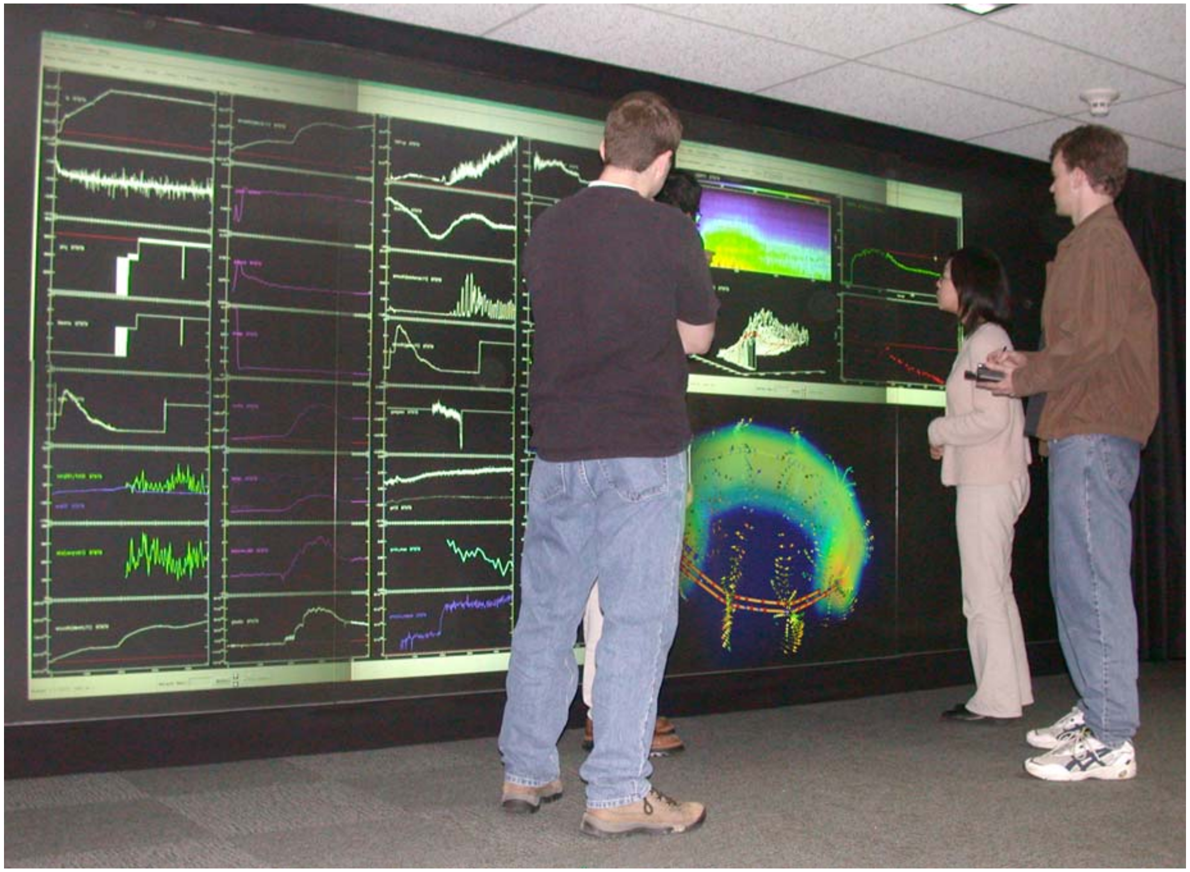
Collaborative Technologies for Distributed Science: Fusion Science and High-Energy Physics



By

David Schissel

Presented at

The Computing Techniques Seminar

FNAL

October 17, 2006

DIII-D NATIONAL FUSION FACILITY

GENERAL ATOMICS

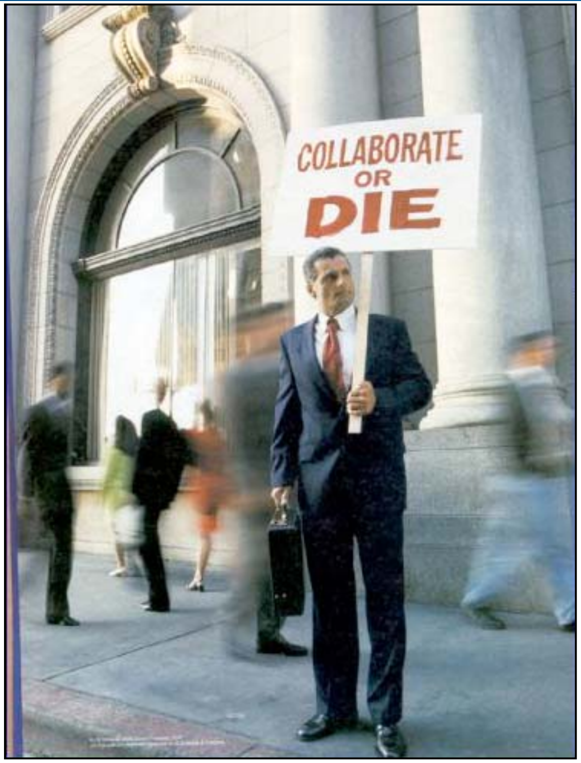
Acknowledgment

- **The National Fusion Collaboratory Project Team Members (Distributed)**
  - C-Mod (MIT), DIII-D (GA), NSTX (PPPL)
  - Argonne National Lab, Lawrence Berkley Lab, Princeton University, University of Utah
- **The Staff of the DIII-D National Fusion Facility**
- **Work is supported by the USDOE Department of Energy**
  - SciDAC: Office of Advanced Scientific Computing Research
  - Fusion Research: Office of Fusion Energy Sciences

Office of Science U.S. DEPARTMENT OF ENERGY

Presentation's Key Points: Fusion Energy Perspective

- **International collaboration is our future:**
  ITER will be the most important facility for 20-25 years
  - For the US to get the most from the project, we must be prepared to exploit the machine remotely



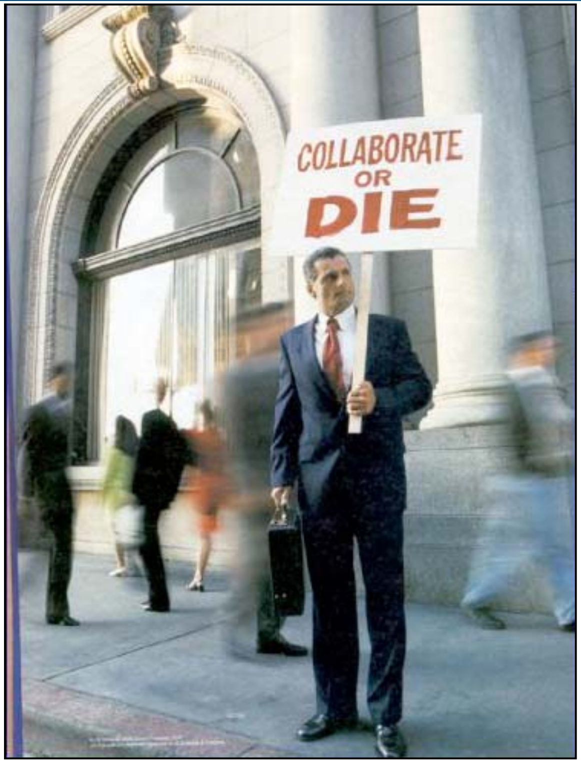
- **Remote collaborations on domestic facilities will continue to be important**
  - Preparation and support for ITER

- **National Fusion Collaboratory Project (FusionGrid)**
  - Scientists using NFC developed tools to enhance current collaborations
  - Remote collaboration: session leadership becomes routine

- **Extend our existing tools to meet future needs (Orbach's FESAC charge)**
  - Functionality, international FES, other SC programs (e.g. HEP)
  - Prototype tools and methodology for ITER

Presentation's Key Points: In General

- **Collaboration needs of FES and HEP are similar**
  - ITER, LHC@FNAL, ILC
- **Collaborative Workspaces**
  - Ad-hoc and structured communication
  - Shared applications
- **Secure Computational Services**
  - Improved security
  - Continuous data sources: MDSplus
- **Requirements overlap justifies joint research**
  - Benefit DOE/SC: OFES, OHEP, and OASCR



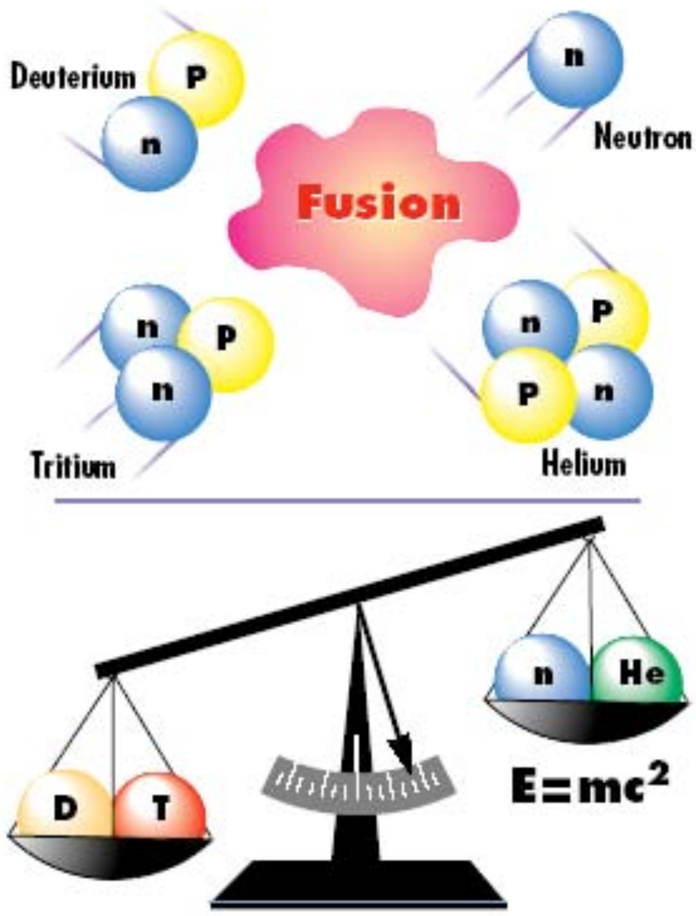
Outline of Talk

- **What is fusion?**
  - Power of the stars and with an aim to harnesses on earth
- **Where is fusion research performed?**
  - International and moving to ITER in France
- **What does fusion research involve?**
  - Demanding near-real-time experimental program
  - Theoretical program on massively parallel machines for simulations
- **What collaboration tools are assisting in the international research?**
  - SciDAC funded FusionGrid: Collaboration and Grid computing
- **How closely aligned are the needs of FES and HEP?**
  - For collaborative activities, they appear very closely aligned

1. What is Fusion?

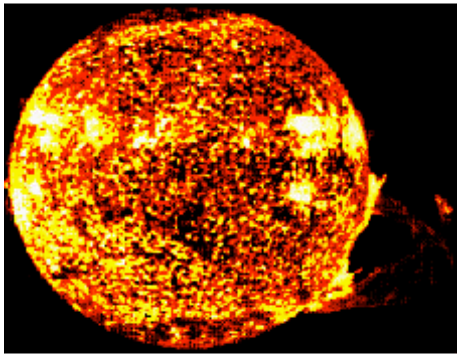
The nuclear reaction whereby the nuclei of light isotopes, like Hydrogen are joined (fused) to form heavier elements, releasing large amounts of energy.

The Fusion Reaction Powers the Stars and Produces the Elements of the Periodic Table



- For 50 years worldwide, teams have been trying to exploit the fusion reaction as a practical energy source

- The promise is for an environmentally friendly method for generating electricity with an inexhaustible fuel supply



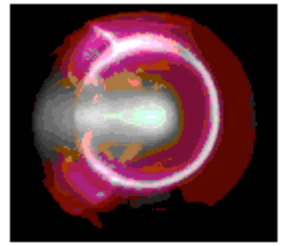
The Fusion Reaction Elastic Scattering Curves Lead us to the Study of Confined Plasmas

- **Even at the optimum energy, the nuclei are much more likely to scatter elastically than to fuse**

- **Nuclei must be confined for many interaction times**

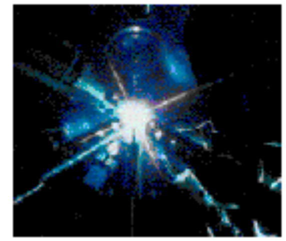
- **Multiple scatterings thermalize the constituent particles**

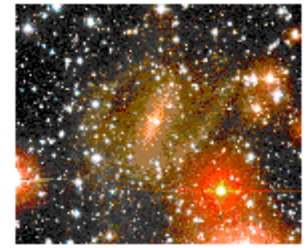
- **At the energies involved (10 - 100 keV) matter becomes fully ionized = plasma**



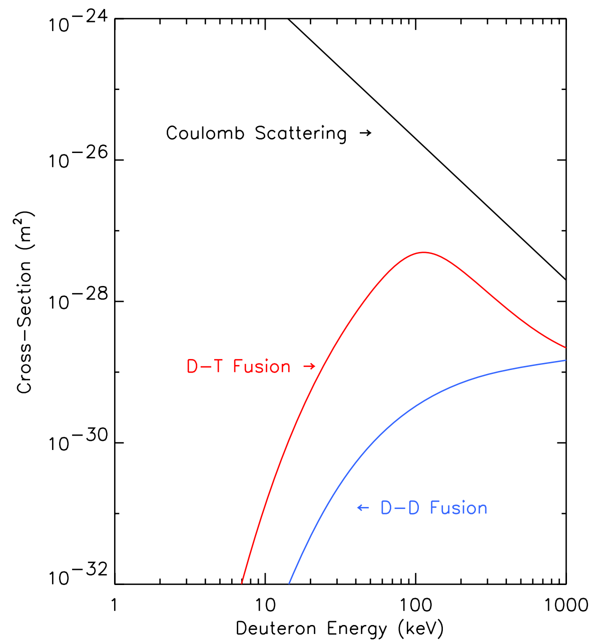
Magnetic Confinement

Inertial Confinement



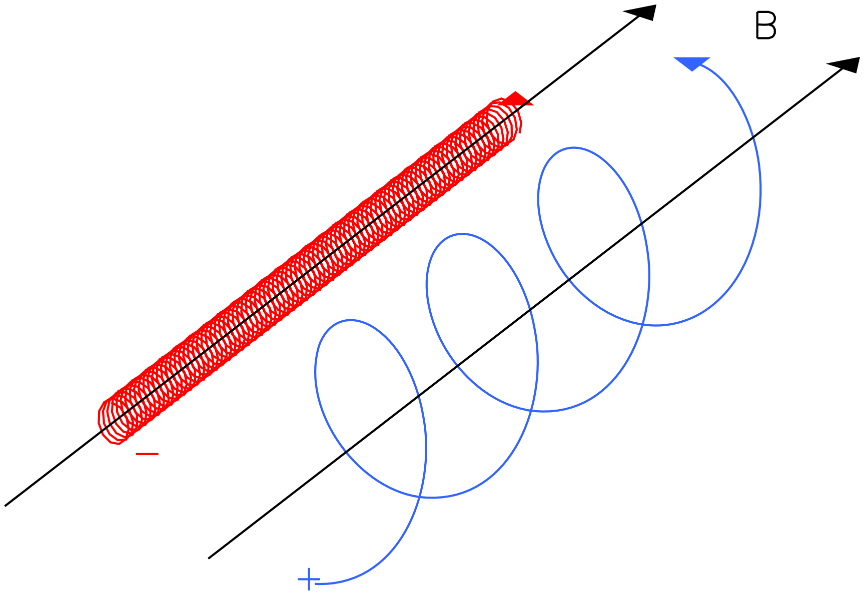


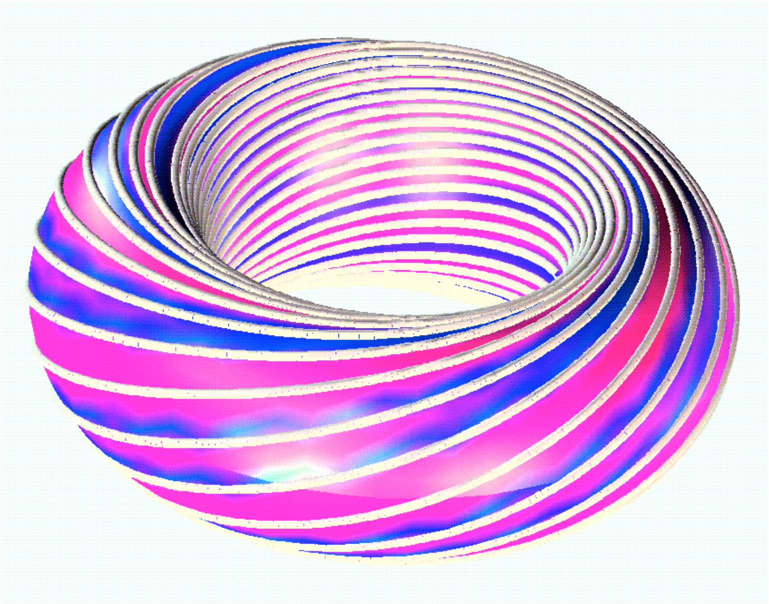
Gravity



Magnetic Confinement - Ionized Particles are bent by the Lorentz Force into Circular Orbits

In the simple example shown, there is no confinement at all parallel to the magnetic field (B)





For a practical device, the end losses must be eliminated

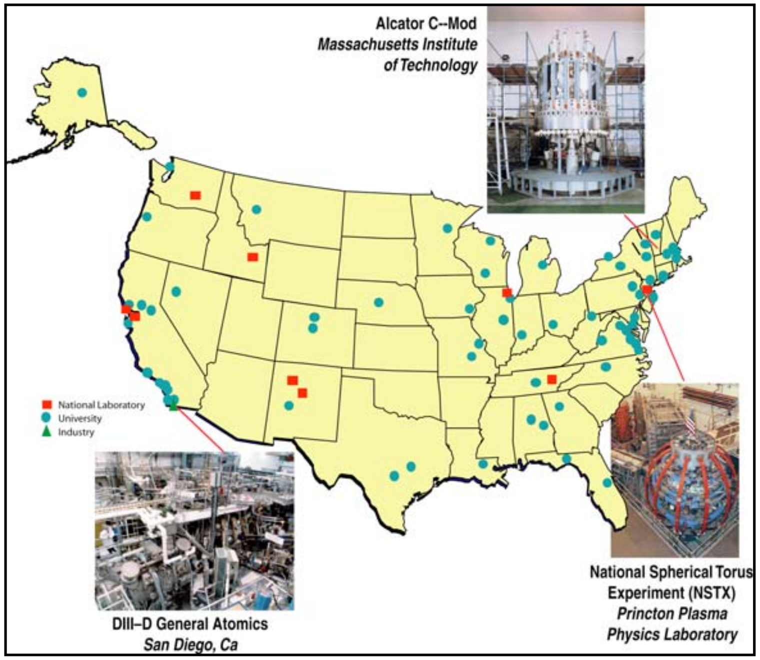
Fusion Research Presents Many Challenges

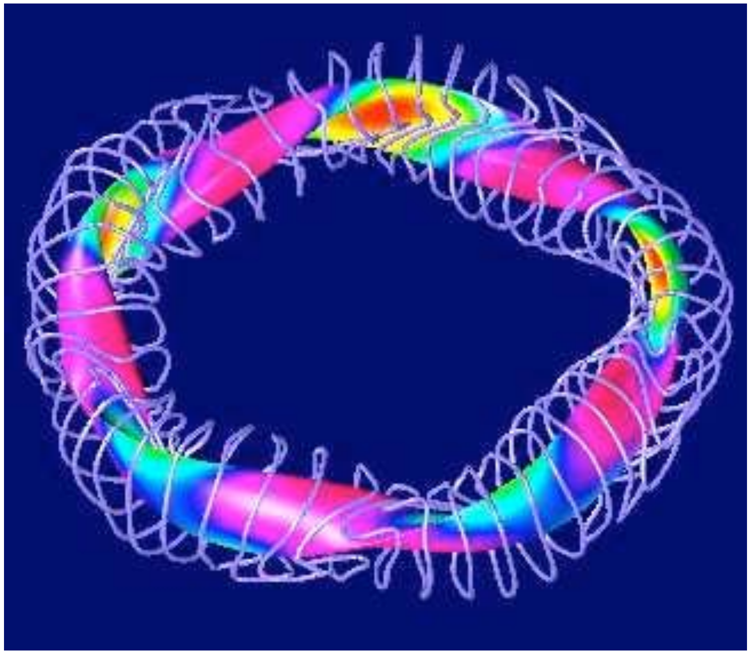
- **Development of physical models for plasma stability and transport**
  - Vast range in space and time which can span over 10 decades
  - 3D motion, extreme anisotropy, free energy driven turbulence
- **Design large experiments**
  - 3D coupling of electromagnetics, structures, heat transfer, neutronics
- **Development of complex diagnostics**
- **Development of plasma heating and fueling methods**
- **Acquisition, analysis, display and interpretation of large quantities of experimental data**
- **All of these are computationally intensive**

2. Where is Fusion Research Performed?

Truly a worldwide open scientific endeavor that is highly collaborative.

Three Large U.S. Experimental Facilities and a Vibrant Theoretical Community





- **3 Large Experimental Facilities**
  - ~$1B replacement cost
- **Numerous theoretical groups**
  - High-performance computing
- **67 U.S. fusion research sites**
  - Over 1500 scientists
- **Efficient collaboration is required**
  - Geographically diverse teams

Fusion Science Today is Worldwide Team Sport

- 90 institutions participate
- 425 active users
- 317 scientific authors
- Students and faculty from
  - 65 universities
  - 28 states

Active Collaborations 2004

US Labs

ANL (Argonne, IL)
LANL (Los Alamos, NM)
LBNL (Berkeley, CA)
LLNL (Livermore, CA)
ORNL (Oak Ridge, TN)
PPPL (Princeton, NJ)
SNL (Sandia, NM)

Industries

Calabasas Creek (CA)
CompX (Del Mar, CA)
CPI (Palo Alto, CA)
Digital Finetec (Ventura, CA)
DRS (Dallas, TX)
DTI (Bedford, MA)
FAR Tech (San Diego, CA)
IOS (Torrance, CA)
Lodestar (Boulder, CO)
SAIC (La Jolla, CA)
Spinner (Germany)
Tech-X (Boulder, CO)
Thermacore (Lancaster, PA)
Tomlab (Willow Creek, CA)
TSI Research (Solana Beach, CA)

US Universities

Auburn (Auburn, Alabama)
Colorado School of Mines (Golden, CO)
Columbia (New York, NY)
Georgia Tech (Atlanta, GA)
Hampton (Hampton, VA)
Lehigh (Bethlehem, PA)
Maryland (College Park, MD)
Mesa College (San Diego, CA)
MIT (Boston, MA)
Palomar (San Marcos, CA)
New York U. (New York, NY)
SDSU (San Diego, CA)
Texas (Austin, TX)
UCB (Berkeley, CA)
UCI (Irvine, CA)
UCLA (Los Angeles, CA)
UCSD (San Diego, CA)
U. New Mexico (Albuquerque, NM)
U. Rochester (NY)
U. Utah (Salt Lake City, UT)
Washington (Seattle, WA)
Wisconsin (Madison, WI)

Russia

Ioffe (St. Petersburg)
Keldysh (Udmurtia, Moscow)
Kurchatov (Moscow)
Moscow State (Moscow)
St. Petersburg State Poly (St. Petersburg)
Triniti (Troitsk)
Inst. of Applied Physics (Nizhny Novgorod)

European Community

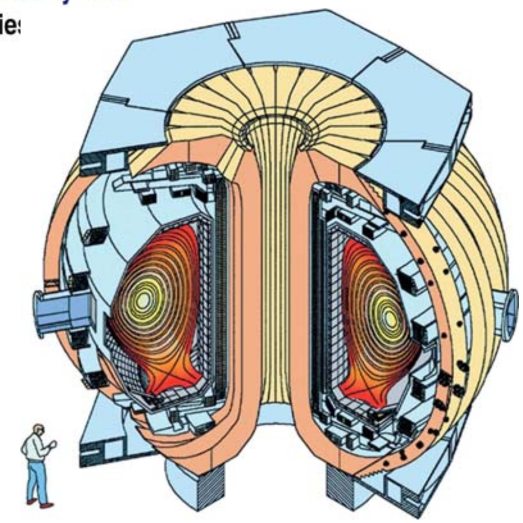
Cadarache (St. Paul-lez, Durance, France)
Chalmers U. (Goteberg, Sweden)
CFN-IST (Lisbon, Portugal)
CIEMAT (Madrid, Spain)
Consorzia RFX (Padua, Italy)
Culham (Culham, Oxfordshire, England)
EFDA-NET (Garching, Germany)
Frascati (Frascati, Lazio, Italy)
FOM (Utrecht, The Netherlands)
Helsinki U. (Helsinki, Finland)
IFP-CNdR (Italy)
IPP (Garching, Greifswald, Germany)
ITER (Garching, Germany)
JET-EFDA (Oxfordshire, England)
KFA (Julich, Germany)
Kharkov IPT, (Ukraine)
Lausanne (Lausanne, Switzerland)
IPP (Greifswald, Germany)
RFX (Padova, Italy)
U. Dusseldorf (Germany)
U. Naples (Italy)
U. Padova (Italy)
U. Strathclyde (Glasgow, Scotland)

Japan

JAERI (Naka, Ibaraki-ken, Japan)
JT-60U
JFT-2M
Tsukuba University (Tsukuba, Japan)
NIFS (Toki, Gifu-ken, Japan)
LHD

Other International

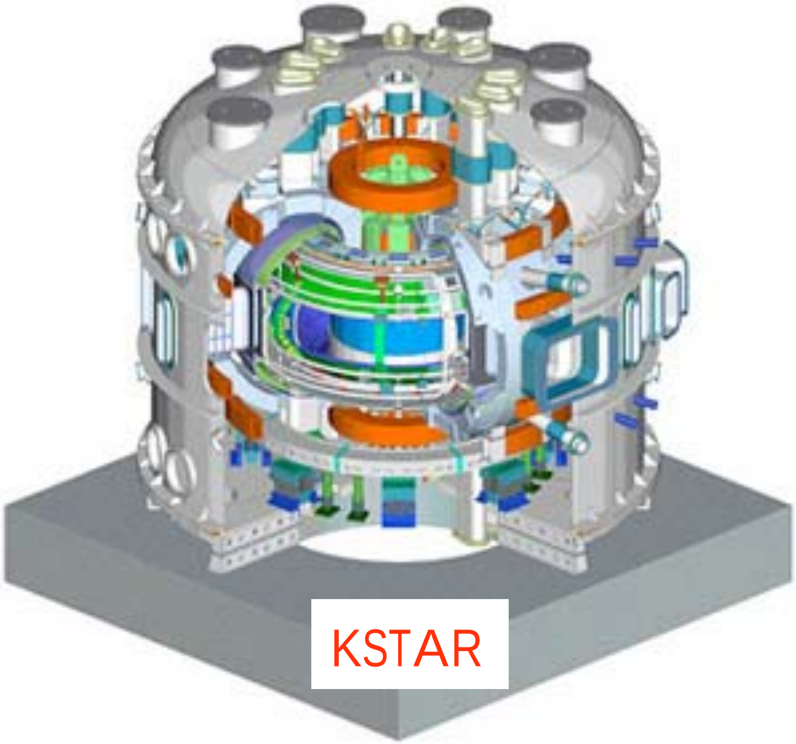
Australia National U. (Canberra, AU)
ASIPP (Hefei, China)
Dong Hau U. (Taiwan)
KBSI (Daegon, S. Korea)
KAERI (Daegon, S. Korea)
Nat. Nucl. Ctr. (Kurchatov City, Kazakhstan)
Pohang U. (S. Korea)
Seoul Nat. U. (S. Korea)
SWIP (Chengdu, China)
U. Alberta (Alberta, Canada)
U. of Kiel (Kiel, Germany)
U. Toronto (Toronto, Canada)



An Example From The DIII-D National Fusion Facility in San Diego

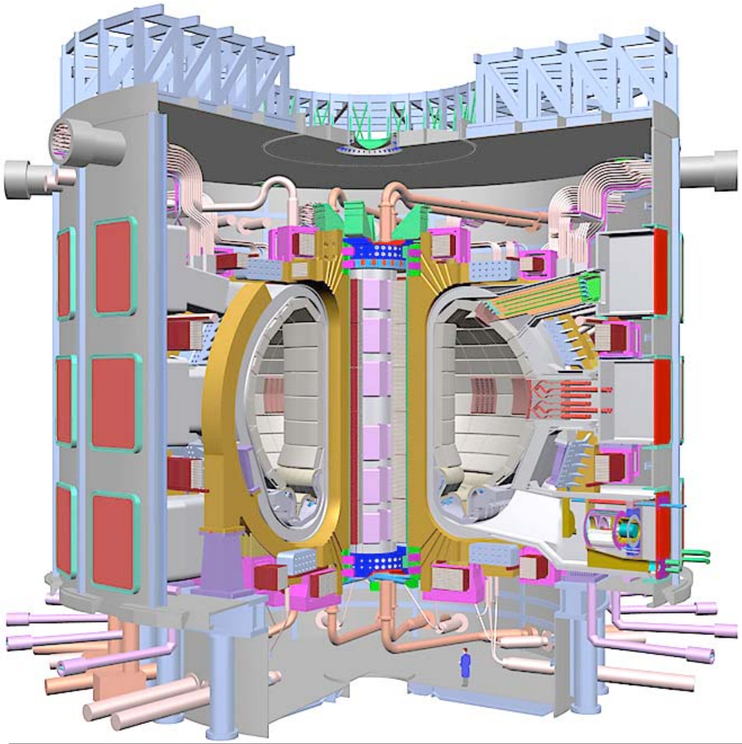
Fusion Science Today is Worldwide Team Sport

Two New Machines in Asia: China and South Korea



- **EAST had first plasma last month, KSTAR under construction**
  - Both have significant U.S. collaboration
- **Excellent test beds for distributed science on ITER**

Next Fusion Device is ITER to be Built in France



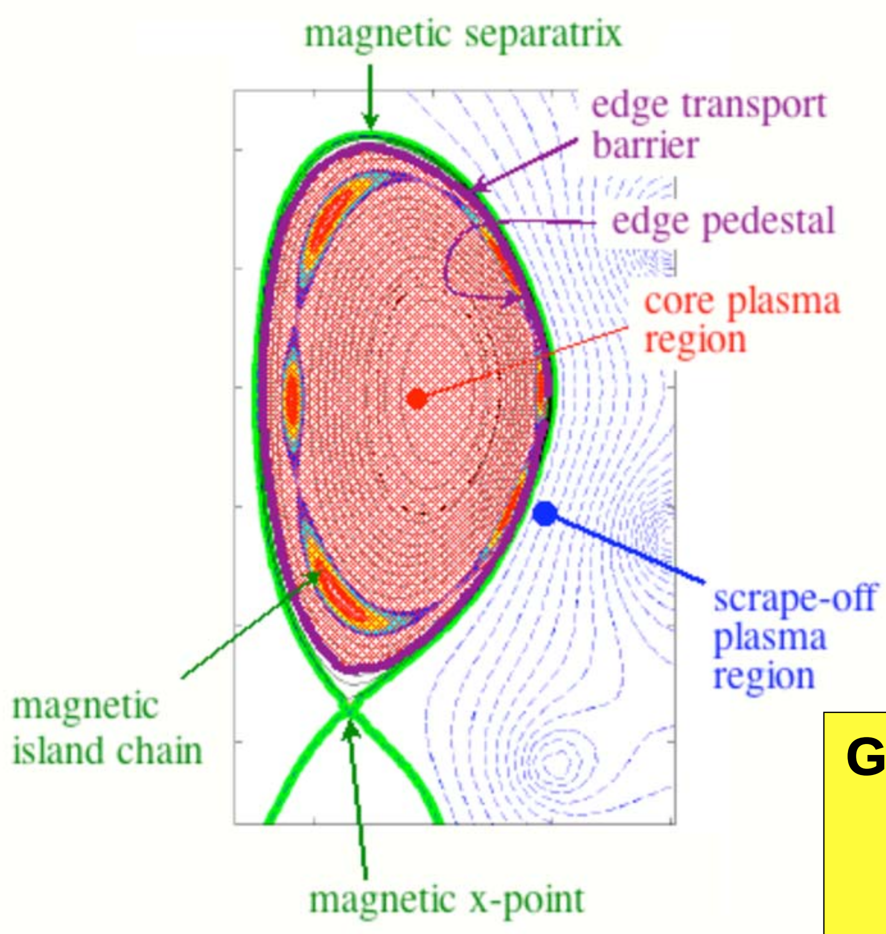
- **China, Europe, India, Japan, South Korea, Russia, United States**
- **~5B total construction cost**
  - First plasma ~10 years
- **Burning plasma experiment**
  - Demonstrate physics viability

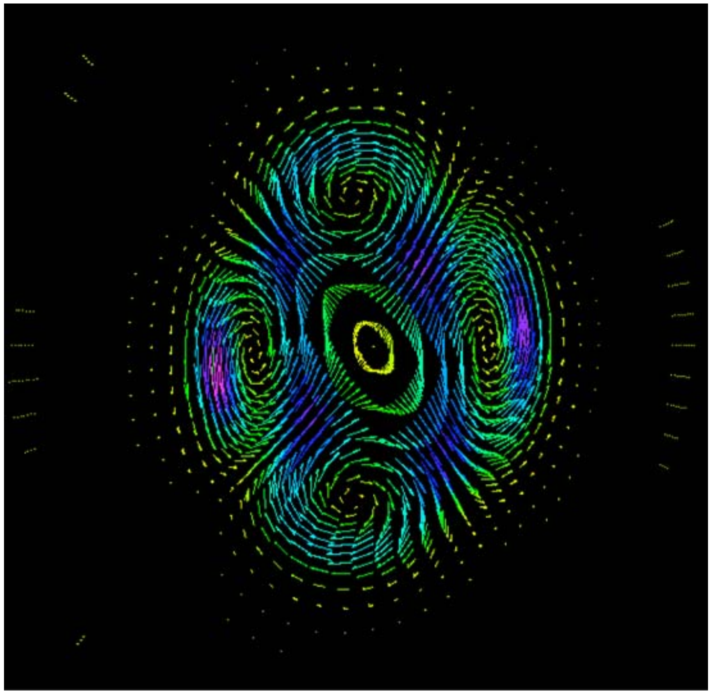
First on our list is fusion. The prospect of limitless source of clean energy for the world leads with our commitment to join the international fusion energy experiment known as ITER.
– Secretary of Energy Spencer Abraham, November 10, 2003
Introducing the Department's 20-year plan for building the scientific facilities of the future.

3. What does Fusion Research Involve?

The combination of complex simulation codes running on massively parallel computers and very large scale experimental facilities.

Fusion Simulation Project (FSP): Integrated Simulation and Optimization of Fusion Systems





Goals of joint OFES & OASCR Program:
- Comprehensive models
- Architecture for integration
- Computational infrastructure

Experimental Fusion Sciences Places a Large Premium on Rapid Data Analysis in Near-Real-Time

- **Pulsed Experiments**
  - 10s duration plasma every 20 minutes
- **20-40 people in control room**
  - More from remote locations
- **10,000 separate measurements/plasma**
  - kHz to MHz sample rates
  - Between pulse analysis

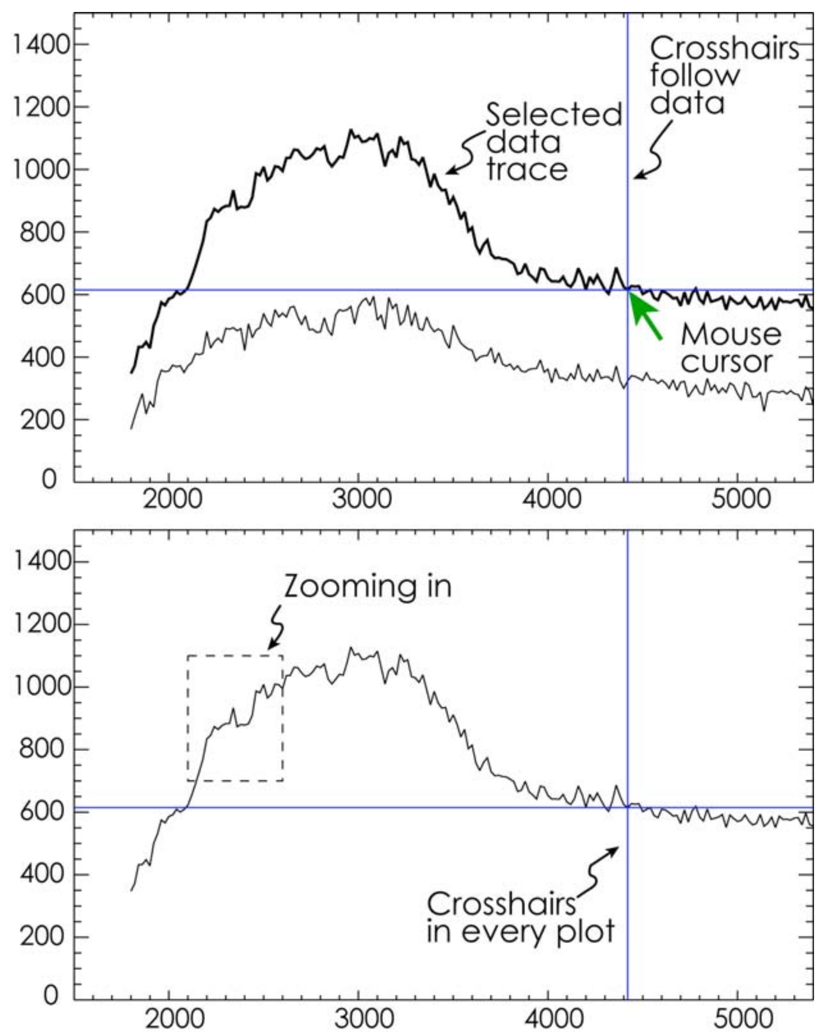
- **Not batch analysis and not a needle in a haystack problem**
  - Rapid near-real-time analysis of many measurments
- **More informed decisions result in better experiments**
  - The collaborative control room

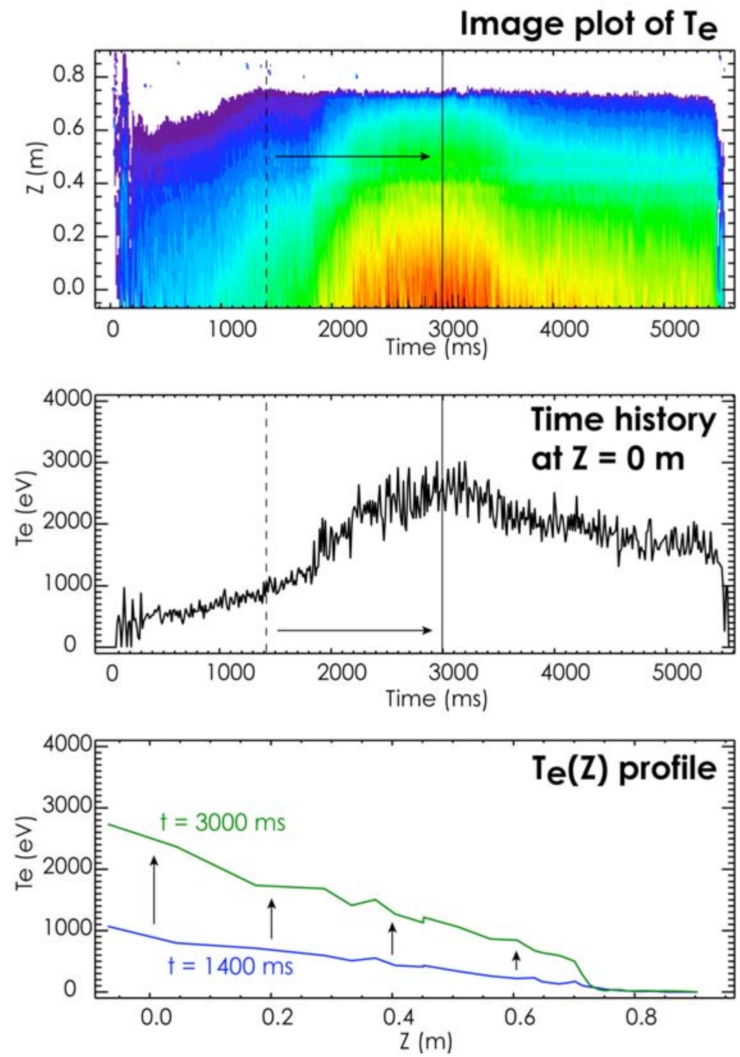
Fusion Tokamaks have Extensive Diagnostics

DIII-D NATIONAL FUSION FACILITY

GENERAL ATOMICS

Interactive Graphic Libraries have been Developed



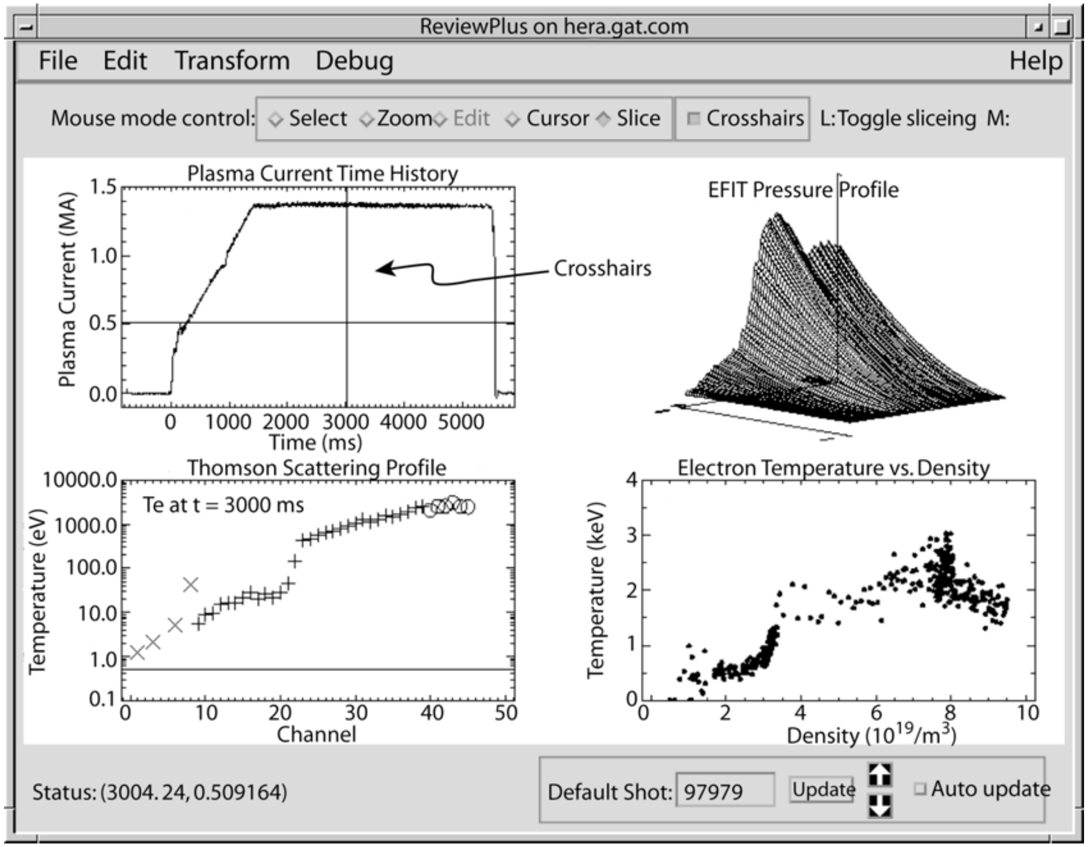


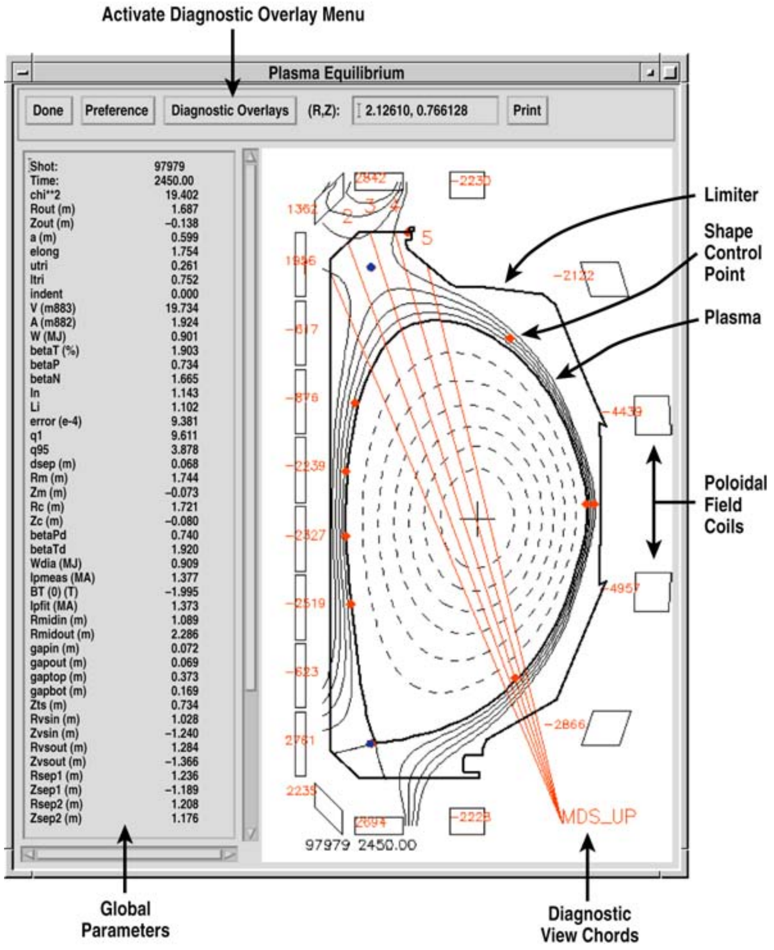
Move crosshairs

Automatic crosshairs

Profile slice changes

Custom Applications Allow Detailed Scientific Analysis Between Pulses

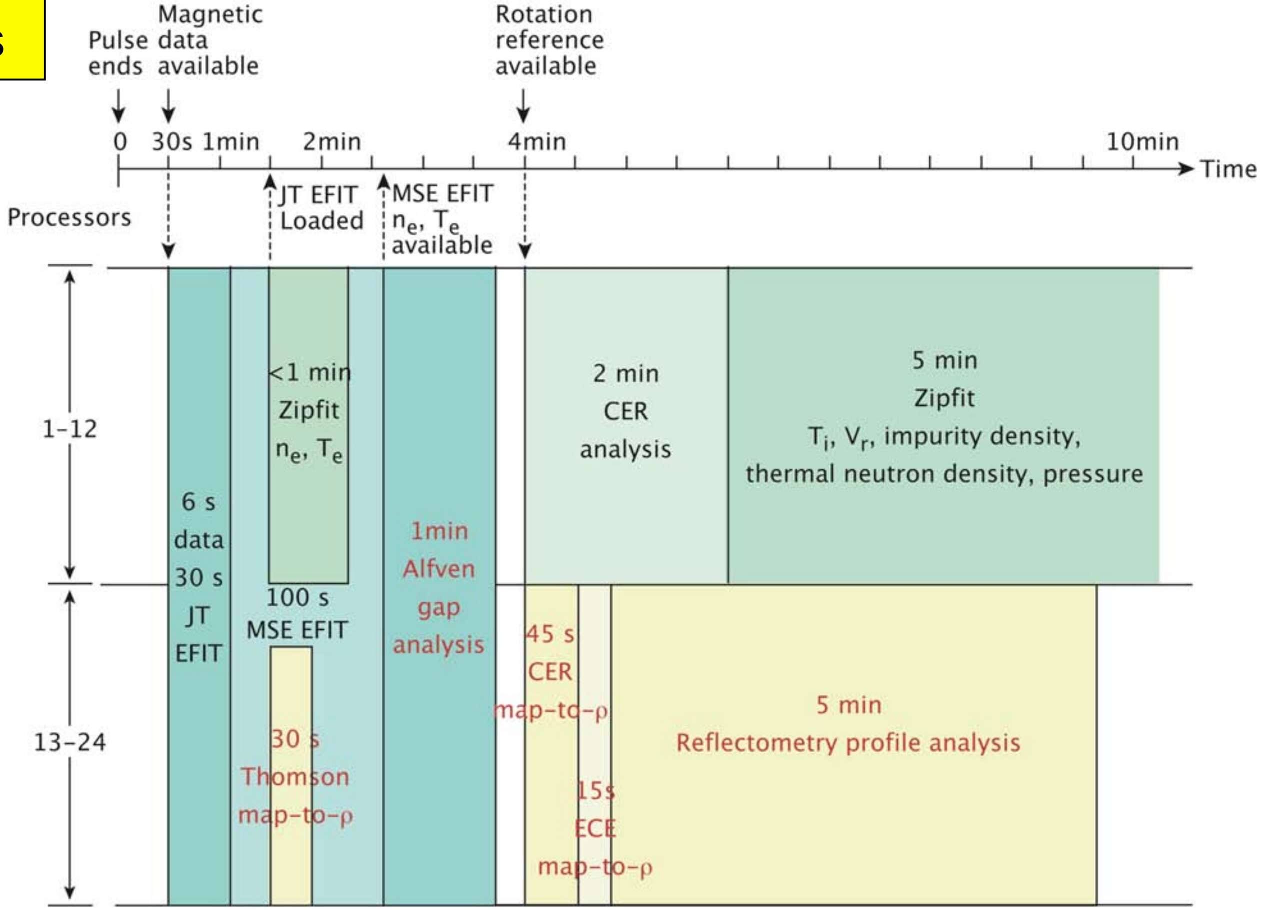




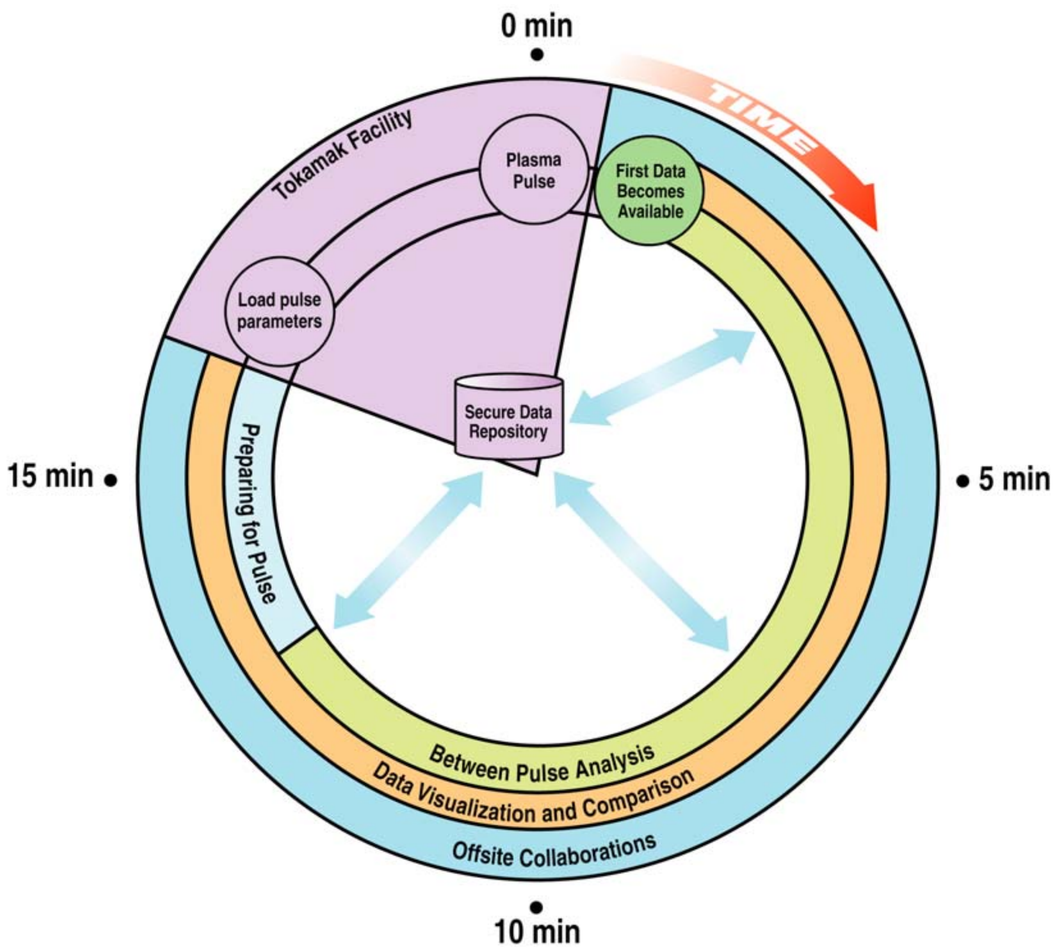
Between Pulse Data Analysis Critical to Decision Making

Dedicated Linux Clusters for between Pulse Analysis

24-CPU DIII-D Linux Cluster



Experimental Fusion Science is an Endless Cycle of Analysis and Decision Making



4. What Tools are being Used to Assist the International Collaborative Nature of Fusion Research?

A wide variety of techniques are being investigated and used but in the United States the SciDAC Program funded FusionGrid: A collaboratory pilot program for fusion energy research.

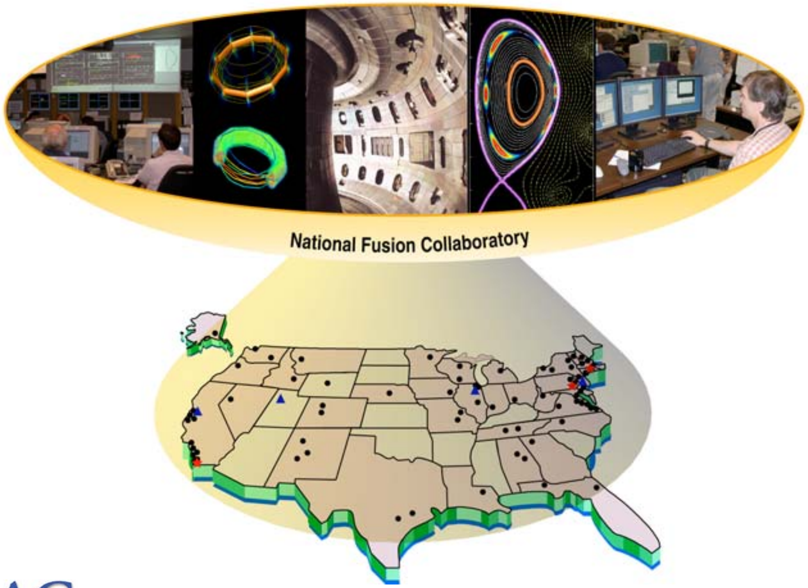
The National Fusion Collaboratory Project (FusionGrid)

- **Funded by the US DOE under the SciDAC Program (2001-2006)**
  - A distributed team: C-Mod, DIII-D, NSTX; ANL, LBL, PCS, Utah
  - Started as a pilot project but has transitioned to production usage

- **Unify distributed MFE research into a U.S. Virtual Organization**

GOALS

- More efficient use of experimental facilities
- Integrate theory & experiment
- Facilitate multi-institution collaboration
- Create standard tool set



ASCR

SciDAC Scientific Discovery through Advanced Computing

FusionGRID www.fusiongrid.org

The Vision for the NFC's Technologies: Optimize the Most Expensive Resource – People's Time

- **Remote operation of experiments**
  - Near term: US Domestic program
  - Longer term internationally: e.g. JET, EAST and KSTAR

- **Data, Codes, Analysis Routines, Visualization Tools should all be thought of as network accessible services (SOA)**
  - Access is stressed rather than portability
  - Transparency and ease of use are crucial elements
  - Not CPU cycle scavenging or "distributed" supercomputing

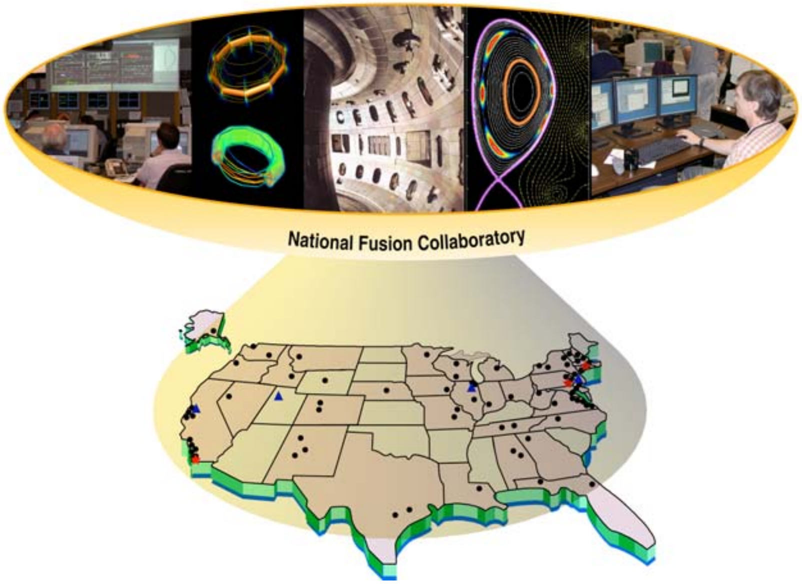
- **Shared security infrastructure with distributed authorization and resource management**
  - Ease of use: "security with transparency"
  - Tools for authentication, authorization, and resource management

FusionGrid: Unified Security Model with Data Access

- **Authentication: PKI via X.509 certificates**
  - FusionGrid CA & RAs
  - Centralized certificate management
  - Onetime login

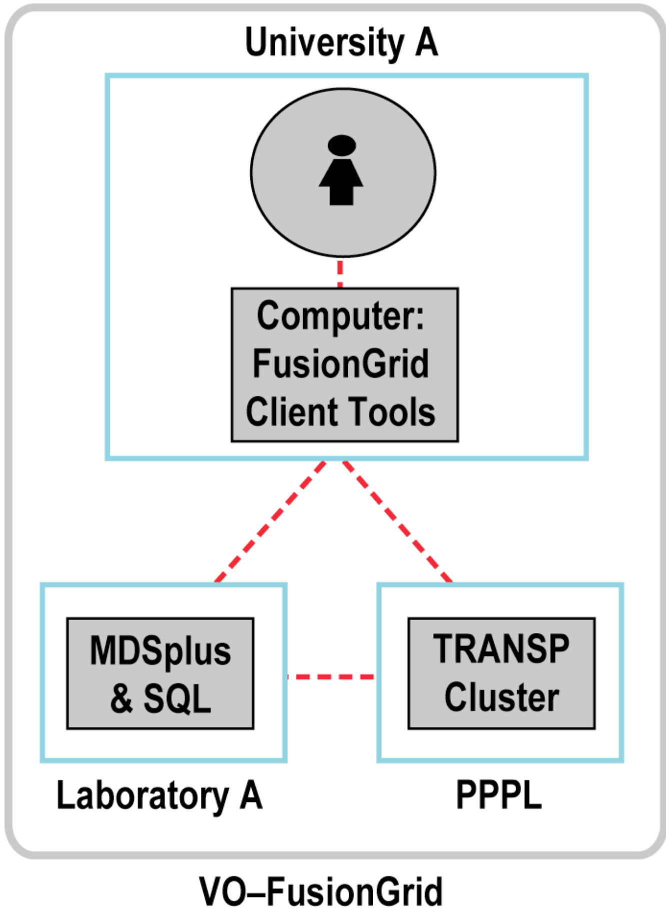
- **Authorization: Centralized ROAM**
  - Controlled by resource providers
  - More secure & easier to use

- **Data: Secure via MDSplus**
  - Client-server model
  - Not file transfer



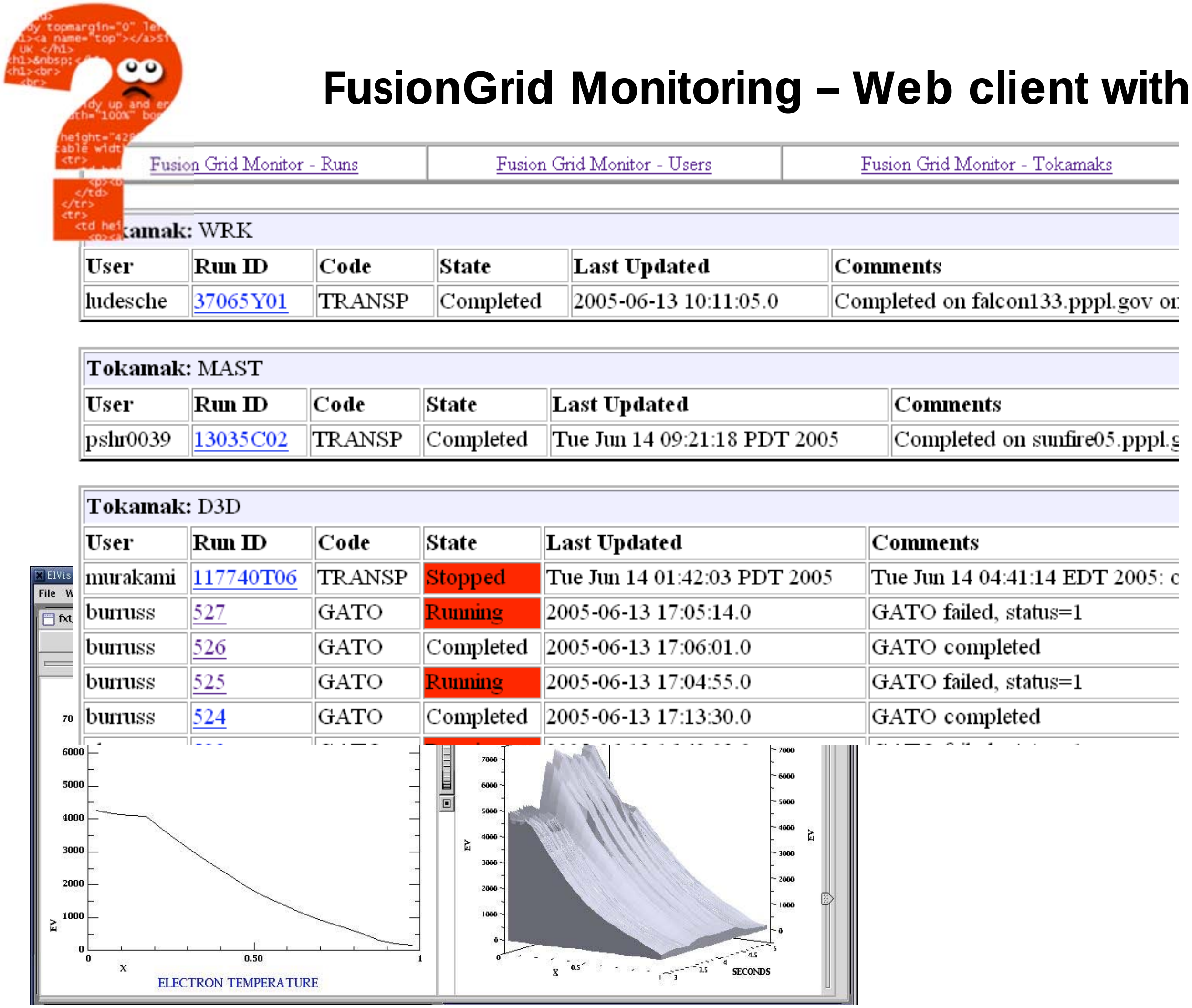
Successful Grid Computing for Fusion Science

- **The U.S. TRANSP Service**
  - 7,500 cases, 50,000 CPU hours
  - 10 fusion experimental machines
- **Centralized expertise for better support**
  - Debugging, maintenance, monitoring
- **Reduced administrative work at other labs**
  - Smaller sites to use bigger codes
- **Model for other codes**
  - GATO, ONETWO, ELFresco, FWR, GENRAY/CQL3D



FusionGrid Monitoring (FGM) for Scientists

FusionGrid Monitoring – Web client with real-time graphics


Fusion Grid Monitor - Runs | Fusion Grid Monitor - Users | Fusion Grid Monitor - Tokamaks

Tokamak: WRK

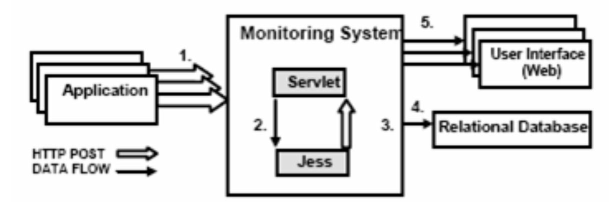
User	Run ID	Code	State	Last Updated	Comments
ludesche	37065Y01	TRANSP	Completed	2005-06-13 10:11:05.0	Completed on falcon133.pppl.gov or

Tokamak: MAST

User	Run ID	Code	State	Last Updated	Comments
pshr0039	13035C02	TRANSP	Completed	Tue Jun 14 09:21:18 PDT 2005	Completed on sunfire05.pppl.g

Tokamak: D3D

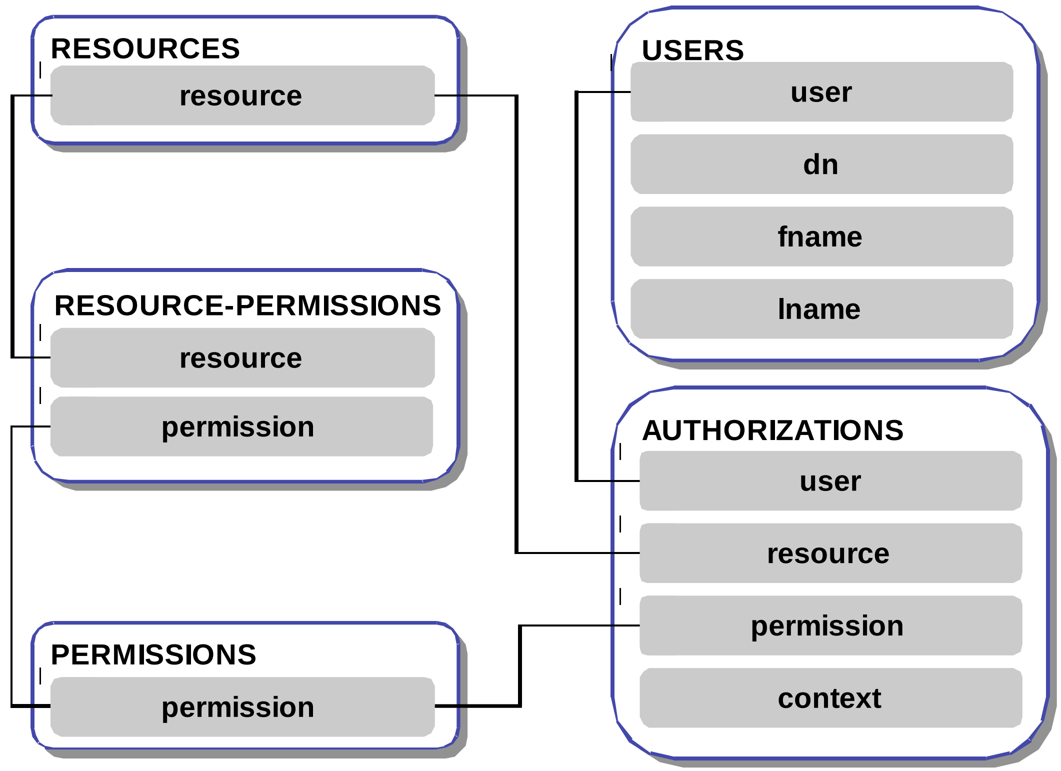
User	Run ID	Code	State	Last Updated	Comments
murakami	117740T06	TRANSP	Stopped	Tue Jun 14 01:42:03 PDT 2005	Tue Jun 14 04:41:14 EDT 2005: c
burruss	527	GATO	Running	2005-06-13 17:05:14.0	GATO failed, status=1
burruss	526	GATO	Completed	2005-06-13 17:06:01.0	GATO completed
burruss	525	GATO	Running	2005-06-13 17:04:55.0	GATO failed, status=1
burruss	524	GATO	Completed	2005-06-13 17:13:30.0	GATO completed



- **Web browser client**
- **Java servlet**
- **Expert system**
  - Registered posts
- **Relational database**

A New Authorization System (ROAM) Deployed

- Resource Oriented Authorization Manager
- Focus on resources
- A resource can be a code, a database, an entire site
- Access via Web client
- Empower stakeholders to specify types of permissions



Access Grid: Real Time Complex Communication

Scientific Leadership of JET in UK from US

January 2004, San Diego

Remote Participation from JET to DIII–D

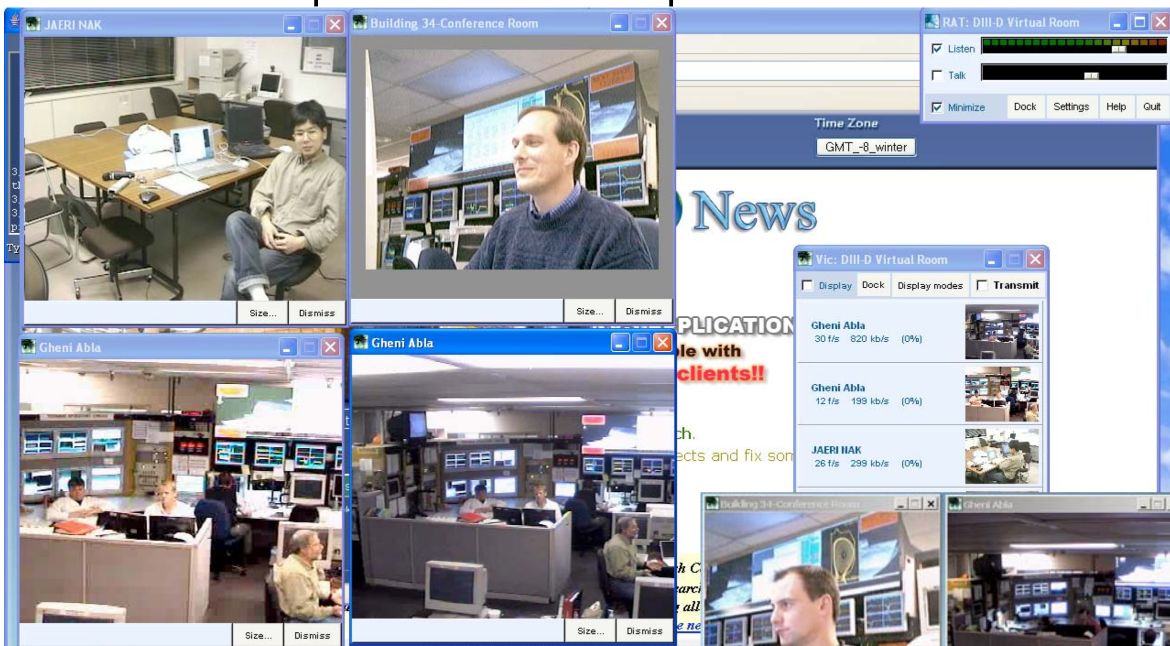
May 2004, DIII–D Tokamak Control Room

- **Has been used for seminars, working meetings, tokamak operations**
  – Linux, Windows, and Macintosh OS X; small to large immersive nodes

- **For tokamak operations, the collaborative control room**
  – Software framework: sharing humans data, applications, info

VRVS Being Used For Its Web Based Client

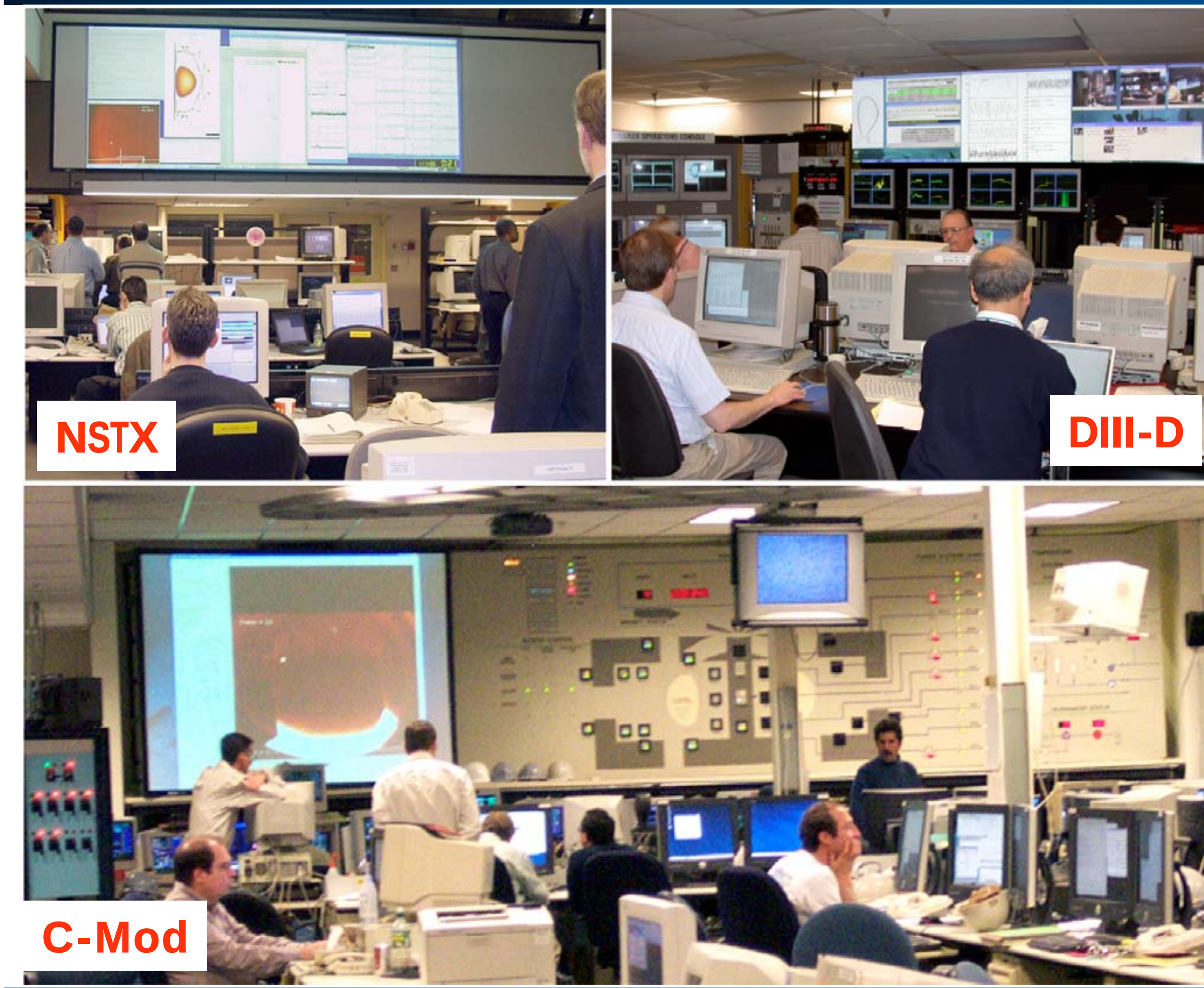
JAERI Participation in DIII-D Experiment March 2005



- **Closed source model limits customized expansion**
- **Web client is easy for scientists to get started**
  - 2004 - 2005:
    341 unique Users
    539 unique computers
    3480 total hits

JET Participation in DIII-D Experiment January 2005

Shared Display Walls Installed in Fusion Control Rooms



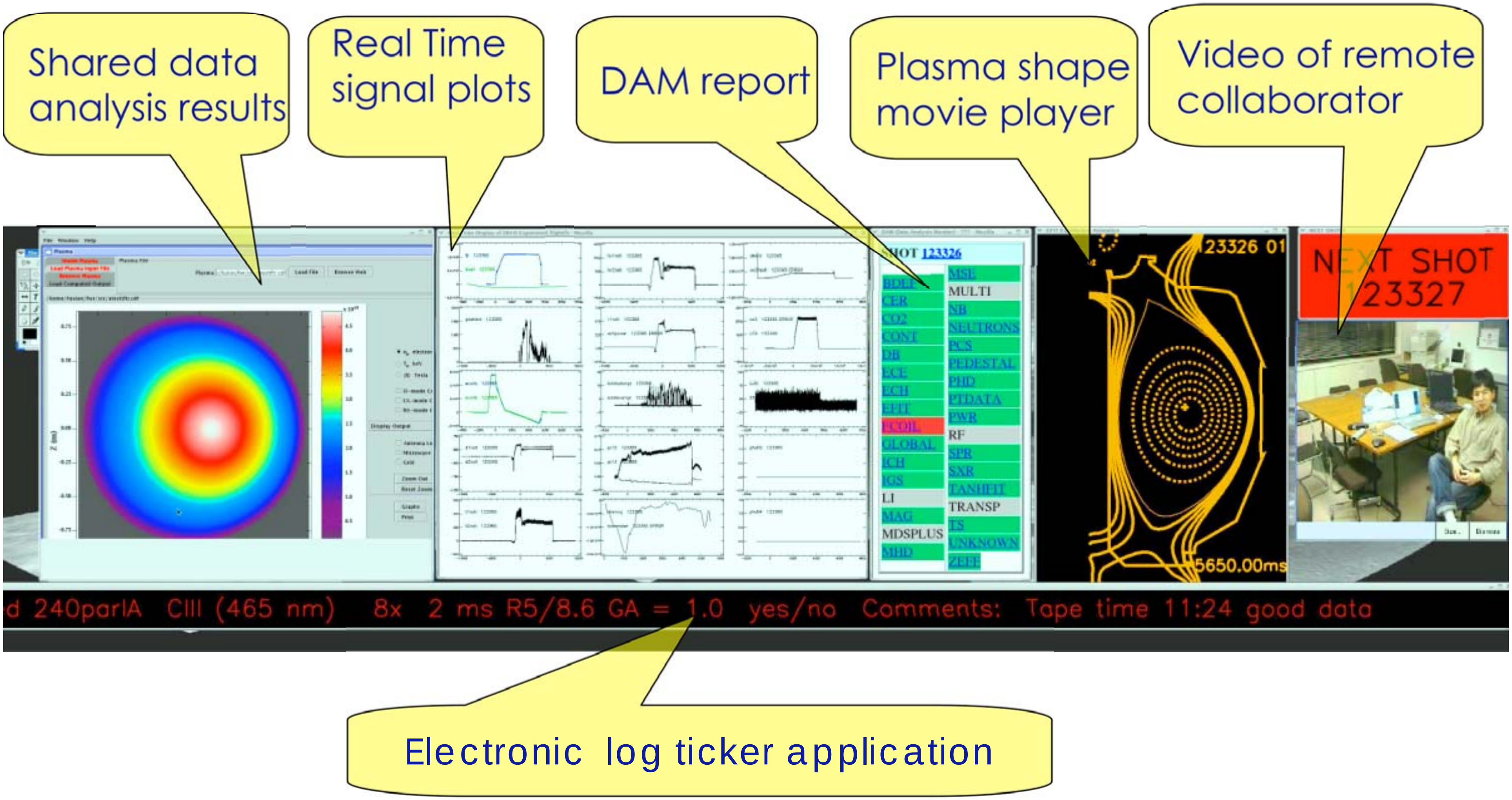
- **Customized Apps**
  - Display Walls
- **Sharing to the group**
  - Collocated
- **Sharing from off-site**
  - "See my graph"

Shared Display Wall has Enhanced Remote Participation in the control room

- **Shared display also has been helpful for remote participation**
  - Video streams of Access Grid and VRVS from remote sites displayed
  - Life sized video of remote participant gave a realistic impression

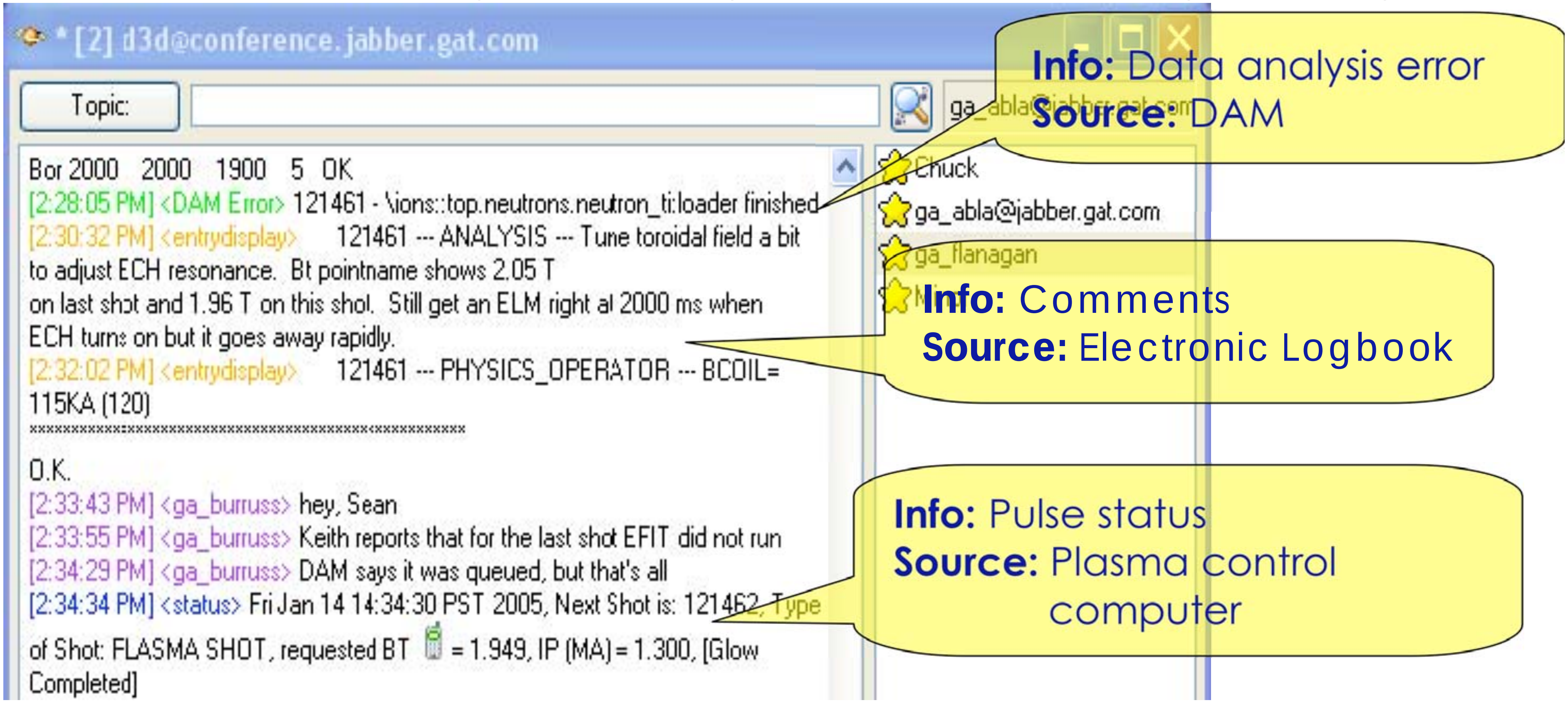
- **Controllable camera is for monitoring the shared display and control room**
  - Remote participants can control (Pan/Tilt/Zoom) through the web
  - Example: Remote scientist in UK shared his screen snapshot image to the wall and then used camera to make sure his image was displayed correctly

Shared Display Wall has Enhanced Remote Participation in the control room (2)



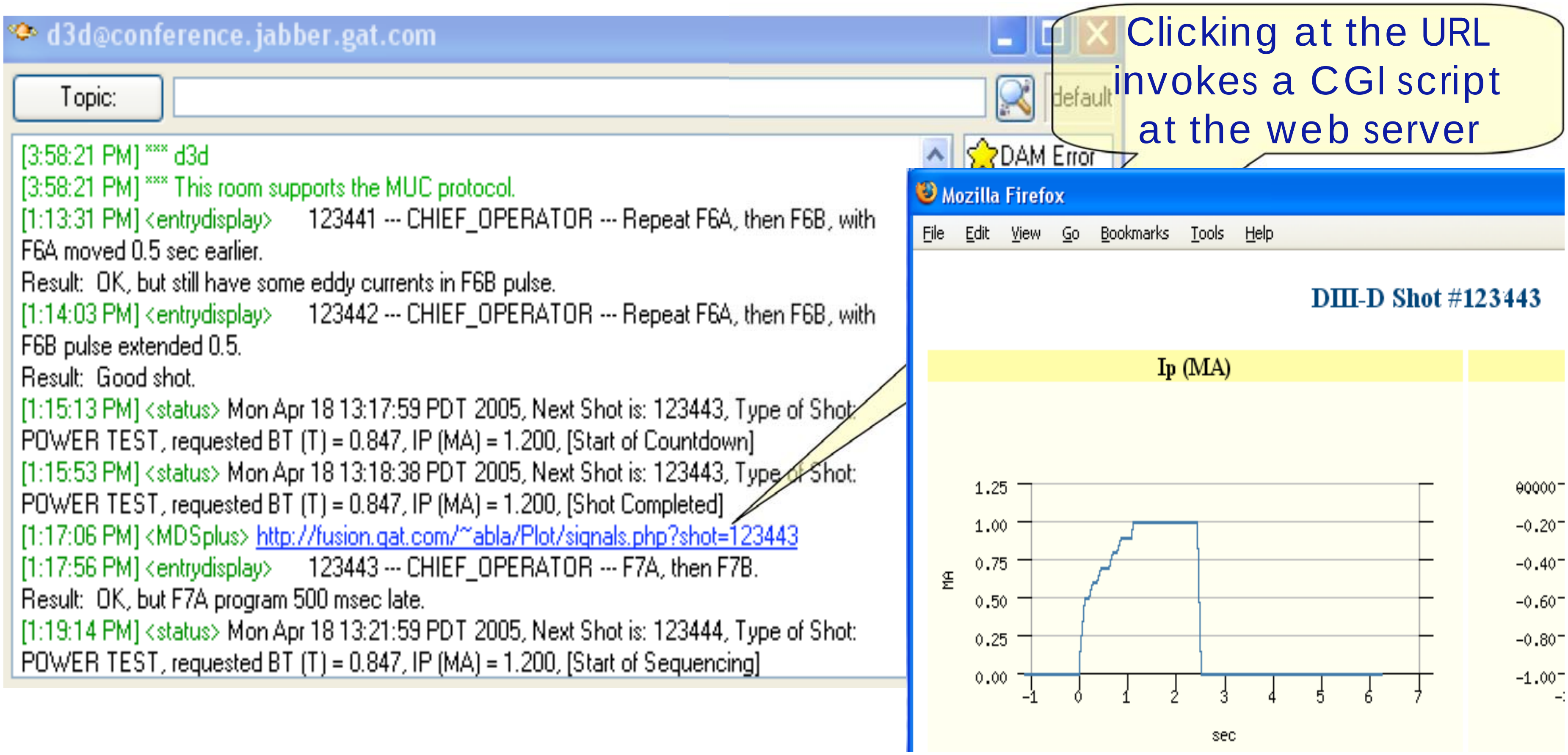
Instant Messaging has been Extended and Offers a Good Lightweight Collaboration Space

- **Jabber users automatically receives experiment status info**
  - C/C++ based shared Jabber client library has been created
  - Several data analysis tools integrated with Jabber using this library

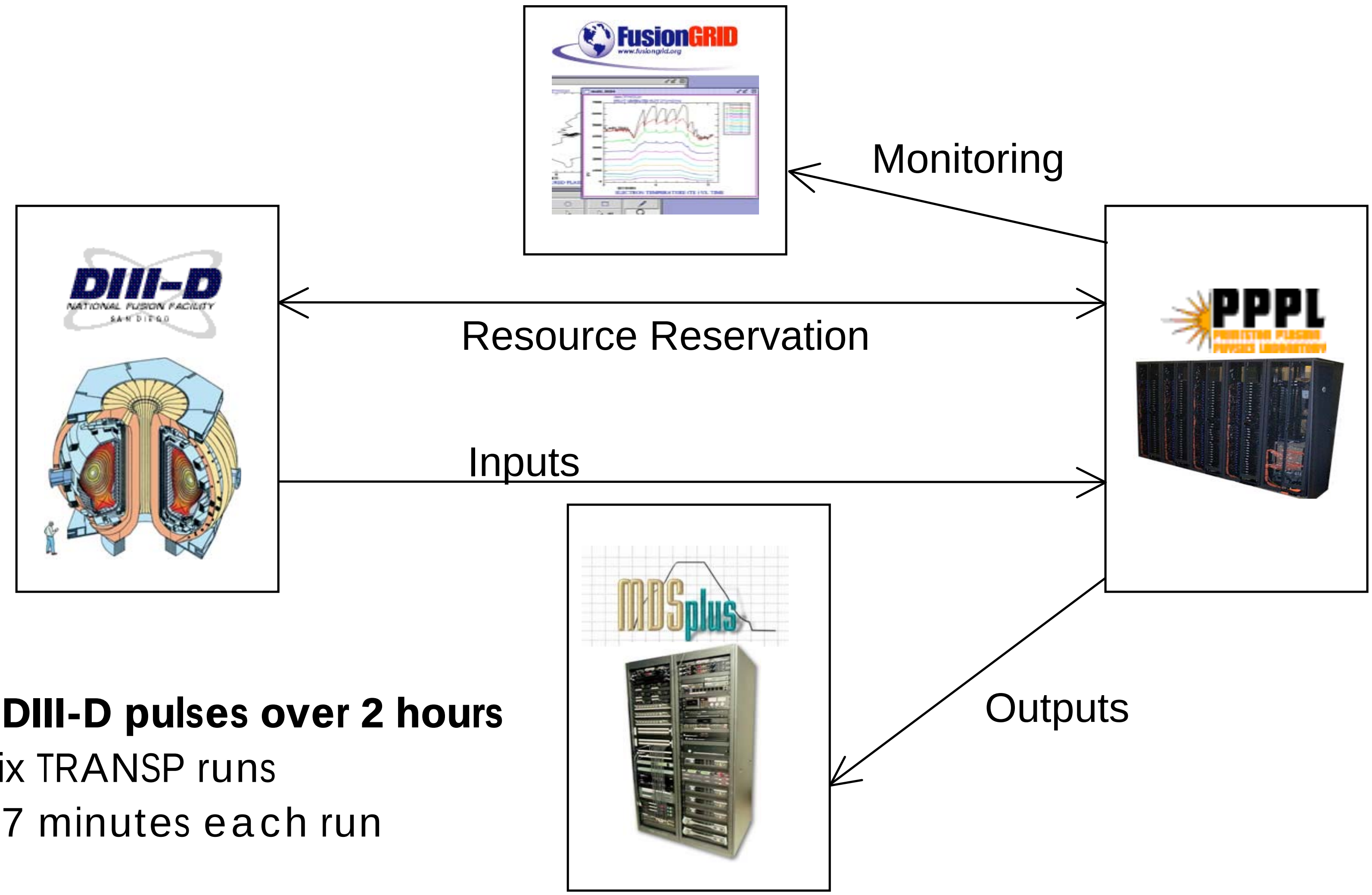


Instant Messaging has been Extended and Offers a Good Lightweight Collaboration Space (2)

- **2D visualization plots of selected signals automatically available**
  - The URL of webpage are posted at d3d Jabber room after each pulse



FusionGrid Tested for Between Pulse Data Analysis



- **Six DIII-D pulses over 2 hours**
  - Six TRANSP runs
  - ~7 minutes each run

5. How Closely Aligned are the Needs of FES and HEP?

The requirements of LHC@FNAL appear to be very similar to the needs of running ITER in France from the United States.

Challenges: International, Remote Participation

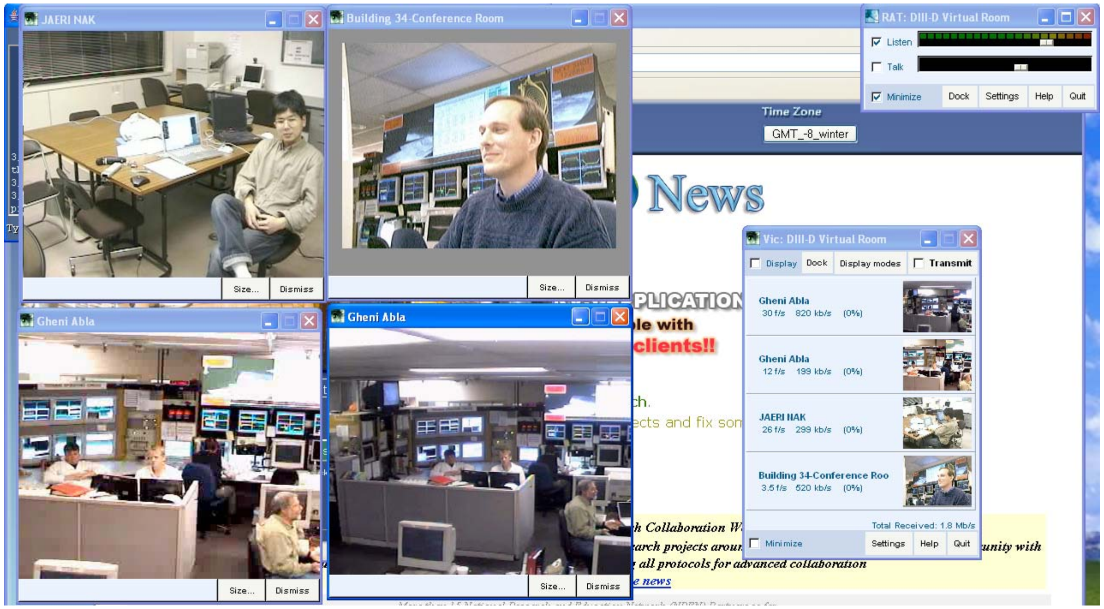
- **Scientists will want to participate in “live” experiments from their home institutions dispersed around the world**
  - View and analysis data
  - Manage ITER diagnostics
  - Lead experimental sessions (we are already breaking new ground here)
- **Requirements are full-time: before, after, and during an experiment**
- **Collaborations span many administrative domains**
  - Resource management - local control is essential
  - Trouble shooting/end-to-end problem resolution
- **Cyber security must be maintained, plant security must be inviolable**

Both Remote and Distributed have Challenges

- Informal interactions in the control room are crucial part of the research
- We must extend this into remote and distributed operations - lower barriers and friction
- Fully engaging remote participants is challenging
- Working as distributed team beyond just the experimental day
- The NFC Project has given the U.S. fusion program a good start

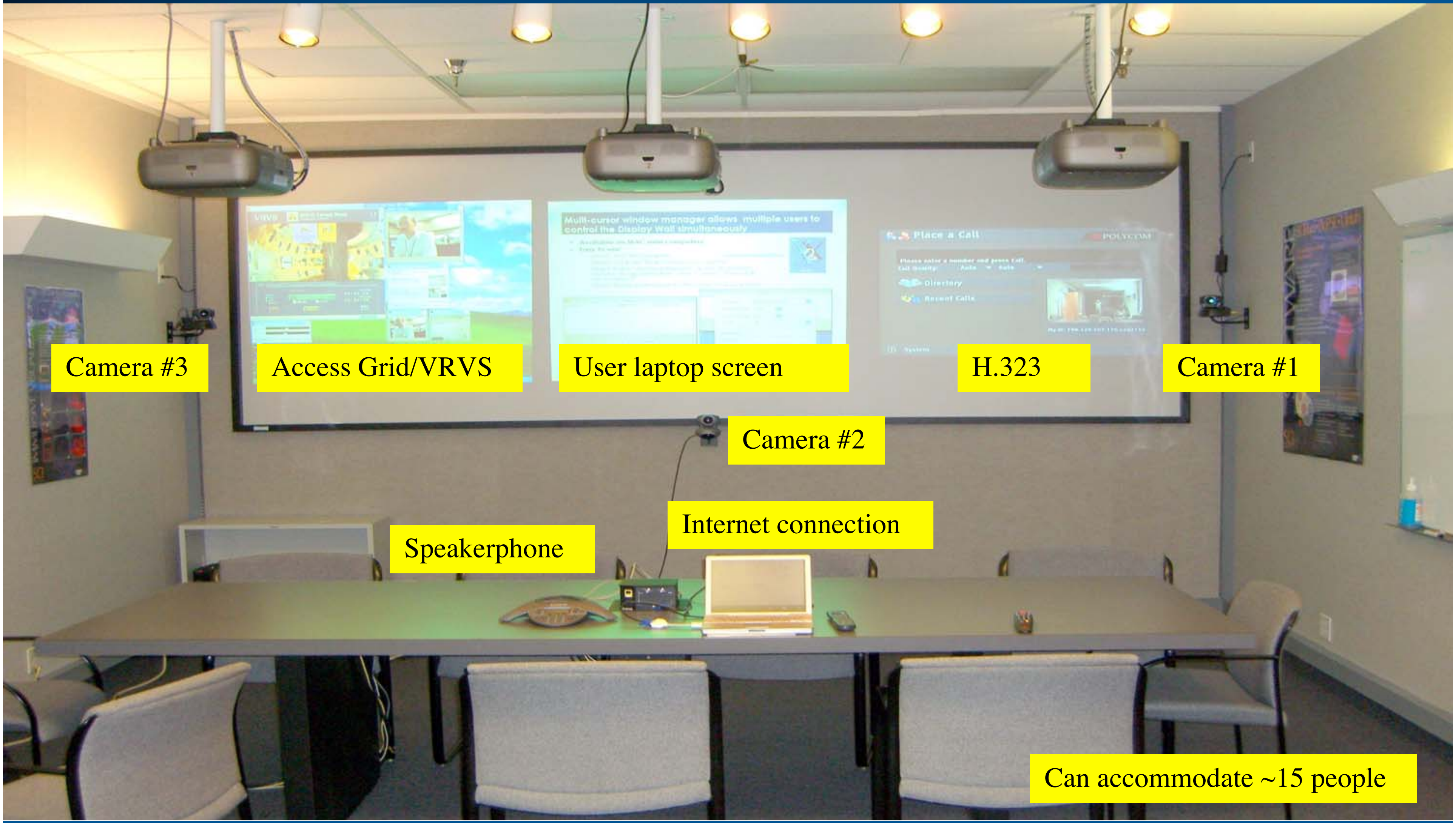
Components for Successful Remote Participation

- **Transparent remote access to data**
  - Secure and timely
- **Real-time information**
  - Machine status
  - Shot cycle
  - Data acquisition and analysis monitoring
- **Shared applications**



- **Provision for robust ad hoc interpersonal communications**
  - e.g. Informal discussion, time-critical contact
- **Provisions for robust structured communications**
  - e.g. Planning and review

Rooms in the Fusion Community are being Constructed for Remote Participation



LHC@FNAL Remote Operations Center

- Provides access to information in a manner similar to control rooms at CERN
- Where members of the LHC community can participate remote in CMS & LHC
- A communications conduit: CERN and North American LHC community
- An outreach tool where visitors can see current LHC activities

The Requirements of HEP for LHC@FNAL are Similar to the Needs of FES for ITER

- **Substantial requirements overlap between FES and HEP**
  - Support remote operations of experimental facilities, distributed code development, computing & visualization, planning activities
- **Access to mission critical controls data**
  - Continuous MDSplus
- **Sharable applications and displays**
  - LHC@FNAL providing access to LHC & CMS control rooms at CERN
- **Improvements in interpersonal and group communication integrated with extensible data services**
  - Role-based security and communication
- **Prototype control room for design efforts for International Linear Collider**

Potential Work Scope Divided into two Main Areas

1. **Collaborative Work Spaces**
   - Extensible Collaboration ad hoc & structured communication
   - Standards-based, modular, role-aware, presence-aware, web-friendly
   - Enhanced user agents: Access Grid and VRVS/EVO
   - Shared displays and applications

2. **Secure Computational Services**
   - Continue on FusionGrid's proven path of success
   - Long-pulse or continuous MDSplus
   - International FusionGrid interoperability:Russia (RDIG), Europe (EGE)
   - New computational services: CQL3D, GS2, TGLF, GKS, GCNM(P)

Security enhancements apply to both areas equally

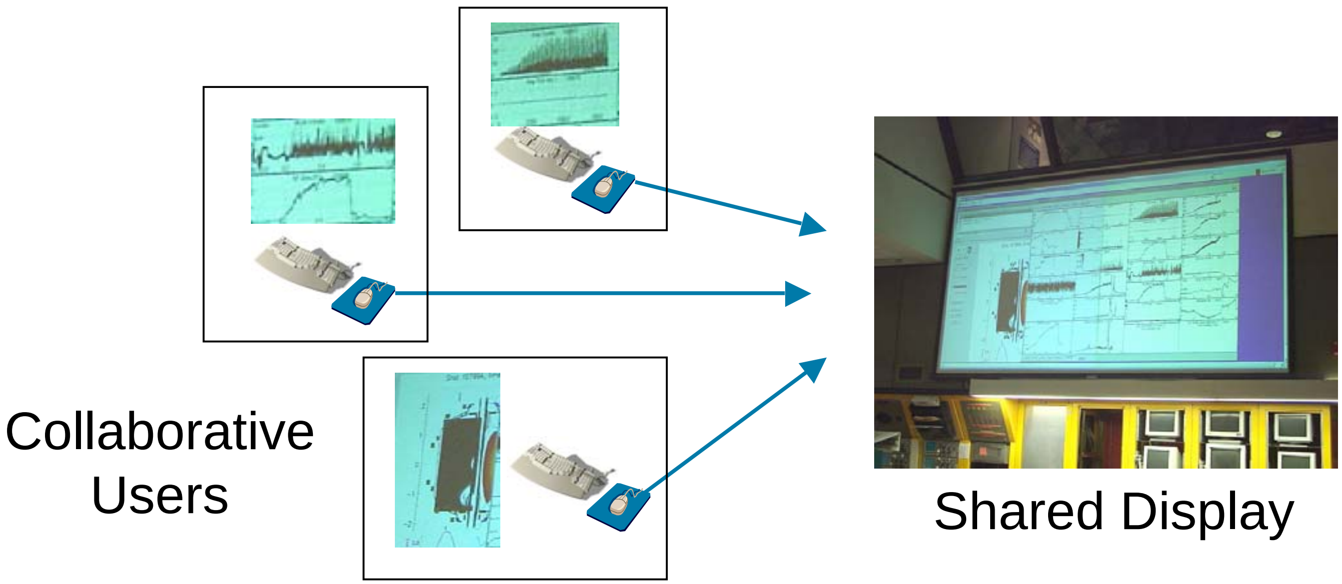
- UCAM: Enhanced easier-to-use security
- Integrated and role-aware

Remote Participation: Ad Hoc Communications

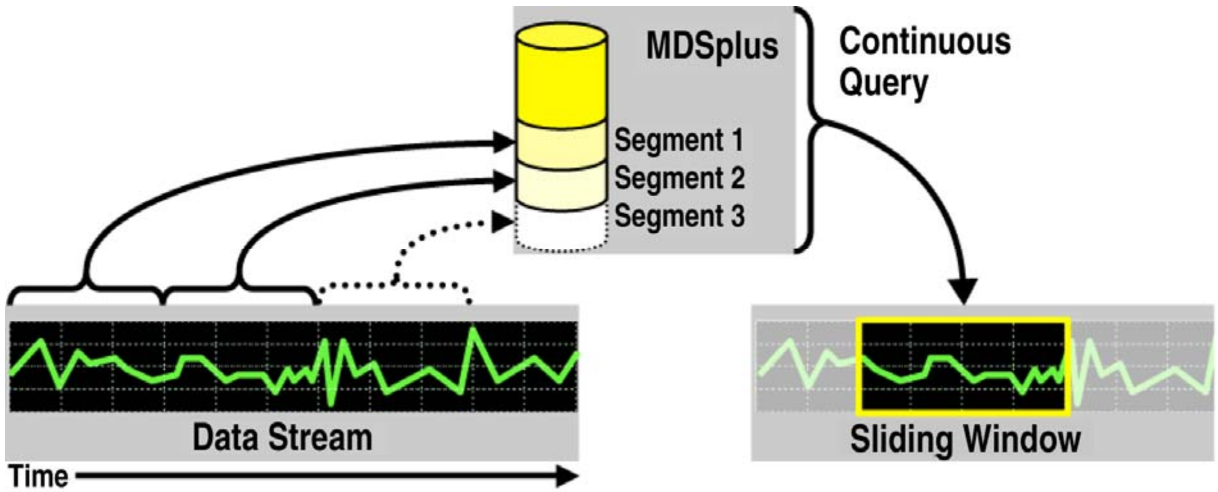
- **Exploit convergence of telecom and internet technologies (e.g. SIP)**
- **Deploy integrated communications**
  - Voice
  - Video
  - Messaging
  - E-mail
  - Data
- **Advanced directory services**
  - Identification, location, scheduling
  - “Presence”
  - Support for “roles”
- **AG & VRVS/EVO SIP Integrated**

Shared Displays and Applications

- **Distributed shared display walls**
  - Remote collaborative visualization: control rooms and desktops
  - Distributed shared display protocol required
  - Multi-party updates with appropriate security
- **Network bandwidth optimization**
  - Overcome network latency and bandwidth limitations
  - Combine intelligent caching with compression



Long-Pulse and Continuous MDSplus Data System



- **Requirements for concurrent writing, reading, analysis, visualization**
- **Data will span range >10^9 in significant time scales**
- **Will require efficient tools**
  - To browse very long records
  - To locate and describe specific events or intervals

Federated Security Implementation

- **Intrinsic conflict between transparency/ease-of-use and strong security**
  - Build on existing FusionGrid work
- **UCAM: User, Credential, and Authorization Manager**
  - Appropriate for international grids with heterogeneous authentication
  - OTP and dynamic firewalls
- **Federated Web Portals with PubCookie**
  - Single sign-on for a US based FusionGrid web system
- **Richer authorization policy**
  - Dynamic role-based authorization
- **FusionGrid interoperability**
  - Russian Data Intensive Grid (RDIG) and European Data Grid (EDG)

Work of Interest to DOE/SC: OFES, OHEP, OASCR

- **OFES**
  - Maximize the benefit of ITER to the U.S.
  - Longer term internationally on a path to ITER: EAST and KSTAR
  - Near term: US Domestic program
  - Build an integrated simulation team for the FSP

- **OHEP**
  - Successful remote collaboration from the U.S. on LHC via LHC@FNAL
  - Technologies required for a successful ILC

- **OASCR**
  - Collaborative technologies critical to many SC programs
  - Collaboration with CS community has clearly aided SciDAC program
  - Needs of science programs go beyond petascale computing

Concluding Comments

- **The NFC Project is implementing & testing new collaborative technology**
  - Attacking problems defined by fusion scientists
  - FusionGrid services being used to benefit daily FES research
- **Service oriented computation on FusionGrid has proved successful**
  - Optimize the most expensive resource - people's time
- **Clear vision & work scope forward to the Collaborative Control Room**
  - Real-time support for experiments is critical
  - Clear software enhancements required for success
- **Helps to position US to exploit ITER**
- **Requirements of HEP appear similar and justify joint research**
- **Path to success that benefits FES and HEP and leverages OASCR work**