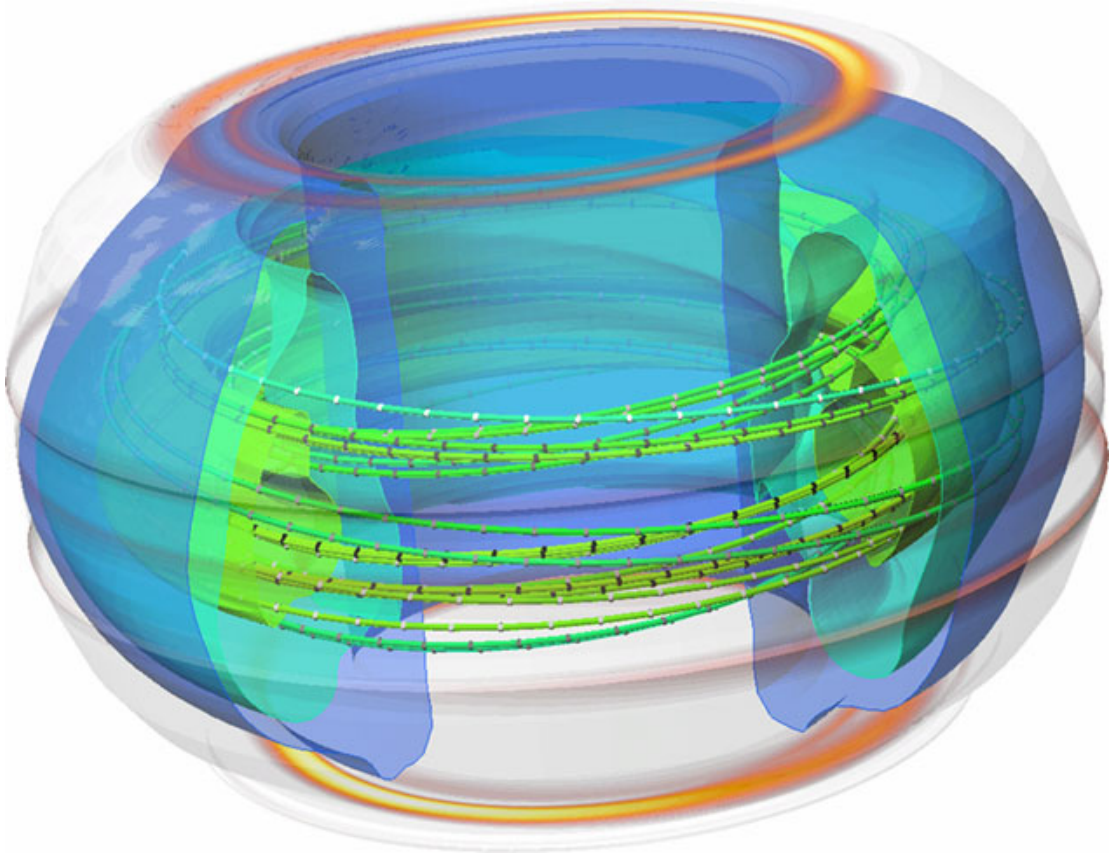
GRID PERFORMANCE AND FUSION SCIENCE



PRESENTED BY
JUSTIN BURRUSS

GRID PERFORMANCE WORKSHOP 2005

EDINBURGH, SCOTLAND
JUNE 22-23, 2005

burruss@fusion.gat.com
http://web.gat.com/~burruss/

FusionGRID
www.fusiongrid.org

ACKNOWLEDGEMENTS

- **U. S. Department of Energy**
  - OFES & OASCR (SciDAC)

- **DIII-D National Fusion Facility**
  - Operated by General Atomics

- **FusionGrid collaborators**
  - MIT, PPPL, LBL, ANL, Utah CS, Princeton CS

- **National eScience Centre**

OUTLINE

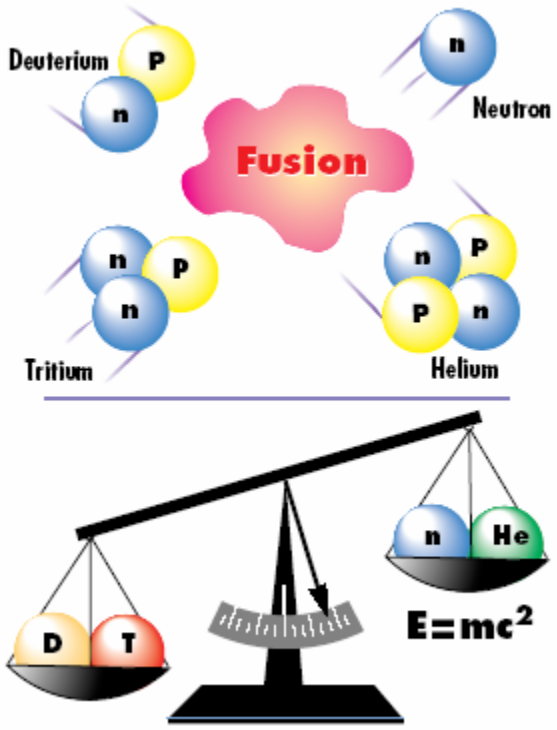
- **Background**
  - What is Fusion?
  - How is Fusion Science advanced?
- **Application**
  - Between-shot analysis
  - Monitoring & failure recovery
  - Moving from LAN to WAN
- **Performance Metrics**
  - Data analysis throughput
  - Reliability
- **The Perfect Performance Monitoring & Management System?**

KEY POINTS

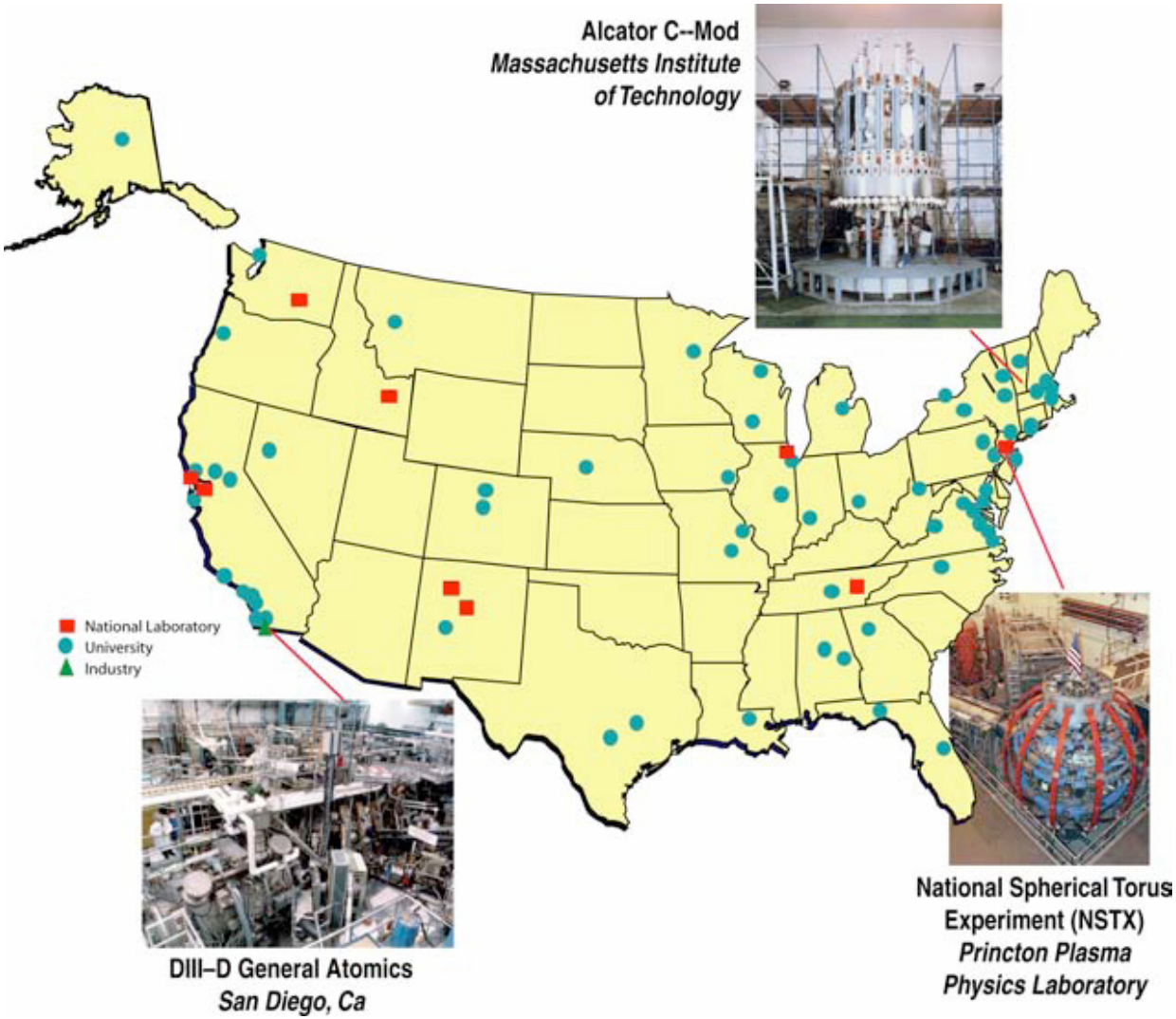
- **Fusion science seeks a new power source and is advanced by experiments on fusion devices (tokamaks) located worldwide**
- **Experiments require rapid data analysis**
  - Not batch analysis
  - Not "needle in the haystack" searching
  - Thousands of measurements that build a coherent picture
- **Data analysis throughput and reliability are the important performance metrics**
- **Moving data analysis from LAN to WAN brings many benefits, but introduces new performance issues**
- **We look forward to collaborating to make this happen**

FUSION SCIENCE SEEKS ENVIRONMENTALLY & ECONOMICALLY ATTRACTIVE POWER PLANT

- **Fusion is when you combine two atoms into one atom**
- **Energy is released from this fusion reaction**
- **An attractive power source**
  - Abundant fuel available to all nations
  - Environmentally friendly
  - No proliferation risk
  - Can't blow up/melt down
  - Not subject to weather/seasonal issues
  - Concentrated relative to wind/solar



FUSION RESEARCH TAKES PLACE ACROSS THE U. S.



FUSION RESEARCH TAKES PLACE WORLDWIDE

FUSION RESEARCH TODAY IS A TEAM EFFORT

- 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)

DIII-D has active collaborators on four continents

FusionGRID
www.fusiongrid.org

FUSION RESEARCH WILL CONTINE TO BE A TEAM EFFORT

Seoul
Korean Participant Team

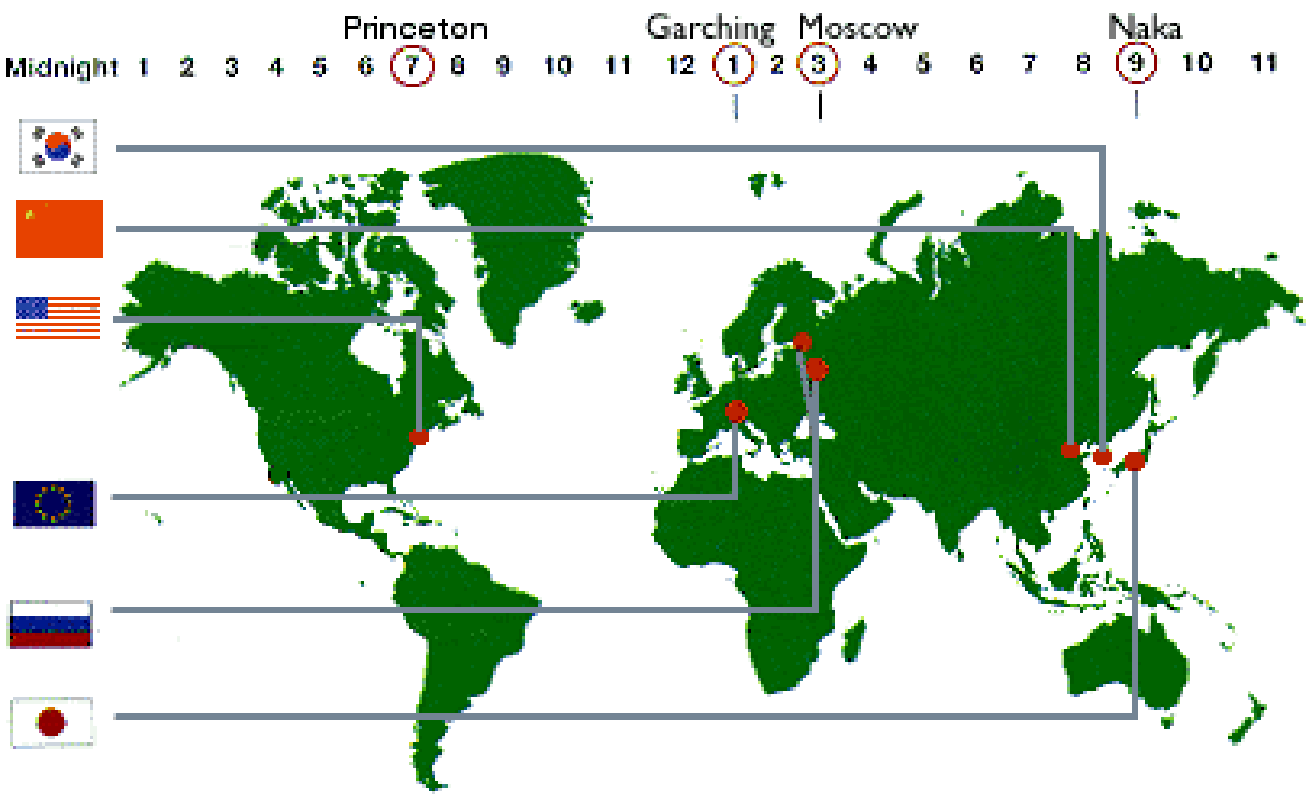
Beijing
Chinese Participant Team

Princeton
US Participant Team

Garching Joint Work Site
International Team
European Participant Team

Moscow/St.Petersburg
Russian Participant Team

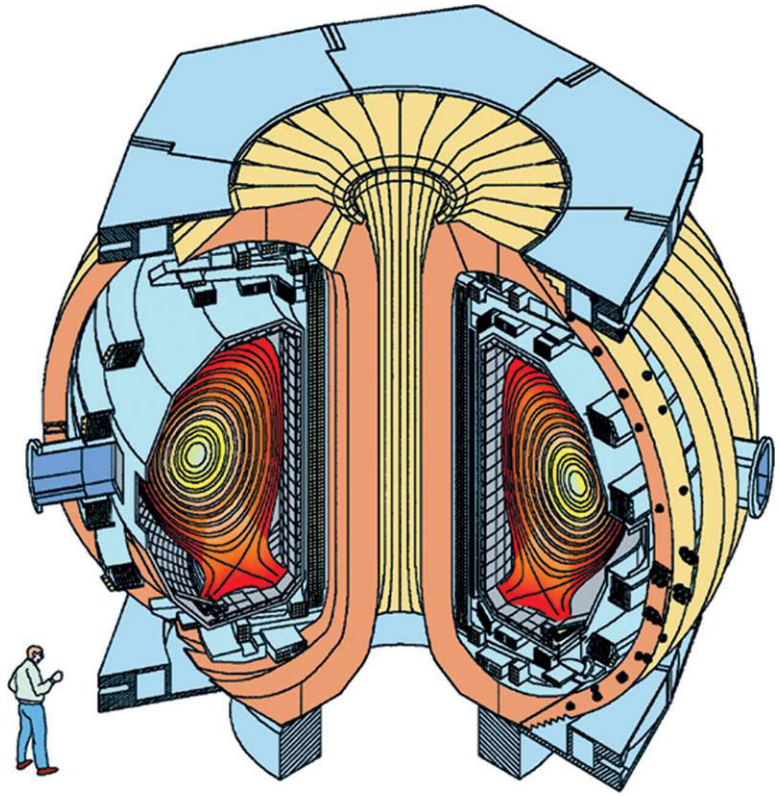
Naka Joint Work Site
International Team
Japanese Participant Team



The Six ITER Partners

FUSION EXPERIMENTS CONSIST OF PULSES OR "SHOTS"

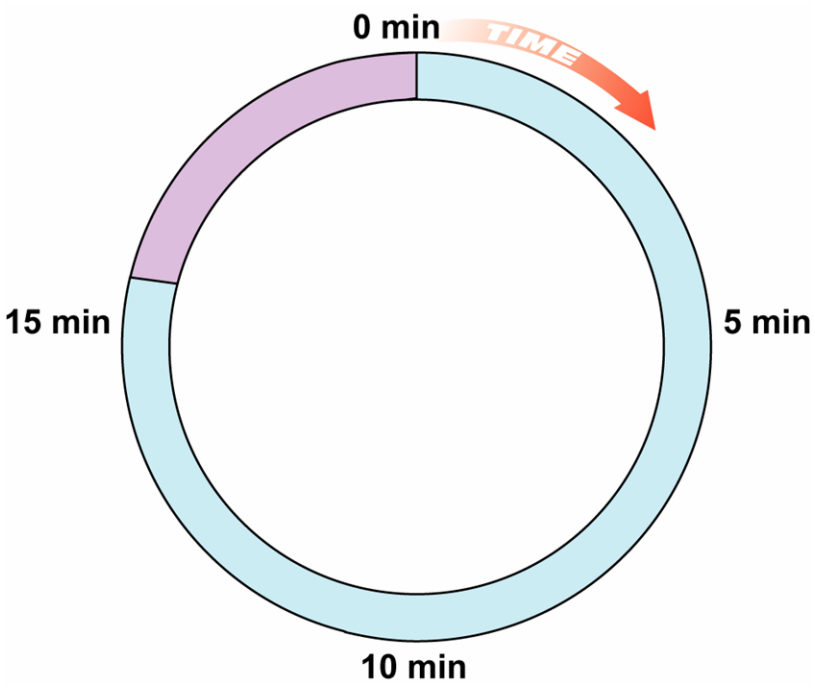
- **Fusion devices called tokamaks operate in a pulsed mode**
  - Turn on, then back off again
  - Called a "pulse" or "shot"
- **Raw data from sensors is collected during each shot**
- **Raw data is analyzed by codes, producing analyzed data**
- **Summary data is collected from analyzed and raw data**
- **The scientists can "see" what happened during a shot by looking at this raw, analyzed, and summary data**



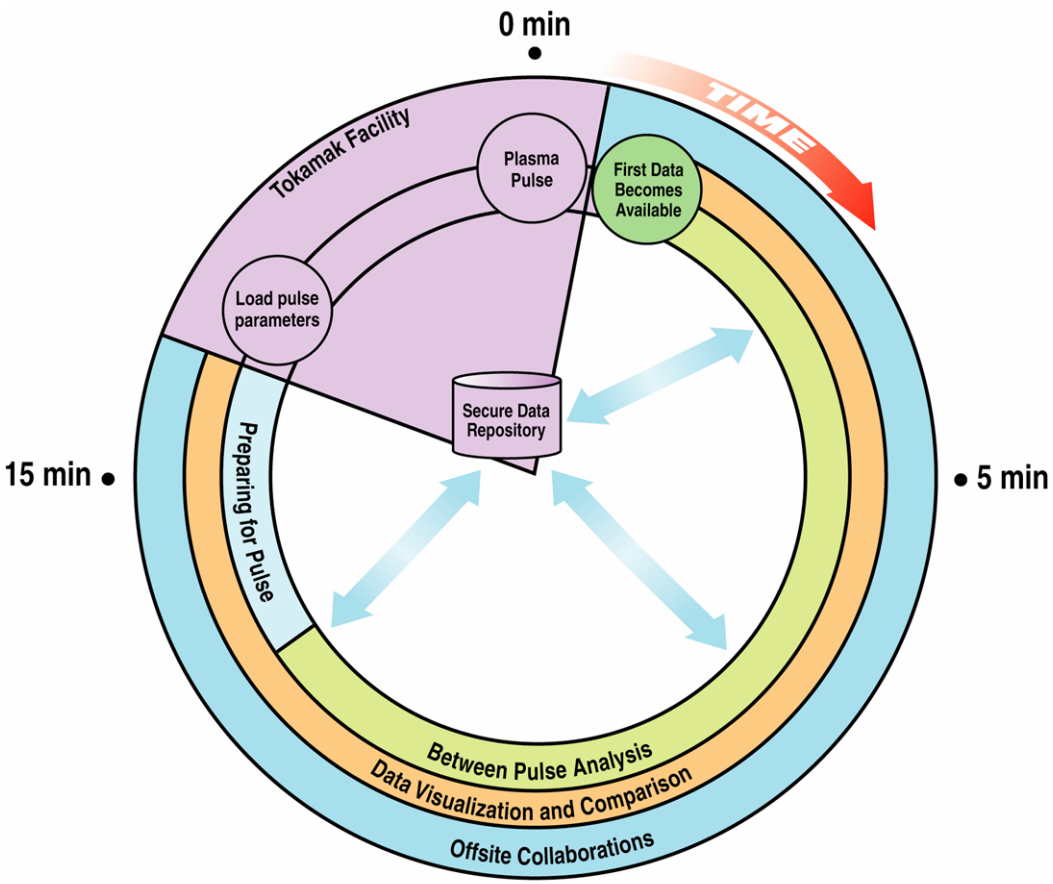
DIII-D Tokamak

DECISIONS MADE DURING FUSION EXPERIMENTS REQUIRE RAPID DATA ANALYSIS ON TOKAMAK PULSES

- **Fusion experiments are not done in a batch mode: scientists "take a shot", analyze the results, and move on to the next shot throughout the experiment**
- **This requires rapid data analysis**
  - Scientists view results of previous shot to make informed decisions about how to configure the next shot
  - Decisions must be "locked in" at some point before shot
- **Each tokamak shot is like a space shuttle launch—but done every 20 minutes**



EXPERIMENTAL FUSION SCIENCE IS AND WILL CONTINUE TO BE A DEMANDING ACTIVITY

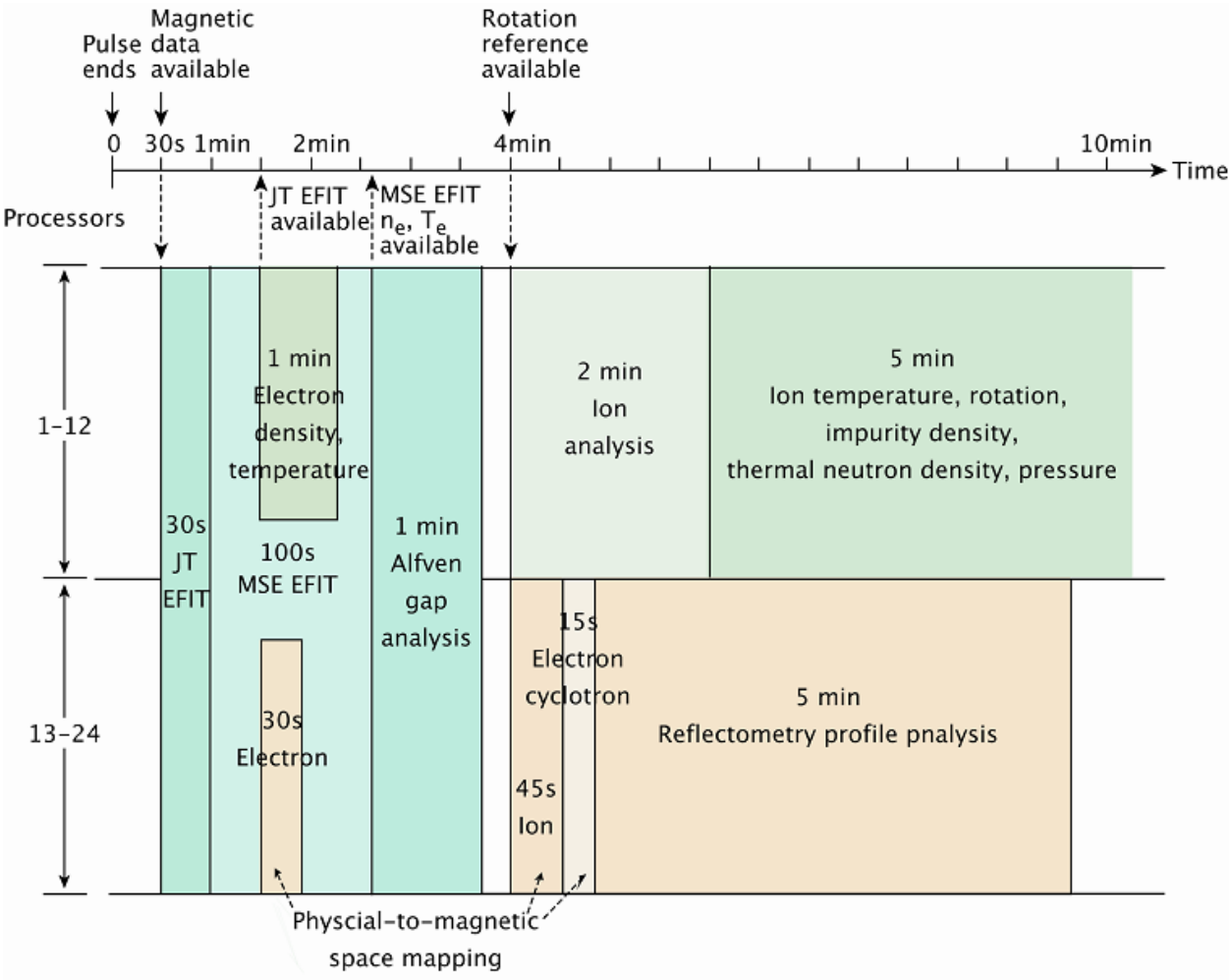


Thousands of measurements and a team of specialists needed to build a coherent picture of what happened during a shot

FusionGRID
www.fusiongrid.org

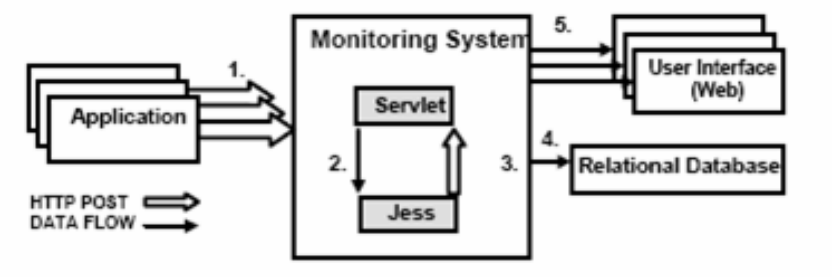
TODAY MOST BETWEEN-SHOT ANALYSIS IS LOCAL

24-CPU DIII-D Linux Cluster



DAM USED FOR LOCAL BETWEEN-SHOT MONITORING

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



GENERAL ATOMICS

DAM (Data Analysis Monitor) - Current Sh

Data Analysis Monitor - Current Shot | Data Analysis Monitor - All Shots | Data Analysis Monitor - Code Group | Data Analysis Monitor - Errors Only

SHOTS
116790
116789
116788
116787
116786
116785
116784
116783
116782
116781
116780

SHOT 116790 DATE:2004-01-29

Time	Name	Code	R. Officer	Logfile
16:49:56	\NEUTRALS::ASDEX:LOADER FINISHED	load_asdex	schacht	View Log
16:49:54	\TOP.D3DRDB:DBLOADER1 FINISHED	db_update_basic	burruss	View Log
16:49:53	\NEUTRALS::GASFLOW:LOADER FINISHED	load_gasflow	mdsadmin	View Log
16:49:52	\NEUTRALS::ASDEX:LOADER STARTING	load_asdex	schacht	View Log
16:49:51	\NEUTRALS::GASFLOW:LOADER STARTING	load_gasflow	mdsadmin	View Log
16:49:48	\TOP.D3DRDB:DBLOADER1 STARTING	db_update_basic	burruss	View Log
16:49:06	\SPECTROSCOPY::TOP.PHD.D_ALPHA_FAST:LOADER FINISHED	load_phd	burruss	View Log
16:48:58	\SPECTROSCOPY::TOP.PHD.D_ALPHA_FAST:LOADER STARTING	load_phd,1		View Log
16:48:41	MSE_FINISHED	load_mse.csh	makowski	View Log
16:48:34	\SPECTROSCOPY::TOP.PHD:LOADER FINISHED	load_phd	burruss	View Log
16:48:33	\OPERATIONS::TOP.BCOIL.CPTWIST:LOADER FINISHED	load_cptrad	taylorpl	View Log
16:48:27	\NB::TOP:LOADER FINISHED	load_nb	burruss	View Log
16:48:26	\OPERATIONS::TOP.BCOIL.CPTWIST:LOADER STARTING	load_cptrad	taylorpl	View Log
16:48:23	\OPERATIONS::TOP.FCOIL.CHOPPERS:LOADER FINISHED	load_choppers	burruss	View Log
16:48:22	\NB::TOP:LOADER STARTING	load_nb	burruss	View Log
16:48:17	MSE_STARTING	load_mse.csh	makowski	View Log

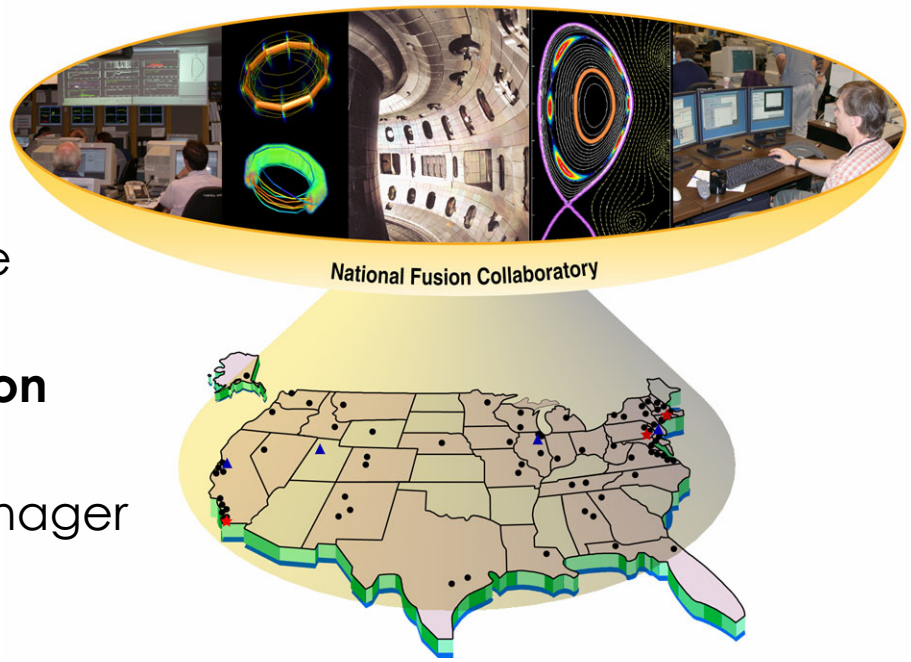
SHOT 116790 DATE:2004-01-29

Time	PHASE
16:49:44	IGS PHASE
16:49:11	ECH PHASE
16:49:09	PWR PHASE
16:48:53	PHDF PHASE
16:48:27	NB_MDS PHASE
16:48:23	BDEF PHASE
16:48:18	EFIT PHASE
16:48:17	NB PHASE
16:48:11	PHD PHASE
16:48:04	CO2 PHASE
16:48:01	MAG PHASE
16:47:35	NEW PHASE
16:44:33	EFITRT PHASE

FusionGRID
www.fusiongrid.org

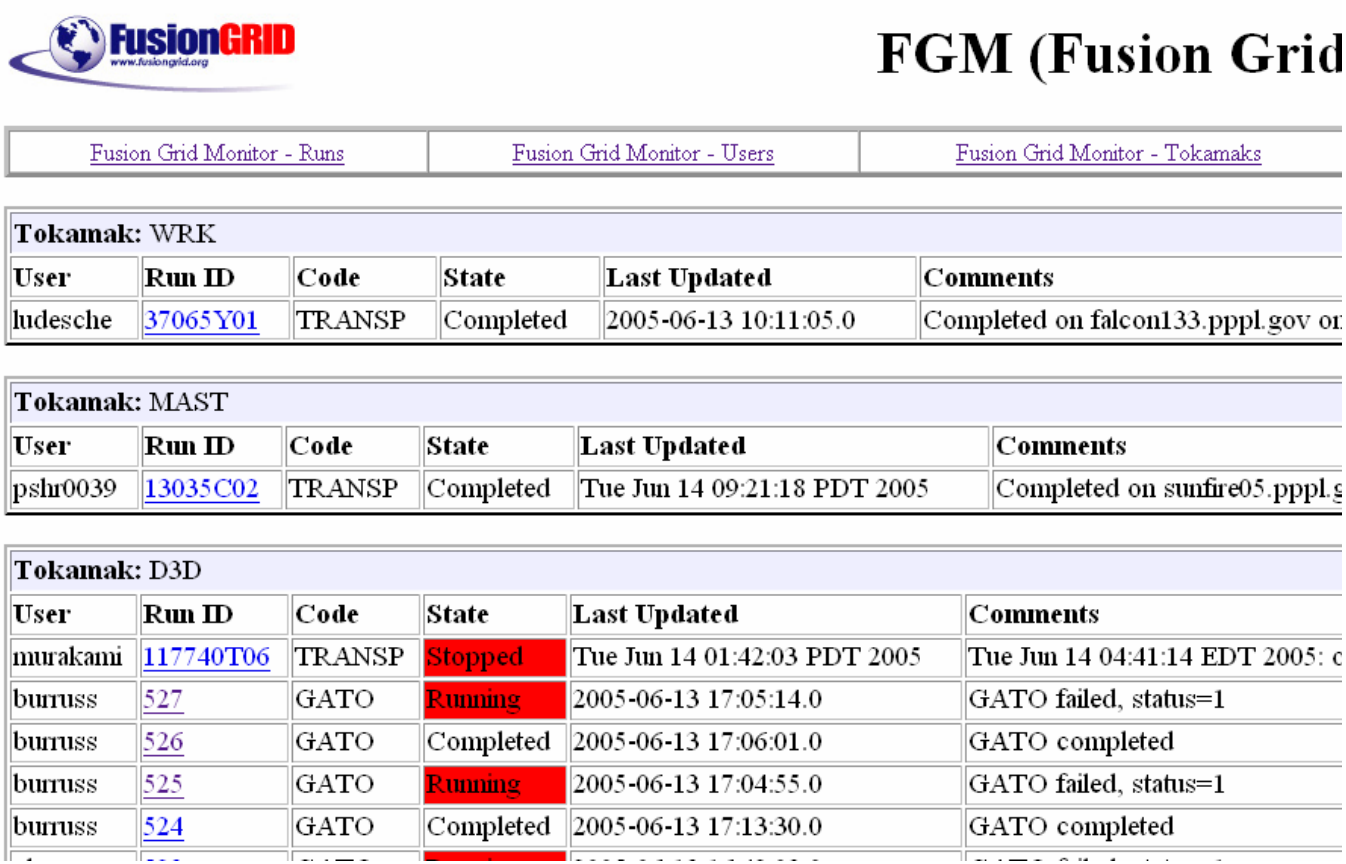
FUSIONGRID CREATED FOR BETTER USE OF RESOURCES

- **U. S. Fusion Grid (FusionGrid) aims to make more efficient use of computing resources**
  - Access is stressed rather than portability
  - Not CPU cycle scavenging or "distributed" supercomputing
- **Share resources between sites**
  - Reduce duplication of effort
  - Exploit comparative advantage
- **Develop a common tool set for fusion**
  - Globus Toolkit (GRAM & GSI)
  - FusionGrid CA & credential manager
  - ROAM
  - Secure MDSplus
  - Access Grid and VRVS



FGM USED FOR GRID-WIDE MONITORING

- For general grid-wide monitoring, the FusionGrid Monitor is used
- Like DAM, but more generalized



FusionGRID www.fusiongrid.org

FGM (Fusion Grid

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 on

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

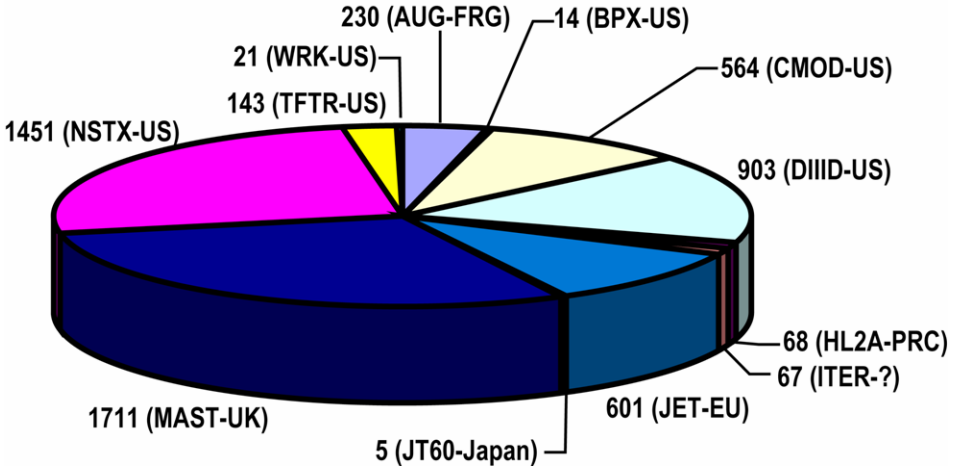
Tokamak: D3D

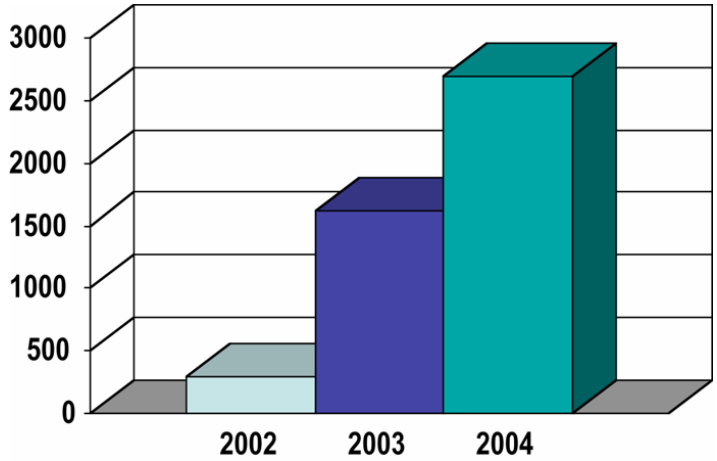
User	Run ID	Code	State	Last Updated	Comments
murakami	117740T06	TRANSP	Stopped	Tue Jun 14 01:42:03 PDT 2005	Tue Jun 14 04:41:14 EDT 2005: c
burruss	527	GATO	Running	2005-06-13 17:05:14.0	GATO failed, status=1
burruss	526	GATO	Completed	2005-06-13 17:06:01.0	GATO completed
burruss	525	GATO	Running	2005-06-13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

FusionGRID www.fusiongrid.org

CURRENTLY ADDING FUSIONGRID SERVICES

- **First FusionGrid service: TRANSP transport code**
- **The TRANSP code runs at PPPL and used throughout fusion**
  - Improved support
  - No code distribution issues
  - Result is users are more productive
- **But…what happens when you move to between-shot analysis?**

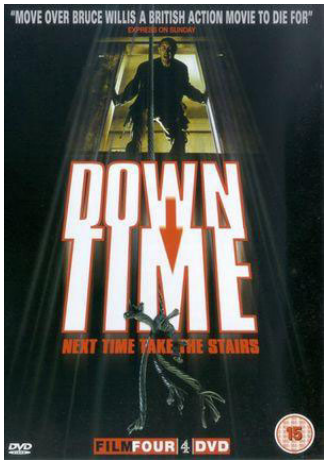




FusionGrid TRANSP runs

FAILURES ARE SENT TO DAM AND FGM

- **Analyses post errors to DAM and FGM**
  - User can click link to view details in the log file
- **Can post "service is down" messages**
- **Authorization failures also sent to ROAM**
- **Errors may require closer debugging work to figure out underlying problem**
  - e.g. "failed to connect to server"
- **The ability to view logfiles through DAM/FGM has proved very useful for debugging**

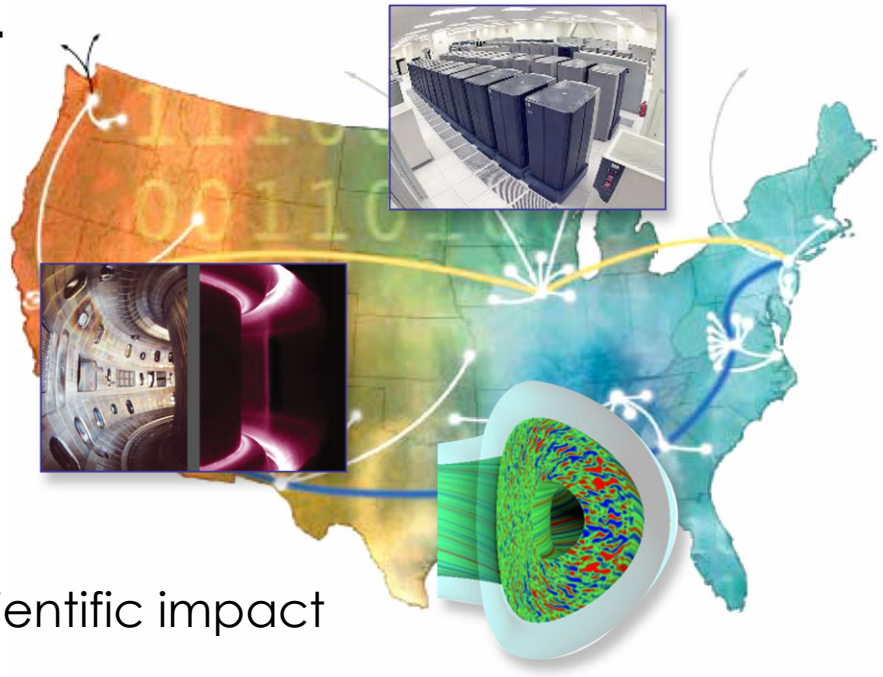


FAILURE RECOVERY IS NOT GENERALIZED

- **Failure recovery is specific to each component of FusionGrid**
- **Individual analyses may have their own mechanisms specific to that code**
  - e.g. try to connect three times before quitting
- **Secondary MyProxy and authorization servers are used**
  - Works like primary/secondary DNS
  - If clients fail to read from primary, they try secondary

MIGRATE BETWEEN-SHOT COMPUTING TO THE WAN

- **Deploying a supercomputer to support pseudo real-time analysis**
  - Network QoS
  - CPU scheduling
  - Faster CPUs and algorithms
  - Data management
  - End-to-end performance



- **Substantially enhanced data analysis**
  - Historically this had made a huge scientific impact

- **Can have a safety impact for future devices**
  - e.g. ITER: <10% of high power discharges can disrupt

FIRST PROOF-OF-CONCEPT BETWEEN-SHOT GRID ANALYSIS REVEALED PERFORMANCE ISSUES

- **A between-shot TRANSP proof-of-concept was done in March**
- **TRANSP (typically 12-36 hour code runs) adapted for shot cycle of 15 to 20 minutes**
- **Revealed that data transfer took more than 45 minutes!**
  - Negligible for 24 hour runs, but not for between-shot analysis
- **First round of optimization: 12x**
  - 48 minutes -> 4 minutes
- **Second round of optimization: 12x**
  - 4 minutes -> 20 seconds

DATA ANALYSIS THROUGHPUT IS THE MOST IMPORTANT PERFORMANCE CONSIDERATION

- **Anything that can increase data analysis throughput is useful**
  - More volume, less time
  - Anything else that helps scientists work faster
    - e.g. easy-to-use, well documented codes

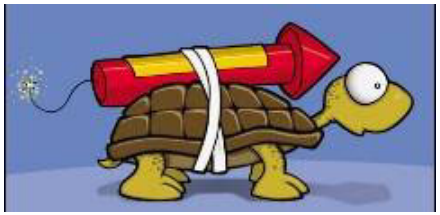
- **While it would be nice to have a supercomputing center at each tokamak site, that is too expensive**
- **We seek to increase data analysis throughput through resource sharing and more efficient use of resources**

RELIABILITY IS ANOTHER IMPORTANT CONSIDERATION

- **To be useful during an experiment, computational infrastructure must work**
- **A lost shot is a lost opportunity**
  - Example: the next-generation ITER device will be ~ $1 M per shot
  - Miss a shot, miss a million-dollar opportunity
  - DIII-D is an order of magnitude less expensive, but still not cheap
- **Must be reliable, dependable**
- **Would you take the bus to work if it showed up only 95% of the time?**

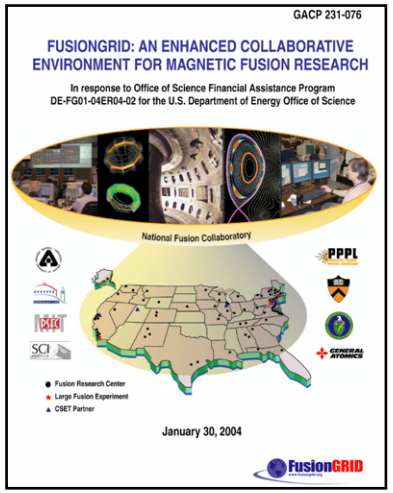
THE PERFECT PERFORMANCE TOOL?

- **The perfect performance monitoring & management tool would…**
- **Monitor Network QoS, CPU, mass storage, status of analyses, physics data consistency**
  - Automatic detection
- **Easy enough for scientists to use**
  - Tells them if something is wrong and who to contact for help
- **Easy enough for developers to use**
  - Open source, easily extended & have a simple API



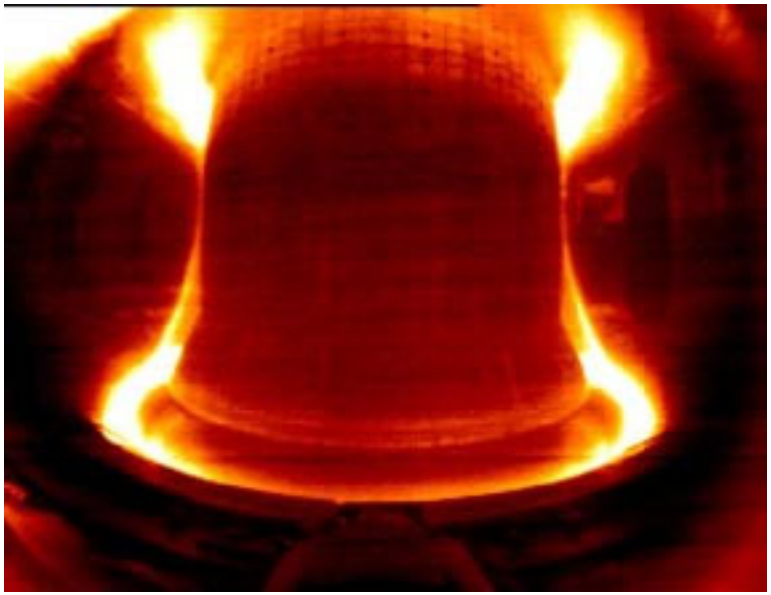
WE SEEK PARTNERS FOR DEVELOPING AND IMPLEMENTING THESE IDEAS

- Develop application requirements
- Determine relevant computer science components
- Examples of good application and computer science partnerships are the National Fusion Collaboratory and Network QoS Projects
- We would be eager to participate in exploring potential ideas



CONCLUSION

- **Fusion science is advanced through experiments**
- **Requires rapid between-shot analysis**
- **To increase data analysis throughput, moving from local-only data analysis to the WAN**
  - Performance issues
- **Important performance metrics are data analysis throughput and reliability**
  - Increase volume of analysis between shots
  - Help scientists to work faster
  - Must be a reliable infrastructure
- **We look forward to teaming together**



Inside a tokamak during a shot

QUESTIONS?

BACKUP SLIDES

- **Backup slides follow…**

CUSTOM TOOLS INCLUDE MDSPLUS, DAM, FGM

Performance Monitoring

Interpretation of Recent Data (hydra)

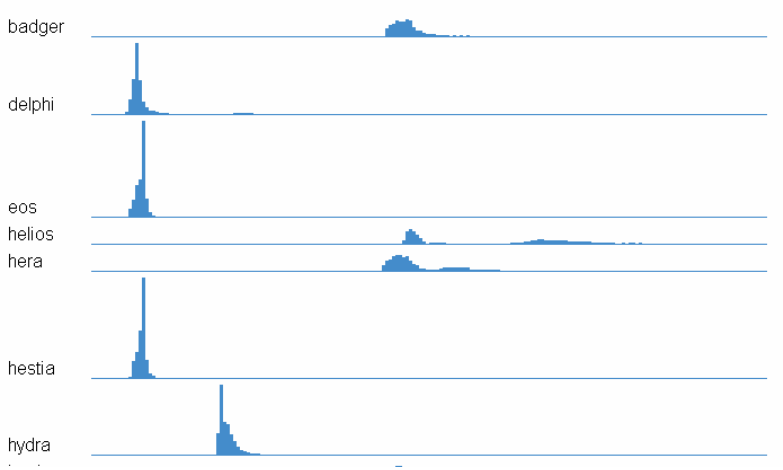
Home
FAQ
Links

Gathering statistics for hydra...

Most recent data indicate:

CPU utilization is normal
MDSplus data performance is typical
System load is low

MDSplus used by worldwide fusion community



FusionGRID www.fusiongrid.org

FGM

Fusion Grid Monitor - Runs | Fusion Grid Monitor - Users | Fusion Grid Monitor - T

Tokamak: TFTR

User	Run ID	Code	State	Last Updated	Comments
budny	88299A23	TRANSP	Completed	2005-06-09 14:56:02.0	Completed on falcon029.p

Tokamak: MAST

User	Run ID	Code	State	Last Updated	Comments
cbrick	13035C01	TRANSP	Completed	2005-06-09 11:03:22.0	Completed on sunfire08.p

FGM adapted from DAM