Disagreements or grievances over publication/conference presentation authorship should be mediated by the appropriate TF/TG leader, with input from ITER Management, if necessary. In addition, the Experiment Team should be open to analysis results of those outside the team, and should be flexible enough to incorporate these into the analysis/publication plans. It is noted that some of the data analysis will be directed by the needs of a publication, and some of it is performed independent of expected publications.

The actual process of preparation of a publication or presentation should be organized by the assigned first author. In particular, support may be needed for additional analysis, preparation of graphics, and writing of individual subsections.

All publications and presentations should be subject to formal review and comment before any submission for publication or public presentation. The first step in the review process will be by co-authors, but participation in these reviews by other ITER team members should be encouraged to facilitate coordination of publications and to obtain useful feedback. Draft publications should be posted electronically for group review, and presentation of oral dry runs should be remotely broadcast to all interested Team Members. Posters should be posted electronically for Team review. The capability for off-site Team Members to provide feedback electronically should be provided. Any disputes or substantive issues about the paper or the presentation that are not resolved among the authors or the Team members shall be mediated by the appropriate TF/TG leader, with input from ITER Management as necessary, in order to arrive at a mutually agreeable resolution prior to submission and/or external use. Conduct of this process in a timely manner is important, and so a time frame for this process should be agreed upon by the ITER management in consultation with the DAs.

ITER Participation and Open Data Requirement (see also Chapter 2, section c)

Data access policy for ITER has not yet but developed, but the potential for conflict with the expected U.S. DOE guidance deserves prompt attention. As of 2014, U.S. researchers must provide public and free access to data displayed in publications. The U.S. policy, which will most likely evolve from this starting point, will be in line with an increasing interest from the U.S. government in "open access" to data and publications that arise from taxpayer supported research (see 2C). Other countries could choose to follow this route also; e.g. U.K. legislation targeted similar policies in 2013, while Germany has recently adopted policies for open access to certain publications. Given these worldwide trends, our partners may agree to compatible policies for ITER data, but this cannot be assumed. Thus there is a possibility of a fundamental conflict. It is worth noting that data access policies from existing fusion experiments - including those in the U.S. have, historically, not been compatible with the open access requirement. For the U.S. this obviously will have to change. It is also apparent that ITER's current approach is to release information only on a limited and "need to know" basis. So the possibility of a policy conflict is not ignorable. On the other hand, we note that the ITER implementing agreement seems to be consistent with a more open policy. Quoting from the annex on information and intellectual property.

- Article 2.1 states "Subject to the provisions of this Annex, the members support the widest possible dissemination of generated intellectual property".
- Article 3 states "Each member shall be entitled, for non commercial uses, to translate, reproduce and publicly distribute information directly arising from the execution of this agreement..."

How these general statements in the agreement are interpreted and translated into policy will be important. Past experience suggests that this is not as straightforward as it might seem.

Our recommendation is that DOE bring up this issue and come to a resolution with the ITER partners as soon as practical. Failure to reach agreement could also compromise U.S. participation in ongoing ITPA activities.

6. Program Review

Periodic assessment of program progress is a key to long-term success. Such assessment would involve reviewers both inside and outside the ITER team. Specifically:

It is recommended to have a mid-campaign Run Assessment to review results to-date, and to identify research gaps and opportunities for additional research during that campaign. ITER management, along with the TG and TF leaders, should, after the mid-run review, decide on any reprioritization of experiments, reallocation of run time, and distribution of contingency. Discussions within the TGs and TFs on these topics should take place prior to the mid-run review. For the anticipated campaign durations of 1-2 years [2], a single mid-run assessment should be sufficient. If campaigns extend beyond two years, progress may need to be re-assessed on a roughly annual basis.

The high-level strategic goals and progress of the overall ITER program should be reviewed periodically to evaluate their appropriateness for, and the progress toward, achieving the defined mission. Additionally, the review should assess whether the Team structure (e.g., TGs and TFs) is appropriate for addressing the research goals. This review should be coordinated by ITER management, with input from all ITER participating countries and outside review teams.

Table 1. Activities and decisions involved in planning and executing the ITER research program.

Activity	Decision Points	ITER Process or Responsibility
Long Term Planning (5 years)	Define high level strategic focus and strategic goals including facility evolution (equivalent to US 5 year planning)	ITER management (see Appendix 1A for definition)
	Define set of long-lived topical groups to support strategic goals	ITER management
	Define topical groups leadership and membership	ITER management
Short Term Planning (1-2 years)	Define facility-wide one or two-year objectives	ITER management with input from TG
	Define short-lived task forces and task force objectives as needed	ITER management
	Define task force membership and leadership	ITER management
	Define one or two-year objectives for each topical group	Topical groups with review by ITER management
Campaign Planning	Define overall operations schedule (how much run time and when) . Define facilities capabilities and limits	ITER management
	Develop task force and topical group priorities and plans, identify program gaps, propose run time priorities.	TG/TF leaders
	Review task force and topical group priorities and plans	ITER management
	Set high level priorities and allocation of run time to TGs and TFs	ITER management
Develop Experimental	Identify possible experiments	Research forum
	Consolidate/combine experiments with common	Topical group and task

Proposals	requirements	force leaders
	Assign experimental proposal writing, insuring adequate representation among parties	Topical group and task force leaders
	Create experimental team, insuring adequate representation among participants, and select Experiment Leader	Topical group and task force leaders
	Prepare, submit and review preliminary pulse segment schedules	Segment Developer, supervised by Experiment Leader
	Review experimental proposals	ITER management, topical group and task force leaders
Run scheduling	Define detailed allocation of run time	ITER management with input from topical group and task force leaders
	Set detailed run schedule	Operations Managers with input from Experiment Leaders
Experiment Preparation and Execution	Assign operational roles (Diagnostics Coordinator, Analysis Coordinator…) for each set of pulses/experiments	Experiment Leader
	Assign Physics Operator	Operations Management
	Submit and review final pulse segment schedules	Segment Developer
	Combine pulse segments from different experiments and validate entire pulse program; submit and review final entire pulse schedules	Physics Operator
	Define run-day assignments for diagnostic coverage, run-time analysis, etc.	Experiment Leader

	Review any detailed plans not explicit in experimental proposal	Operations Managers and Session Leader with Experiment Leaders
	Make run-time decisions on shots (choose from menu of approved pulse schedules)	Experiment Leaders via Session Leaders
Experimental Data Analysis	Submit requests for data analysis and review	Experiment Leader
	Assign data quality measures	Diagnostician/analyst
Results Dissemination	Identify possible publications or presentations through author-submitted intent to publish)	Any Team Member
	Propose allocation of publications and presentations, insuring adequate and fair representation among participants. Resolve conflicts among potential authors.	TG/TF & session leaders with oversight by ITER management *
	Assign leads for publications and presentations	Experiment Leaders with oversight by TG/TF leaders *
	Define organization for preparation of publication/presentation (who does needed analyses, preparation of graphics, subsection writing	First authors
	Review publications and presentations	Experiment Leaders, TG/TF leaders with oversight by ITER management*
Program Review	Conduct mid-run or other periodic progress assessments	ITER management
	Determine whether reasonable and adequate progress towards strategic goals and ITER Mission being achieved. Recommend course corrections or facility upgrades.	ITER management, DAs, outside review teams

* **Issue:** Responsibility to be determined

References:

[1] U.S. Burning Plasma Organization working group: Modes of Participation in ITER, "Sample Data Analysis Workflows from Alcator C-Mod, DIII-D, and NSTX", June 21, 2013
[2] "The ITER Research Plan", https://www.iter.org/doc/www/content/com/Lists/Stories/Attachments/1361/David2.pdf

Chapter 2: US team formation and participation in ITER

Discussions with our Fusion Energy Sciences (FES) leaders at the U.S. DoE identified a set of questions on which FES requested input from the BPO. These questions were:

A. *What is the process by which the US ITER Team should be formed?*
B. *How should the US ITER team be managed?*
C. *How should the research part of the program be accomplished, including data access and storage issues?*

Each of these questions is addressed in the subsections below. As one might expect, there are topical connections between this chapter and those in the first chapter of this document; such connections are identified where appropriate.

2A. ITER Team Formation Process

A successful long-term scientific collaboration by the US on ITER will require a team of scientists with a range of skills from a variety of US institutions. It is envisioned that members of the US-ITER Team will fall into a few broad classifications of scientists and engineers to most effectively plan and perform experiments, analyze data, run simulations, support experimental operation, operate diagnostics and hardware systems, participate in educational activities, and to ultimately build the knowledge base to develop fusion energy systems. The required scientific skills mix will go beyond traditional fusion physicists, who have specialization in plasma physics. For instance, surface physicists/chemists, materials physicists, and control/feedback specialists will all be needed for the U.S. to fully contribute to and benefit from the scientific advances via ITER. Thus the term scientist in this section represents participants with specializations in one or more of these areas. With this in mind, we envision the following classes of US-ITER scientists; most of these would be selected via peer-reviewed proposals:

1) Core scientific team: Staff who are largely based on-site at ITER-Cadarache, and working mostly or entirely on ITER (80-100% time per FTE.) This category includes physicists, engineers, and technicians. Terms would be linked to renewable proposal cycles, which are usually of three to five year duration. These participants are likely to have leadership roles in the ITER scientific program coordination (e.g., Task Force or Topical Group leaders), perform a core role in TF/TG mission (e.g., develop and maintain core competencies in analytical skills, codes, and mentorship of post-doctoral researchers), perform physics operator roles, provide operational support (software, plasma control, specialized testing, etc.) for US experiments, and/or operate key diagnostic systems. Special attention is needed to ensure mentorship in areas outside traditional plasma physics, where on-site expertise may be limited. Most would be employed by a US entity (e.g., national laboratory, university, or scientific research company.) A few longer-term US scientific staff might be ITER IO employees (pending party negotiations for supporting ITER during the operations phase.) Responsibilities of the core team would include managerial coordination of US-ITER participation to ensure that US-based interests (DOE-FES, BPO, institutes) are adequately represented and managed.

2) Cyclical scientific staff: Participants who devote a significant fraction of their time to ITER research in a given task group, but may also have other US responsibilities, such as working at domestic fusion research facilities, educational institutions, other international collaborations, non-ITER computational work, etc. These researchers would bring a set of skills to the particular Task Force or other medium-term effort. They would spend some time at Cadarache on a short-term basis, but likely not be based there. They could participate remotely for planning and preparing for experiments, discussion of results, etc. Terms would be linked to scientific proposals targeted to particular projects, experimental campaigns or analysis activities. Proposal and funding would be similar to that for current university collaborative research projects, which are usually correlated to renewable proposals of three or more years duration.

3) Short-term/task-specific participants: Researchers who propose a specific ITER task (e.g, contributing to or running an experiment, analysis tasks, running codes) either individually or as a group. FES funding would be on a proposal driven basis. May be appropriate for professors to engage in summer or sabbatical research projects. Duration and fractional time expected to be 0.5-2 years at 10-100%.

4) Graduate students and post-doctoral researchers: Early career participants for 2-3 years for thesis or well-defined research projects appropriate for young scientists, likely spending at least part time on-site. Full and part-time staff would ensure that research projects are commensurate with their institutional goals and missions, and provide a healthy research environment with publication and conference/workshop opportunities. Effective on-site mentorship of these participants will be critical for professional development and to ensure a meaningful contribution to the ITER program. As in the core scientific staff category, special attention is needed to ensure mentorship in areas outside traditional plasma physics, where on-site expertise may be limited. See additional note at end of this section for more discussion related to university participation (A1).

5) Operations staff: Participants located on-site for assignment durations of 5 years, renewable. The major hardware systems and diagnostics being provided by the US (or other party) will require teams of technical expertise, whether that is provided by the entity that designed and built the equipment or a new entity that is fully trained. Knowledge transfer is key to successful long-term safe and reliable operation for such systems. In addition, ITER provides a unique opportunity for US engineers to gain valuable experience in fusion power systems, experience that would be valuable in the design of FNSF and power producing devices. Maintaining a long-term, strong engineering and technology workforce (as opposed to contract/term engineering labor) is needed to capitalize on this opportunity and transfer expertise back to US institutions.

6) US-based support staff: Personnel that provide required technical expertise to maintain ready data access and information technology to all interested participants, including: capabilities for remote participation in meetings and experiments (audio, video, screen/data share, etc.) from both dedicated remote control room sites as well as various affiliated institutions as well as technical support for on-site engineering staff. As ITER operation is of a scale not experienced in fusion experiments, we anticipate that remote

participation may become a very effective use of resources, and perhaps the most effective means of large scale US participation, as experienced e.g. in other large scale science experiments (e.g. NIF, LHC, telescopes…). These personnel could be supported explicitly as members of the US-ITER research teams and/or via support for laboratory and university infrastructure. It is unclear whether this class of staff should be linked to proposals.

We feel that the process to be implemented for US team formation should have the following attributes:

(1) Transparency
(2) Inclusiveness
(3) Identification of best people for the team
(4) Continuity
(5) Timely and efficient means of joining a team for a short period

This approach would be most beneficial to the US effort and the ITER project as a whole. We also recommend timely identification of the appropriate international teams and prompt input and advice to the ITER teams (through STAC/MAC/ITPA) about the definition, content and priority of the research thrusts that will govern the US contribution to the research effort.

Additional Areas of Concern for US participation in ITER

In many cases, U.S. research institutions have found it difficult or impossible to participate in ITER tasks because of terms and conditions imposed in ITER contracts – even for work that can be clearly categorized as R&D, rather than fabrication or construction. Among the most contentious issues are those concerning intellectual property (IP) and publication rights. The IO has viewed entities other than the domestic agencies to be conducting “work-for-hire”, and has insisted on ownership of IP and absolute control over publication or dissemination of research results. Neither of these is consistent with current practice for research institutions funded by the U.S. government, nor are they consistent with policies at major universities. In fact, the ITER contract conditions would likely violate the tax-exempt status that many universities qualify for in the U.S., and could threaten critical exemptions from deemed-export rules provided for entities carrying out fundamental R&D. This “work-for-hire” paradigm also appears in contract conditions for payment – contingent upon acceptance of deliverables. While this may be standard for fabrication/construction vendors, it is inappropriate in agreements between research partners, and is contrary to accepted practice for U.S. institutions conducting government-funded research. These issues are serious problems now, having already prevented or delayed several US contributions to ITER preparations. They will become even more critical as we move into the research phase.

Increasingly large computer codes are the product of large team efforts. An important aspect of this is version and quality control. This may be especially important for ITER if the codes were used to validate a discharge scenario. Currently, ITER is developing an integrated modeling effort. How the US contributes to this effort and whether the US scientific teams retain control of our existing codes needs further discussion with the IO.

This issue cross-connects to the Chapter 1 discussion on pulse segment validation prior to execution of the experiments; it needs to be determined which sets of tools will be accepted for such validation, as part of the planning and execution of experiments.

Recommendations:

- Develop an approach to intellectual property rights, publication and other contracting issues that is consistent with current government regulations in the U.S. research community, and acceptable to both participating institutions and ITER. This issue needs to be addressed and resolved well before the start of the ITER research program.
- All members of participants from the US should be considered as part of the ITER team and party/rights to the ITER agreement, i.e. they should not be classified as contractors.
- Diagnosticians, machine operators, etc. should be able to join Task Forces and get involved in proposal writing or analysis.
- A strong cadre of postdocs will be essential to the continued health of the ITER project. Establishing and maintaining such a cadre requires close alignment of specific research goals and programs pursued by the postdocs with ITER goals.
- ITER research needs Long-term participation of engineers should be enabled where appropriate, to benefit from the valuable engineering and technology experiences in ITER.
- As discussed further in Section 2B, there will need to be a thorough assessment within the ITER project, in discussion with DAs, of the number of people, and mix of skills, which will be required to effectively operate ITER. This assessment will greatly affect the resources and planning for teams from all partners, and so this exercise should begin soon.

Finally an important element is the range of participation issues faced by the three largest constituents of the program, i.e. national labs, universities, and industry. By their nature, universities could face the most challenges in coupling effectively into ITER participation. Thus considerations specific to universities are highlighted below, followed by the specific considerations for labs and industry.

A1. University Participation:

Large scientific facilities such as ITER offer unique opportunities for education and workforce development - students are exposed to cutting edge technology, large teams of world leaders in a given field, and a work environment that fosters frequent interaction, collaboration, and scientific integrity. The mode of involvement for faculty and students should be designed to take maximal advantage of these opportunities while supporting the following goals:

- Contribute to the ITER mission
- Advance plasma science and fusion technology
- Develop a competent fusion workforce

Similar to present day magnetic fusion devices and other large scale experiments such as the National Ignition Facility (NIF) and Large Hadron Collider (LHC), it is expected that both faculty and graduate students will have the opportunity to participate in all stages of the ITER project, from initial construction to experimental planning, execution, data analysis and interpretation, as well as assume leadership roles where appropriate.

On ITER, the university community should be encouraged to participate in the development of diagnostic and other packages currently being designed and procured. Similar to LHC and also present fusion devices, we recommend adoption of a policy in which the ITER party that designed the diagnostic or other hardware should be allowed and even encouraged to play a significant role in the commissioning and operation of the instrument. Being involved with diagnostics or related hardware, diagnostic analysis tools, and subsequent upgrades allows some level of direct participation by both faculty and students and has been shown to be scientifically productive, as well as an effective means for integrating both into the broader team; indeed this hands on work will likely be essential during the commissioning phase to get diagnostics into a fully functional state.

As on present magnetic fusion devices, it is expected that both university faculty and student experimental proposals will be allowed through the ITER Research Forum in some manner and that these proposals will be judged according to the same criteria as proposals from any other member of the ITER Team. Similarly, both faculty and students will be considered as potential "Experimental Leaders" and judged by the same criteria as other ITER Team members. Likewise, their proposed publications should meet with the same review process as the broader Team. While many of the initial experiments will likely be focused on achieving specific high-profile ITER milestones led by senior researchers, vast amounts of data will be generated and it is expected that students will be allowed to contribute through assigned analysis tasks, application of a new analysis technique developed by the student or advisor, or simulation of a given experimental result. In order for the student workforce to take advantage of these opportunities, it will be essential for the bulk of ITER physics data to be broadly accessible by the Team, and for academic participants to have a clear path to membership in the Team. Particular consideration and prioritization should be given to resources required for students to complete a thesis dissertation.

An important aspect of optimally engaging the university community is to efficiently communicate what topics or work areas would be of the greatest benefit to ITER, and appropriate for student involvement. Some mechanism should be put in place for publicizing 'ITER topics of interest' - for example, analyzing in a specific way an existing or planned stream of data. In this way, potential topics may find natural fits with a particular university program. Before students or faculty begin long term study at ITER, or analysis of ITER data, a process should be in place for identifying the approximate goal of the work and vetting this topic with the ITER community or relevant Topical Group. The goal of the vetting process is to avoid work which will be of marginal interest or benefit to ITER and therefore will not get necessary resources, support for grant applications, etc. While most research on ITER will be mission-oriented, a small fraction of resources and data access should be provided for exploratory research on fundamental science topics that are not immediately central to the ITER Mission, but for which the ITER experiment may provide unique capabilities. This process will increase the attractiveness of ITER to the academic community and broader university environment and thereby increase scientific support for the project as well as potential sources of high quality graduate students.

A2. Potential Issues Associated with National Lab participation

Participation by scientists and engineers from the National Laboratories in ITER can be as Visiting Scientists or as "secondees" from their home institution or on a Leave of Absence as a member of the IO. As Visiting Scientists or "secondees", they retain their affiliation with their home institution. Similarly upon the completion of a Leave of Absence, they can resume their affiliation with their home institution. Currently, there are no national lab employees working on ITER in this capacity, i.e. national lab personnel have instead become ITER employees. The Visiting Scientist option would be beneficial for staff from the National Laboratories on long-term assignments to ITER, and would encourage the transfer of their experience back to their home institutions. As discussed above this is particularly critical for engineering staff.

A mechanism for direct participation in ITER projects is by responding to solicitations from the IO. National Labs fall in the Federally Funded R&D Center category; as such, they are not allowed to compete with domestic universities or industry, nor with any international entities. The only way for labs to respond to such solicitations is via a subcontract from a bidding entity, or if the IO chooses to enlist the lab as a sole source. This presently represents a major hurdle to direct National Lab participation.

The IP rights issues are not as problematic for Nationals Labs as other groups. Since the National Laboratories are under contract to the US DoE, the contractual issues associated with the ITER Legal Agreement such as IP are incorporated directly into their contracts.

A3. Potential Issues Associated with Industry Participation

There are a number of existing impediments to industrial participation in ITER that should be addressed and resolved if possible. Briefly, these issues relate to intellectual property (IP), labor rate disclosure, subtleties related to international partnerships, and strict French nuclear regulatory requirements.

With respect to the question of IP rights, ITER's position is that if work for hire is done, then they own the IP. Of particular concern is the fact that this could include background IP that was brought to the table prior to contract, which from industry's perspective, shall not be given away. This needs clarification by the US and IO.

Another ongoing concern is that of labor rate disclosure and the need for detailed breakdowns in some cases and not in others. This is not so much of an issue for academia and national laboratories, whose rates are a matter of public record but it is an issue for industry. This has been finessed in some cases where the rates are assumed to be those audited or verified by an independent accounting standards agency, yet future rates are deemed to be acceptable if they are verified in the future. Fees, which are a natural component of industrial contracts, make things more complicated.

A few issues arise when teaming with international partners is involved. One issue derives from a basic difference in the definition of a consortium in the US and Europe. In the former there is an assumption of shared liability, whereas in the latter the liability is typically not transferable to partners. This results in a US domestic industrial problem because of the poorly defined liability exposure. The consequence of this poor definition can result in legal debates over contractual terms and conditions; these typically take several months to a year for resolution. At present, the simplest method to enable partnering between a US company and an international party, with the US company as the primary, requires a series of small, well-defined contracts. This is problematic for the development of flexible task agreements. Tracking the work is seen as more difficult than the technical issues in some cases, with the need to build an organizational chart for each task element. One remedy to this bureaucratic inefficiency might be the development of an "accounting standards group" for foreign entities, so that task agreements could be simplified, in principle.

One apparent advantage to industrial participation is the presence of people who are well versed in and cognizant of the French nuclear regulatory agency rules and requirements. This is particularly relevant for hardware development and delivery. Similarly, ITER is applying industrial level standards and practices to their software development and will likely require the same for most contributions from parties. Although these standards and practices can certainly be learned, industry is most experienced with them.

2B. US ITER Team Management

We envision three levels of management and oversight for the U.S. ITER Team

1. Selection of U.S. personnel
2. Oversight of U.S. research and technical involvement
3. On-site management of U.S. ITER Team

1. Staffing by U.S. personnel

FES should be ultimately responsible for the selection and funding of the U.S. ITER research proposals in much the same fashion as they do now for domestic programs, through proposal reviews. Based on a definition of the U.S. role in ITER (see 2. Below), FES should solicit proposals for funding these teams. As is the case now, peer reviews will consider the qualifications of the PI and proposed team, but staffing will continue to be an essentially institutional responsibility.

FES will need to work closely with the IO. For instance, FES funding must be made taking into account decisions made by and guidance from the IO in order to align the U.S. research goals with those of ITER. FES should keep some contingency funding available to address research areas that were not previously identified.

Guidance is needed on the envisioned scale of this program. For instance, is a 20 FTE or 100 FTE level effort foreseen? This will affect how people can be switched in and out of working on ITER research, and will shape the selection of the ITER Research Team members to a significant extent.

Projection of the resources needed for the US to fulfill its commitment, namely 13% of effort during the operational phase, is needed but such estimates have substantial uncertainty. Since an order of magnitude estimate is desired, we approach this two ways: from reported resource usage on US fusion facilities, as well as large scale high energy physics international collaborative facilities. The range of anticipated US participation from these estimates was 200-400 FTE.

The large uncertainty in the appropriate size of the ITER team can only be resolved by a 'bottoms up' assessment of the operational needs of the ITER project, including for example, engineering, technical, software, control and diagnostics expertise. Experience from other projects will be a useful check that the assessment is realistic. In parallel, FES and the US fusion community should have a broad discussion of the scientific and technical areas on which it wishes to have the strongest efforts. These two exercises will form the basis for a negotiation between the USDA and IO on the size and composition of the US ITER team.

Whatever decisions are made in this regard, FES needs to make sure that the funding for team members is above some critical minimum. As a general rule, our feeling is that the threshold (for most of the Team members) is around 0.5 FTE. At this level, scientists could actively be involved in the ITER project and make meaningful research contributions. Similarly, we feel a group effort which is providing specific expertise (for example, plasma control system or software solution development) could also contribute meaningfully at this support level. A 0.25 FTE level will not allow for that. University professors, however, might effectively participate at the lower level (i.e, summer salary level), provided their graduate students or post-docs would be funded at the higher level.

Once selected as part of the ITER Research Team, US personnel will participate in ITER and follow publication guidelines as will be developed by ITER. Our recommendations in this regard were given in Chapter 1.

Given the large number of institutional, national, governmental and international administrative domains that will be involved, there is a critical need to minimize administrative overhead and multiple reporting requirements as much as possible.

1.1 Rough estimates of needed US workforce for ITER

The FY13 level of effort on the three large US facilities was obtained from a self-reported survey [Source: Estimates provided upon request by the Department of Energy, Office of Science, Fusion Energy Sciences], and for JET from L.D. Horton (private communication), see the table below. The FTE counts include scientists (research scientists and faculty), engineers and technicians, and postdocs and graduate students.

	C-Mod	NSTX	DIII-D	JET
Total FTEs	94	226	169	700

Note that the balance of manpower differed between the facilities. Typically Alcator C-mod had the highest fraction of participation from students. Specifically in FY13 NSTX-U was under construction, and so a large fraction of its manpower was for temporary technical labor. Scaling simply by device major radius, from e.g. JET, gives a projected ITER manpower requirement of 1470 FTE. If the US provides 13% of this manpower, i.e. linearly proportional to our projected cost obligation during the operations phase, the total is ~ 190 FTE. The total is actually likely larger than this, with additional resources needed for the substantial auxiliary heating systems, but this estimate gives a lower bound on the likely required level of effort. One can also see from this table that the increase in manpower from C-Mod/DIII-D to JET is faster than with major radius. Scaling with e.g. R^2 gives an expected total manpower requirement of about 3070, with the US portion about 400.

The resource needs for ITER might also be estimable by comparison with other large-scale international physics collaborations. For reference, CERN employs ~ 2400 full-time employees and 1500 part-time employees, and hosts some 10,000 visiting scientists and engineers, representing 608 universities and research facilities. At its peak with the Tevatron running, Fermilab had a staff of approximately 2000 employees. Approximately 2500 scientists from 230 universities and laboratories in 35 states and 30 countries carried out research at this energy frontier. The scale of ITER is larger than either of these facilities, hence 2000 FTEs overall may represent a minimum. The US share of such a venture would be ~ 250 FTEs. With these two methods to estimate level of effort, the US portion could be expected between ~ 200-400 FTEs.

2. Oversight of U.S. research and technical involvement.

A single domestic entity should provide the technical oversight of US ITER participation, both scientific and technical tasks. This entity would "take over" from the US IPO and BPO, or perhaps evolve from the BPO by assuming responsibilities appropriate for this new phase of the program. In this section we will use the abbreviation "xPO" for this entity. Given the potential multiplicity of management bodies (IO, DoE, xPO, ITER task forces, participant home institution, etc.) streamlining management and reporting processes will be crucial for the efficiency and effectiveness of US activities.

The role of the xPO will be to:

1. Coordinate planning activities with the ITER IO and on-site Team managers (see 3. Below) including timely development of the likely subtopics within each Task Force to allow enough lead team for the proposal writing and review cycle. These need to be seen as a natural flow-down from the ITER Research Plan to ensure consistency of our effort.
2. Interface to the ITER Task Forces to help develop research goals and specific research thrusts of various Task Forces, and then disseminate these to the U.S. fusion community.
3. Facilitate discussions among groups to develop strong US-wide research teams to participate in ITER Task Forces.
4. Mediate any disputes with the US team and ensure that the interests of smaller groups are represented.
5. Review progress and level of participation, success in achieving deliverable goals, etc.

The eventual role of the ITPA is not presently known. Either ITPA Topical Groups will morph into the ITER Topical Groups, or they will remain in their present incarnation and supporting global, and not just ITER-centric, fusion research. Similar questions of scope will need to be decided for the USBPO. In either case, the xPO should proactively interact with these groups in order to facilitate input from the international fusion effort into the US domestic research and ITER Research Team planning activities.

3. On-site management of U.S. ITER Team

As discussed in Section 2A, there should be an on-site managerial coordination of US-ITER participation to ensure that US-based interests (DOE-FES, xPO, institutes) are adequately represented and managed. This management team should reside on-site, and serve as the primary contact of the xPO and FES for assessment of US research goals and personnel on the (on-site and off-site) ITER Team. This includes both physics and technical participation. It is appropriate that the lead person on the team, as the US Team leader, be on the “ITER Physics Program Committee” that develops both the high-level and more focused research goals for the experimental campaigns.

2C. Execution of the research program, and data access/storage issues

Effective mechanisms for remote participation and timely access to data will be critical for the success of U.S. research on ITER. In this section we discuss principles and processes for ensuring those capabilities and make recommendations to resolve particular policy and technical issues well ahead of ITER operation. In addition, we address the issue of data caching, and also the applicability of the model for data sharing that has been adopted by the High Energy Physics community for the Large Hadron Collider experiments.

Main takeaways:

a) In accordance with the ITER implementing agreement, IP Annex, all raw and processed data should be made available to *all* members of the ITER team, as broadly defined in 2A.
b) ITER team members need to be able to work effectively no matter where they physically located.
c) At the same time we recognize the advantageous sociology in groups of researchers working at the same physical location - whether at the Cadarache site or at designated remote control/participation sites.
d) The computer and communications architecture should support all modes of participation to the fullest extent possible.

1. Definitions: we need to clearly distinguish between remote control, remote participation and remote data access.

- Remote data access: The ability to read any piece of ITER data from locations **outside** the ITER plant operating zone (POZ). Generally we want this to be possible wherever a user happens to be and with minimal infrastructure requirements on the user. (We do however assume an adequate level of cyber-security - passwords, tokens, certificates, or whatever will be required for access to the ITER data store). Access to the ITER POZ is tightly controlled and restricted, but it is part of the ITER plan to “export” all data from the POZ to systems outside this zone. This exPOZ data archive is the repository that would be accessed remotely (including by ITER participants in their offices at Cadarache). The results of any subsequent data analysis, amendment or correction should (we believe must) be written back into this exPOZ archive (which may be monolithic or federated) for sharing with the rest of the team. (Recommendation: The US should develop a consistent position on this issue.)
- Remote participation: This refers to participation in ITER activities from a remote site, especially machine operations, diagnostics and research, but also includes attendance at meetings, seminars, etc.
- Remote control: Extends remote participation to the ability to operate or configure hardware or software, including diagnostics, used as part of ITER experiments. Policy on remote control is still unclear, though a prototype “operations request gatekeeper” was developed by GA & MIT under an ITER contract to demonstrate secure control, particularly for diagnostics. While we do not see a technical

barrier to full remote physics control – in the sense of setting up and dispatching ITER pulses - this may be forbidden by policy. An alternate approach, and one used at large facilities in other scientific disciplines, is to allow for remote experiment leaders, with physical control of discharge parameters remaining in the hands of operators at the ITER site. Remote control of specified diagnostic settings would seem to have much less risk associated and more obvious advantages. An assessment of potential risks and benefits to the various levels of remote control should be carried out before policies or system architectures are finalized. (Recommendation: The US should develop a consistent position on these set of issues including a possible tiered approach which allows remote control of some functions but not others.)

2. These capabilities are hierarchical

- Remote participation implies remote data access
- Remote control implies remote participation

3. These capabilities increase scientific productivity as they allow broader US participation and reduce unnecessary, and possibly rationed, international travel. At the same time, a significant US presence in Cadarache will be required and is assumed. US participants, not on assignment to Cadarache, but carrying out ITER research would almost certainly spend some time in Cadarache as part of their activities. The importance of establishing personal relationships with other international team members is clear and cannot be achieved through remote participation alone.

4. There is also value in working as a team on particular experiments from a particular remote site. Effective interpersonal communication among a geographically distributed team is challenging even in the internet era. Thus it may be advantageous to set up one or more US facilities for remote participation. It is not clear yet how much infrastructure would be required, but the US program has lengthy and extensive experience already in this area which suggests that it would not be beyond the capabilities of any institution with a significant program presence. It is probably too early to decide on the number and location of such sites, but rather to assume their existence for now. Capabilities for remote control would depend on the eventual ITER policy and concomitant security infrastructure and support requirements.

5. We believe it is critical that all ITER data be shared by the whole team. This is in fact, written into the implementing agreement IP annex. Care must be taken in the data management architecture to ensure this principle is followed and supported in practice. The utility of the ITER experiment will be compromised if data access is restricted by rule or made cumbersome by technology. "Private" data archives should be actively discouraged. That said, we note that there is always a class of provisional data that is produced in the course of analysis or modeling and is not archived or shared. Our view is that when an ITER researcher archives a piece of data it becomes ITER data and should be shareable - and absolutely no data should appear in publications or presentations that is not archived and shared.

6. This data sharing principle means that:

a) There should be a single well-known "source" for any piece of data. Thus any copies of data should be viewed as volatile "caches", not independent archives. This helps ensure data consistency – a critical requirement. In order to make data widely useful, sufficient metadata should be provided for all data items including descriptions of data provenance – that is the origin and history of each data item. Thus while it may be that some large sets of analysis or simulation data become obsolete and could be deleted from the archive, it will be good practice to hold on to the corresponding metadata indefinitely.

b) All processed data that appears in or provides a basis for results in publications or presentations must be written back to a globally accessible archive which could be a single physical archive or a set of federated archives (this is consistent with DOE policies on data access). A federated archive could map multiple archives, maintained on different systems or in different locations into a single virtual data structure that appears monolithic to data consumers or providers. In principle data should be transparently accessed by some extended naming convention without regard to where or how it is physically stored. This is simplest if the data is included in a central archive, but this is not a requirement. Clearly the same API should be used for all of the ITER archives.

7. "Smart" data caches can, in fact, greatly enhance scientific productivity by bringing copies of data items closer to end users. Network latency, the temporal delay for any particular network packet is obviously limited by the speed of light, even for very high throughput connections. In our current experiments, users can have a very different experience with respect to data access depending whether the data is on a disk in their own computer, on the same local area network, on a domestic wide area network or available through an international connection. The difference is mainly due to the increase in network latency, although throughput also tends to decrease at the higher latencies associated with international connections. So, especially for display-intensive applications, we want to be able to push data as close as possible to the end user. As noted above, for cpu intensive applications, it is usually best to move the computer power to the data.

8. Thus, rather than envisioning a central US repository for ITER data, we should think about a data caching architecture that is flexible and adaptable. One approach, that may be tested for the current set of international collaborations is to support local caches that fill based on user data read requests. All data writes, however, would be back to the original archive. Thus the cache never contains unique data – only copies – and can be cheap to implement. That is, it does not require the sort of redundancy and reliability associated with data centers. This would reduce the cost of the caches by about an order of magnitude.

9. Back-up procedures for ensuring the protection of ITER data is a separate issue. Certainly an off-site copy or copies may be desirable – but this is not a research issue per se.

10. We do not believe the LHC data and processing infrastructure is a good model to follow for ITER participation for the following reasons.

a) The tiered LHC approach is designed around a research model that depends on high throughput of mostly automated analysis. Data is accessed primarily in batch rather than interactive mode. Network latency, which is a problem for interactive work, was not a principal consideration.
b) The LHC approach moves huge quantities of data unnecessarily. Political considerations prevailed and led to creation of moderate scale computing centers that were assembled at "domestic" sites around the world. The result was an enormous requirement for network bandwidth, whose provision cost considerably more than the computer centers themselves.
c) This approach leads to greater complexity for supporting the sharing of all data and for maintaining data consistency.
d) If the US needs a major computing center for processing ITER data, it is probably much smarter to locate that center in Cadarache, as close to the data as possible. (Recommendation: The US should ensure that provision for such centers at Cadarache is not overlooked - note, these were specified and assumed in the CODAC conceptual design.)

11. U.S. participants will need to follow emerging regulations for data management and access which apply to all government sponsored researchers (summarized at: http://science.energy.gov/funding-opportunities/digital-data-management/). We will need to ensure that rules and procedures adopted by ITER are consistent with these requirements. In addition, the FES office in DOE has explicitly stated that these requirements apply to analysis and simulation codes and provides guidelines for their application (http://science.energy.gov/fes/funding-opportunities/digital-data-management/). For ITER, it seems certain that codes developed by U.S. researchers will be widely used and also likely that large codes will be developed or extended by teams that include members from both the U.S. and our international partners. Since these codes will be critical for the safe and efficient exploitation of ITER, consistency between U.S. and ITER rules for code sharing and support will also be crucial. This issue is clearly connected to intellectual property issues discussed above as well.

Summary of recommendations

- The US (government and community) should develop and articulate a consistent position supporting remote data access and data sharing principles on ITER.
- In concert with other parties or on its own, the US should develop an architecture for distributed data caching, consistent with the principles outlined above.
- The US (government and community) should develop and articulate a position supporting remote control of some ITER functions, particularly diagnostics.
    - At an appropriate time, the US should develop the technical requirements and architecture for remote participation sites
- The US IPO should ensure that there is provision at the Cadarache site for locating adequate computing resources for analysis of ITER data by US researchers.

- The US should ensure consistency between rules and procedures for data/code sharing and access that are emerging from the US government and the ITER organization.

Appendix 1A. Definition of terms used

Campaign: A series of sequential **runs** – historically bracketed by facility maintenance periods.

Experiment: A collection of shots or portions of **shots** that are designed to answer a defined research question.

Experimental Proposal: A document that describes in detail the goals and procedures to accomplish an **experiment** including a plan for each **shot segment** required.

ITER management: Any of a number of groups or individuals empowered to make high-level project-wide decisions. We envision that ITER management will consist of both operational and oversight entities and assume that they will include appropriate representation of partner associations.

Operational Roles: These are jobs for executing each experiment or run day that are assigned to as needed to qualified personnel. These are not necessarily done by separate people, e.g. Physics Operators could also serve as Session Leaders. We note that that these definitions differ from those commonly used in the US. These responsibilities include both segment-based roles, and pulse-based roles:

Segment based roles

- Experiment Leader: leads the research team carrying out a particular experiment including design of pulse segments and any shot-by-shot decisions for that experiment
- Segment Developer: prepares and validates schedules for segments of pulses
- Diagnostic Coordinator: ensures that diagnostics are configured as needed for a given experiment and that required data is being recorded
- Analysis Coordinator: ensures that required post-shot and post-run analysis is performed.

Pulse based roles (proposed IO/ITER responsibility)

- Session Leader: responsible for executing the series of shots scheduled for a particular run day; interfaces between the Engineering Operator and the multiple Experiment Leaders for a desired set of pulses
- Physics Operator: combines and validates multiple segment schedules to construct entire pulse program
- Engineering Operator: makes changes to pulse programming as requested by Session Leader, and in consult with Lead Engineer
- Lead Engineer: has "control room" responsible for facility availability, operation, integrity and safety

Operations management: Individuals or groups with overall responsibility for day-to-day facilities operations.

Pulse schedule: A detailed plan for the plasma control system that defines how the facility would operate on a particular **shot**.

Roles: One of a set of job descriptions that are required for the final planning and execution of experiments – each with a well-defined set of duties. Roles could include, for example, Engineering operator, Physics Operator, Session Leader, Diagnostic Coordinator, etc. Individuals are assigned specific roles, usually for the duration of a **Run**.

Run: A series of sequential **shots** – historically performed on a single day. A run will be dedicated to a set of primary **experiments**.

Shot: An entire single pulse of the experimental facility.

Task Forces: Short-lived groups organized to accomplish a well-defined set of tasks (e.g. "learn how to access and control H-modes during current ramp-up").

Topical Groups: Standing (long-lived) groups organized to address open-ended scientific issues (e.g. MHD, Transport, Boundary, etc.) meant to complement short-term task-oriented research activities.

"xPO": A single domestic entity that provides the technical oversight of US ITER participation, both scientific and technical tasks, subsuming present tasks from e.g. the U.S. ITER Project Office (IPO), US Burning Plasma Organization (BPO), etc.

Appendix 1B: Executive Summary of Sample Data Analysis Workflows from Alcator C-Mod, DIII-D, and NSTX

The U.S. Burning Plasma Organization (BPO) has formed a working group "Modes of Participation in ITER". One main purpose of this group is to work with the IO and provide information on modes of operation and analysis in present day devices, to contribute to the formulation of the strategy for ITER experimental operation procedures and support systems. To this end, the BPO group has engaged in discussions with IO employees, A. Winter and S. Pinches, who are tasked with formulating a proposal for ITER experimentation strategy.

As a first step, the BPO group has documented sample between-pulse analyses done in Alcator C-Mod, DIII-D, and NSTX. These advanced analysis tasks have been used to guide the conduct of specific experiments in these devices. An experiment on these devices typically represents a sequence of plasma discharges (often in a single experimental session) aimed at investigating various aspects of a particular plasma phenomenon. Between-pulse analysis allows some modification or refinement of the run plan based on what has been learned from previous pulses. For operation on ITER, this could translate to the adapted selection of pulse schedules from a set of previously validated set of pulse schedules with the aim of making the most effective use of run time. Moreover if multiple experiments are performed in a single pulse, each sub-experiment will require a set of validated pulse schedules, and the entire pulse will need validation to make sure one experiment isn't affected by changes to another.

The workflows for these tasks are described in the subsequent pages, including a cursory evaluation of the implied requirements for ITER. The sample tasks described below include gyrokinetic analysis (C-Mod), energetic particle instability analysis (DIII-D), and transport and confinement analysis (NSTX). While these tasks are presently carried out between pulses in US devices, processing of certain analysis elements that underpin more in-depth analysis concurrent with the ITER discharge will be needed to insure timely completion between ITER pulses.

The common elements in the workflow analysis from these three devices include:

1. Some level of automated data analysis and required data availability
    a. Magnetics data, from which equilibria are reconstructed as a launching point for further analysis, is presently acquired rapidly enough to allow reconstruction of the equilibrium state during multiple time slices.
    b. Plasma data used as input for subsequent calculations (mostly profile data) is required as inputs to multiple analysis tasks; a more accurate equilibrium calculation actually uses plasma profile data, which is typically available a few minutes after the end of the pulse in present devices.
2. Some level of interactive tasks
    a. Review of data to disregard obviously questionable inputs and to validate data for higher level analysis.
    b. Semi-automated creation of scripts to run in-depth analysis

3. There is a hierarchy of calculations: increasingly constrained equilibrium calculations, e.g. magnetics-only, including kinetic + diamagnetic data, including MSE, etc. are used for different purposes; many of these are presently done between pulses.
4. Analyzed data is written, along with meta-data for provenance, to a central archive accessible by all team members who have been granted access to data, typically through project-wide data usage agreements.

These translate to the following requirements for ITER:

1. Rapid access to "critical" data for between shot inspection and evaluation of run progress
2. Rapid access to allow important automated analysis tasks to begin as quickly as possible – ideally concurrent with the plasma discharge, but minimally to complete soon after each shot. Most notable in this regard is the calculation of the magnetic equilibrium.
3. Reserved cycles on local machines for analysis tasks.
4. Ability to declare and distribute 'events' that signal that needed data is available, allowing the workflow to proceed and be monitored; this is important for tasks with both automated and interactive components.
5. Availability of a central repository where analyzed data and corresponding meta-data can be stored and accessed by all authorized team members.

Comparison of the three tasks also highlights certain flexibility needed:

1. The ability to specify the order of priority for diagnostic signals for specific experiments; e.g. fluctuation experiments will need fluctuation data to come in quickly, while other experiments may not need this.
2. The ability to communicate results of analysis that could guide the experiment execution to the session leader and diagnosticians.
3. The ability to modify the configuration or control of the device between successive pulses of an experiment.

Although not discussed in the sample analysis workflows below, higher-level functionality is also desirable and would improve overall coordination and productivity. These capabilities exist piece meal in existing US devices, but will likely need to be developed in an integrated fashion for ITER. Additional specific recommendations will follow in subsequent documents. Examples of the required tools include:

1. <u>Run information database</u>: This information includes descriptions of each day's experiments with links to experimental proposals, run plans, run summaries, logbook entries and data summaries.
2. <u>Data quality system</u>: Infrastructure that allows users to assign and record data quality metrics.
3. <u>Data analysis/review request system</u>: A system that allows researchers to request data analysis or review from other members of the team and to track the status and results of these requests.
4. <u>Data provenance system</u>: Infrastructure that supports the documentation and annotation of raw and processed data through the entire analysis chain.

Appendix 1B1: Sample workflow for data analysis from Alcator C-Mod
Between-shot linear gyrokinetic analysis using GYRO code
M. Greenwald

Task Description

Data from a broad set of diagnostics is marshaled, analyzed and reviewed to prepare inputs for a linear gyrokinetic code (GYRO), running on HPC hardware.

Physics Motivation

When available between shots, the results of linear gyrokinetic calculations are used to help guide machine operations. The relevant experiments require operation at conditions specified relative to stability boundaries (e.g. the boundary when Ion Temperature Gradient (ITG) fluctuations/transport overtake Trapped Electron Mode (TEM) fluctuations/transport) – which would otherwise be computed after the run. By carrying out the analysis between shots, we are able to gather the required data with the fewest number of shots, selecting from a previously prepared menu of planned discharges. Without this capability, we would be operating somewhat blindly and would have to more exhaustively sample the operating space to ensure that we obtained the data needed. Since manual analysis and review are part of the workflow, this approach presents requirements for data availability and coordination among the analysis team.

Workflow

The workflow can be broken roughly into 3 phases. In the 1st phase, automated processes acquire, store data, then carry out low level analysis tasks (steps 1-3). In the 2nd, interactive tools are used for higher level analysis and manual data review (steps 4-7), finally in the 3rd phase, input files for are assembled and the stability jobs are dispatched to an HPC system (steps 8-12).

1. Automated data acquisition & data storage
    a. Required data set includes magnetics, density and temperature profiles from multiple diagnostics, various spectroscopic measurements
2. Automated conversion to physical units, application of calibration factors
3. Automated post-shot analysis
    a. Equilibrium (EFIT – run with several different input options)
    b. Diagnostics: Thomson, Edge Cyclotron Emission (ECE), Motional Stark Effect (MSE), Impurity assessment via effective charge (Z_{eff}), Charge eXchange Recombination (CXR), Xray Ion Crystal Spectrometer (XICS), etc.
    c. Calculation of statistical error bars
4. Set data retrieval parameters for GYRO runs
    a. Select which data to use (which EFIT "tree", which diagnostics for Ti, velocities, etc.)
    b. Set radius and time range of interest

    c. Set radial grid for mapping/interpolation
5. Get "raw" EFIT data
6. Process EFIT data
    a. Compute magnetic flux surface shape profiles
    b. Smooth in time as required
    c. Display and review results
    d. Or compute kinetic or MSE constrained EFIT equilibrium if desired and repeat a-c; these provide more a realistic representation of the equilibrium, as they use more plasma data as constrains
7. Retrieve, fit and map profile data for ne, Te, Ti, V, q, etc.
    a. Typically requires interactive analysis and/or expert review
8. Load scalar parameters required by code
9. Load raw and mapped profiles
10. Set GYRO run parameters (krange, nk)
11. Invoke script to run code
    a. Read data
    b. Write GYRO input file for normalization run
    c. Write batch file for normalization run
    d. Submit GYRO job (write output file)
    e. Read/parse output file
    f. Create k-array data
    g. Write set of input files, batch files
    h. Move files to directory structure
    i. Submit each "directory" (run GYRO)
    j. Check queue for job status
    k. Read files
    l. Write data to archive (MDSplus tree)
    m. Compute derived quantities, write to archive (MDSplus tree)
    n. Clean-up
12. Invoke procedures to display data

Implied Requirement

- Need for some level of self-documenting automatic data processing that integrates more than one plant system

    - (Note: scope might be extrapolated based on C-Mod total, between-shot, automated analysis jobs for all purposes = 165)
- Need to have raw and automatically processed data available to end users as quickly as possible - ideally during the ITER pulse duration
    - This might require only some part of the total data set, though this is hard to determine in advance
- Users need fast access to data – processing servers may need to be physically close to data archive
- May require direct data access from HPC machines
- May require reserved cycles on HPC for these between-shot analysis tasks
- User must be able to write data back into shared archive
- Need some sort of “events” or notification system to allow users to know when raw or processed data is available
    - User applications (including HPC tasks) must be able to set these events/notifications for subsequent dependent tasks
- Need to track/document status and provenance (with full chain of dependencies) of all automated and manual analysis tasks
    - User applications must be able to write this data to archive

Appendix 1B2: Sample workflow for between shot analysis of energetic particle (EP) relevant diagnostic data, using instability studies as an example

M.A. Van Zeeland, W.W. Heidbrink, G.R. McKee, D.M. Thomas

Task Description

Data from several fluctuation (ECE, BES, CO2 interferometers, magnetics), EP (Fast Ion Loss Detectors, BILD (Beam Ion Loss Detector), Ion Cyclotron Emission), and equilibrium related diagnostics are acquired and analyzed between shots for DIII-D EP experiments. Fluctuation diagnostics in particular are processed primarily through Fourier analysis.

Physics Motivation

For DIII-D EP experiments, rapid between-shot access to information about the particular instability under investigation is essential. Measurements of eigenmode stability, frequency, toroidal mode number, radial structure/localization and induced EP transport are all used to inform actions for subsequent discharges. These actions can specifically target a physics aspect of the instability itself, for example, changing neutral beam heating power to alter mode drive, or be more diagnostic oriented, such as radially shifting the BES array to be better centered on the instability location.

Workflow

1. Automated data acquisition and data storage
2. Select time and frequency range of interest
3. From local user terminal, trigger calculation of windowed spectra for several fluctuation diagnostics
    a. Retrieve data
    b. Divide into overlapping time windows
    c. Fourier analyze
    d. Store resultant surface plot as .eps
    e. Repeat for:
        i. Crosspower spectrogram of vertical and radial CO2 interferometer data
        ii. Crosspower spectrogram of adjacent Electron Cyclotron Emission (ECE) channels across array
        iii. Crosspower spectrogram of successive Beam Emission Spectroscopy (BES) channels (if appropriate beam was used)
        iv. Autopower spectrogram of large bandwidth FILD Photomultiplier channels
4. Automated post-shot analysis
    a. Equilibrium (EFIT with and without Motional Stark Effect – MSE data)
    b. Diagnostics: Thomson, ECE, MSE, Charge Exchange Recombination - CER (simple analysis between shots, detailed analysis overnight)
    c. Rapid spline fits (ZIPFIT) to kinetic profiles created

    d. Toroidal Alfven Eigenmode (TAE) Frequency calculated and stored using q_0, q_{min}, q_{95} values from EFIT and line-averaged density –used to help identify observed instabilities

5. If Alfven continuum analysis desired
    a. MSE based EFITS are used in combination with ZIPFIT density profiles to create input file for single toroidal mode number on EFIT timebase
    b. Calculation is triggered on cluster
6. Expert review of spectrogram data to inform next steps
    a. Initial determination of mode stability, frequency range, and rough amplitude made from CO2 Interferometer spectrogram
    b. Frequency and spectral behavior (i.e. steady, chirping, sweeping) combined with calculated TAE frequency and/or continua used to help identify instability
    c. Mode localization determined from ECE and BES data
    d. Assessment of induced EP loss determined from spectra of FILD channels
7. If satisfied with ECE/ECEI/BES array location relative to instability, continue with physics investigation, otherwise adjust BES steering and/or make minor toroidal field adjustment for improved ECE/ECEI location
8. If physics investigation includes ECH probing of instability and between shot adjustment of ECH deposition location relative to instability
    a. Launch EFIT viewing program
    b. Load present gyrotron setup
    c. Adjust gyrotron steering angles
    d. Using EFIT, electron temperature, and electron density profiles create input file for ECH deposition code, TORAY
    e. Run TORAY to obtain ECH deposition profile
    f. Iterate until desired steering is obtained
    g. Communicate desired steering angles for next discharge to ECH group

Implied Requirement

1. Immediate access to fluctuation data. At minimum, data from diagnostics that probe large regions of plasma (e.g. central interferometer chord) should be available first to give overall picture of instabilities.
2. Variable diagnostic information, e.g. geometry, frequency, mapping between variable names and locations, etc. for key fluctuation diagnostics
3. Fast access to equilibrium, density, and temperature profile information.
4. Fast access to neutral beam and other heating waveforms
5. Ability to communicate desired changes to responsible diagnosticians

Appendix 1B3: Sample workflow for data analysis from NSTX
Between-shot TRANSP analysis
S.M. Kaye

Task Description

Data from a broad set of diagnostics is prepared as input for between- (or among-) shots cross-magnetic field transport TRANSP analysis (BEAST)

Physics Motivation

When available between shots, the TRANSP can be used to help guide machine operations. TRANSP is used not only as a means of diagnostic and data validation, but also as a guide to understanding confinement and transport trends of various scans during experiments. Results will indicate if the data is good enough for more comprehensive post-run analysis, and can provide a guide for whether discharges need to be rerun, or the run plan modified during operation. Since manual analysis and review are part of the workflow, this approach presents requirements for data availability and coordination among analysis team.

Workflow

The workflow can be broken several phases, and different options for run preparation are offered.

1. Automated data acquisition & data storage
    a. Required data set includes magnetics and other 1D time evolving quantities (neutron emission, diamagnetic signal, voltage, plasma current, etc). density, rotation, temperature and radiation profiles from multiple diagnostics
2. Automated conversion to physical units, application of calibration factors
3. Automated post-shot analysis
    a. Equilibrium (EFIT)
    b. Diagnostics: Thomson, Charge Exchange Recombination Spectroscopy (CHERS), Motional Stark Effect data (MSE, if available), Bolometer, etc.
    c. Calculation of statistical error bars
4. Set up TRANSP run using "ELVIS" graphical user interface
    a. Specify EFIT version for run (magnetics-only EFITs for between-shots TRANSP, EFIT with profile information for "among-shots" runs)
    b. Set data source for Zeff
    c. Set time range of calculation (time slice or full evolution; the latter will not complete between shots)
    d. Process data for run setup (run ELVIS interface)
        i. "Scrunch" EFIT data
        ii. Retrieve 1D and profile data, automatically smooth/deglitch data
        iii. User examination of input 1D, 2D data if using ELVIS interface
        iv. Enter comments for run
    e. Submit run directly through ELVIS interface
5. Set up TRANSP run using python script, review input and submit run

a. Run python script specifying EFIT version, other parameters
   i. Automatically "scrunch" EFIT data, retrieve/smooth/deglitch input 1D and 2D data, write UFILES,
   ii. Graphical user interface allows for setting time range of run
   iii. Prepare and write namelist
b. Examine input data using trdat, res-mooth/re-deglitch if necessary
c. Preprocess data using tr_start; enter comments for run
d. Submit run using tr_send

6. Results of running by both methods written into BEAST MDSplus tree
7. Email notification when process complete
8. Run ELVIS or personal script, or other routine (RPLOT) to examine results

Implied Requirement

- Need for reserved and dedicated CPUs on linux cluster for BEAST processes (serial runs)
- Need to have raw and automatically processed data (equilibrium, profiles, in 3a, 3b) available to end users as quickly as possible
- Users need fast access to data – processing servers may need to be physically close to data archive
- Need "events" or notification system to allow automated scripts to know when raw or processed data is available.
  - Preparation scripts "wait" until data is available

Appendix 2: Committee Members

Martin Greenwald, university representative – MIT
Don Hillis, national lab representative – Oak Ridge National Laboratory
Amanda Hubbard, university representative – MIT
Jerry Hughes, university representative – MIT
Stan Kaye, national lab representative – Princeton Plasma Physics Laboratory
George McKee, university representative – UW-Madison
Rajesh Maingi*, national lab representative – Princeton Plasma Physics Laboratory
Dan Thomas, industry representative – General Atomics
Mike Van Zeeland, industry representative – General Atomics
Mike Walker, industry representative – General Atomics

*Chair