SOME THOUGHTS ON PRESENT AND FUTURE DATA MANAGEMENT ISSUES FOR THE MAGNETIC FUSION COMMUNITY

Presented by

D.P. SCHISSEL

*for the Magnetic Fusion Team**

Presented to

Workshop on Scientific Data Management

OCTOBER 20–22, 1998
Stanford, CA

GENERAL ATOMICS

MAGNETIC FUSION TEAM

Tom Casper – Lawrence Livermore National laboratory

Tom Fredian – MIT Plasma Science and Fusion Center, Alcator C–Mod

David Gates, Doug McCune, Steve Davis – Princeton Plasma Physics Laboratory, TFTR/NSTX

David Schissel – General Atomics, DIII–D National Fusion Facility

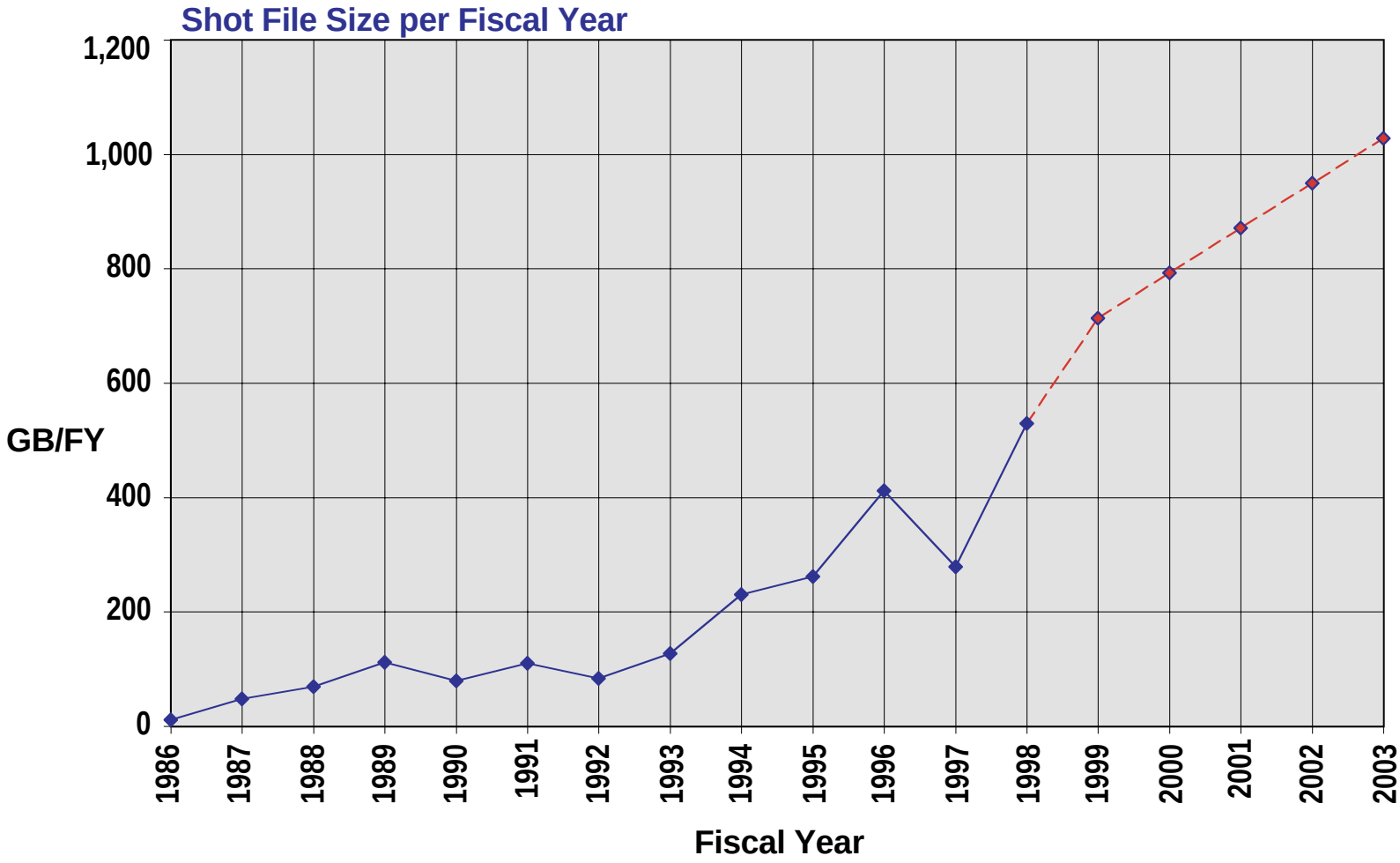
GENERAL ATOMICS

SUMMARY

- *Present experimental and modeling data storage requirements are modest*
  - *on the order of 10 TB for experiments*
  - *on the order of 1 TB for simulations*
- *Real time computing during experiments has a large benefit*
  - *Experiments are pulsed events typically lasting on the order 10 seconds*
  - *Between 30 and 40 pulses per day with 10 minutes between pulses*
  - *Pulse length will increase with newer machines*
  - *This is primarily a CPU issue*
- *Geographically diverse group of research scientists*
  - *Remote participation in experimental and theoretical research is required*
- *Future data storage requirements can grow substantially*
  - *Tools are needed to deal with larger datasets*
  - *More data is acquired per pulse every year with faster sample rates*
  - *More detailed transport simulations will require more CPU power and data storage*
  - *Simulations that move from the transport time scale (100 ms) to the MHD time scale (< 1 ms) will have large storage requirements*

RAW SHOT DATA FILE SIZE CONTINUES TO INCREASE

- More diagnostics, faster digitization rates, potentially longer pulse lengths
- Assume 15 weeks in FY98 and 18 weeks thereafter (today 200 MB/pulse)

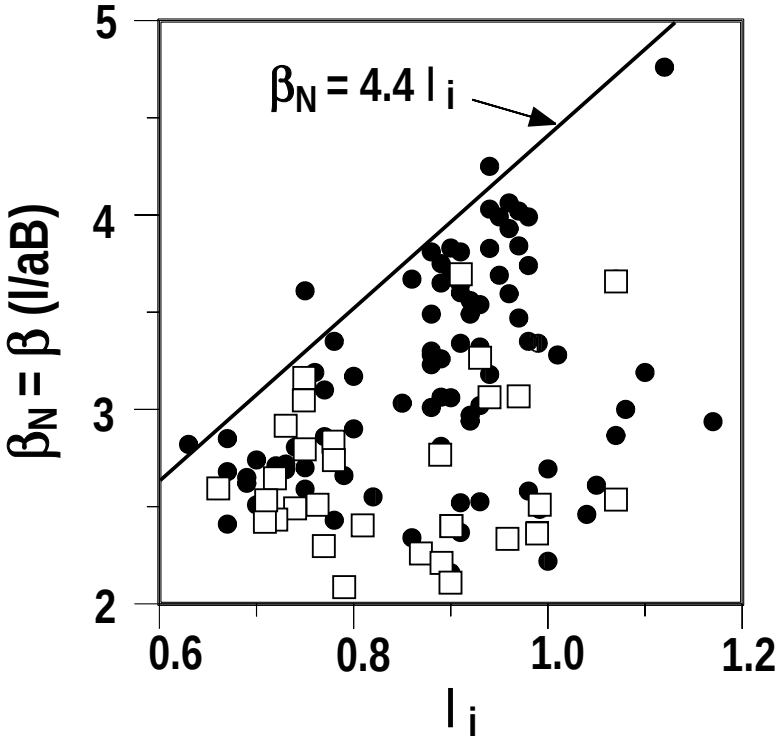


DIII-D NATIONAL FUSION FACILITY SAN DIEGO

GENERAL ATOMICS

DISRUPTIONS ARE AVOIDED BY KNOWLEDGE OF THE MHD STABILITY BOUNDARY

- MHD time scales are several orders of magnitude smaller than for energy transport
- Detailed experimental data and theoretical modeling will have large storage needs



DIII-D NATIONAL FUSION FACILITY SAN DIEGO

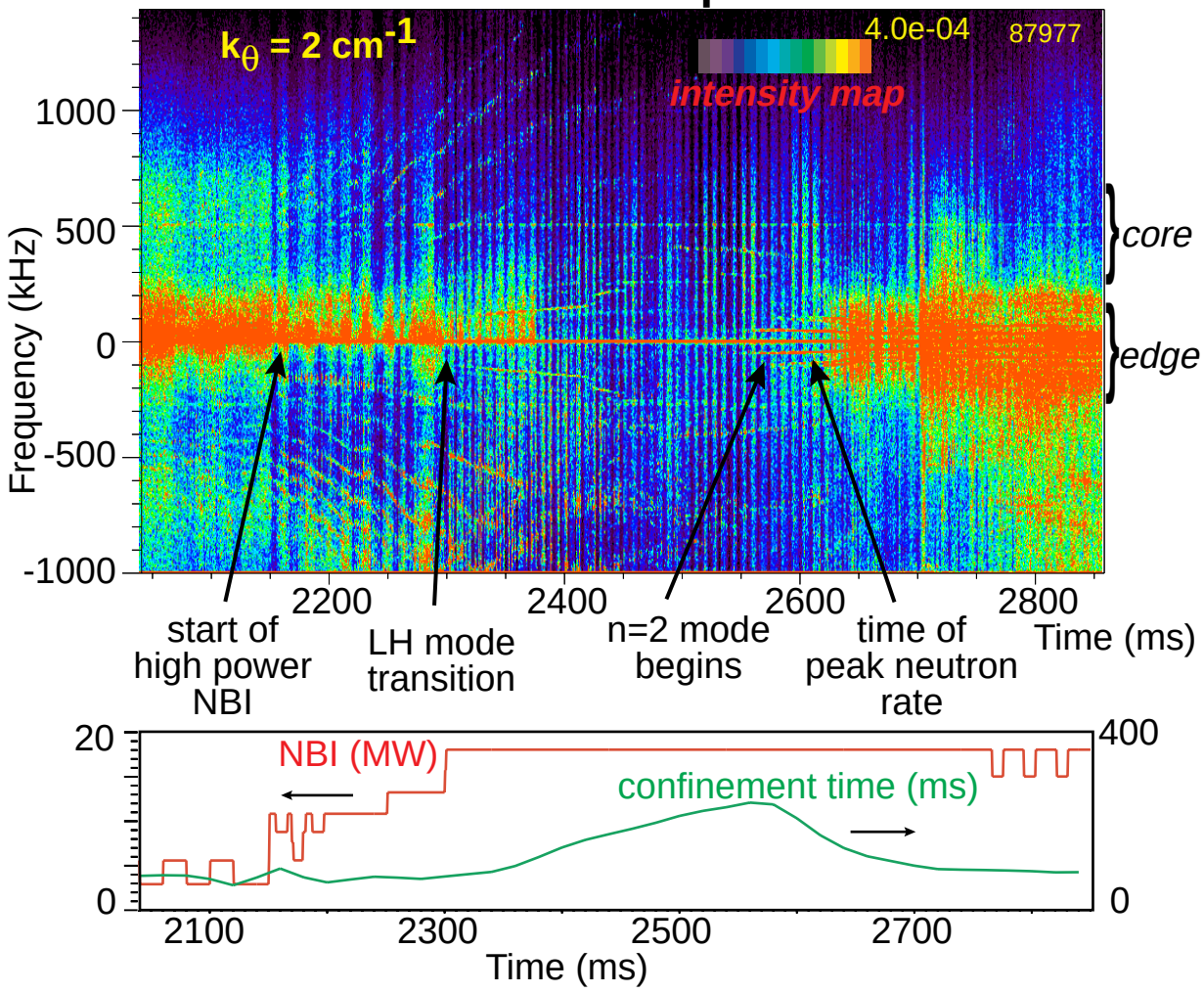
GENERAL ATOMICS

Highest Performance Achieved with Core Transport Barrier when Core Fluctuations Reduced

- **Core fluctuations observed at start of heating phase are completely suppressed prior to peak confinement time and reactivity.**

- **Termination of high performance phase accompanied by larger fluctuations.**

Color Enhanced Contour Plot of Measured Fluctuation Spectrum Evolution



UCLA Electrical Engineering

GENERAL ATOMICS

POTENTIAL LARGER DATA STORAGE

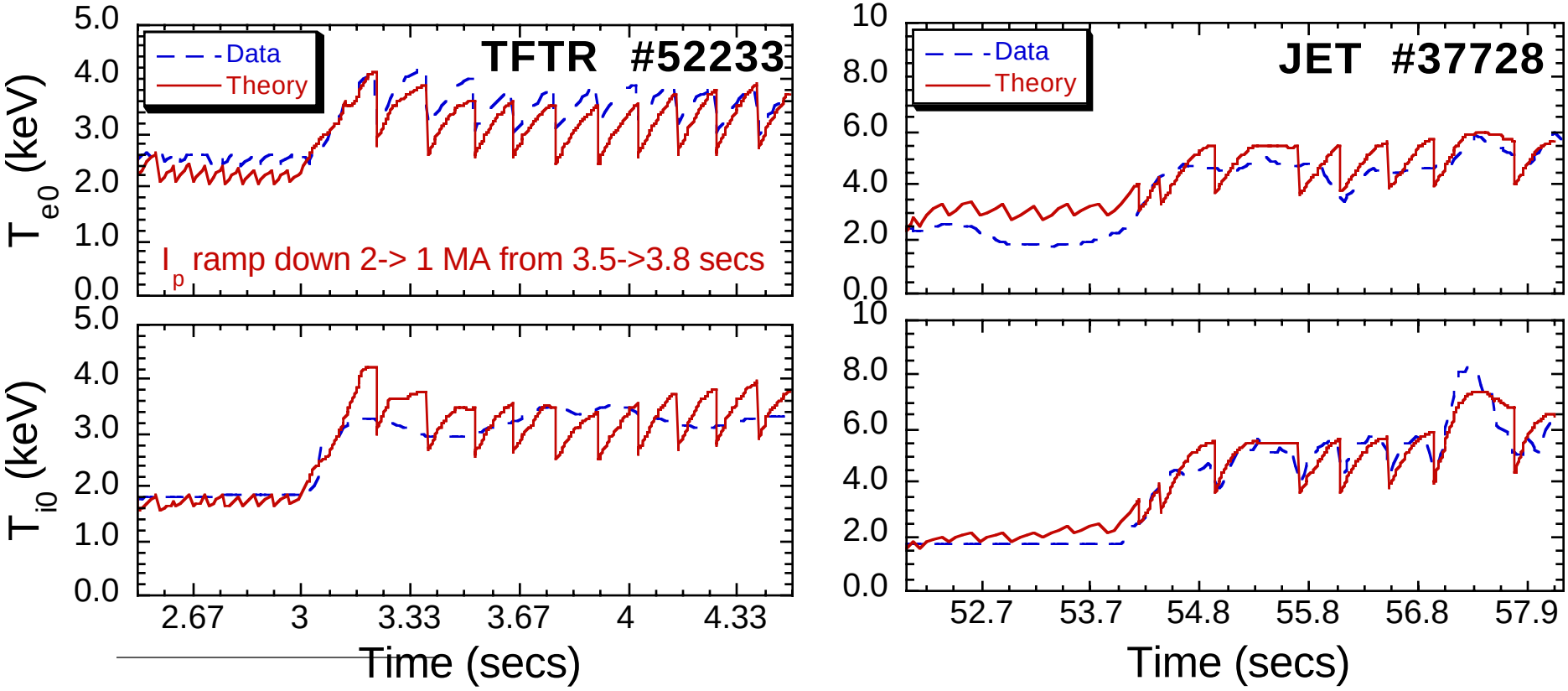
- *Far Infra–red Scattering Diagnostic measures plasma density fluctuations*
  - *Fluctuations correlated to increased energy transport*
  - *Allows close coupling of experiment and theory*
  - *Sampled to give as 1 MHz response for 0.5 seconds*
  - *Sample at 25 MHz for 5 seconds for a total of 500 MB per pulse which will double the size of the DIII–D pulse file*
- *Such increases desirable with other diagnostics (e.g. Reflectometer)*
  - *Electron density every 100 micro–seconds for 5 seconds (2 GB/pulse)*
- *Longer pulse lengths*
  - *From 5 seconds to minutes and possibly hours*

REAL TIME PLASMA CONTROL

- *Better use of experimental time by making something happen*
- *Computationally intensive*
- *Potential near term examples*
  - *Real time magnetic equilibrium reconstruction*
  - *Disruption avoidance*
  - *Real time ray tracing for RF control*
  - *Real time neutral beam power deposition for control of heating profile*

Multi-mode Model Successfully Predicts Full Time History For Many Discharges

- *Time evolution of profiles well described using MM model in the BALDUR transport code*
  - *Predicted temperatures for TFTR Ip ramp-down discharge #52233 and JET low nustar discharge #37728 shown below[1]*



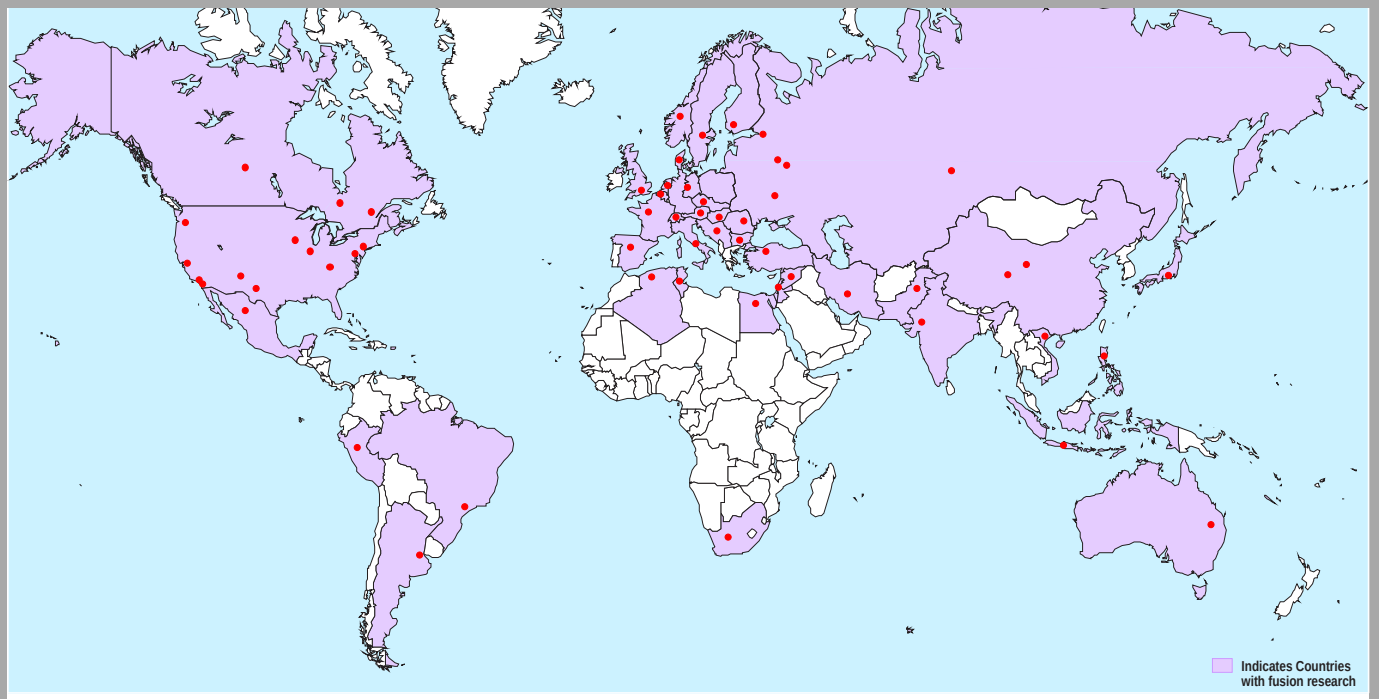
J. Kinsey

GENERAL ATOMICS

GEOGRAPHICALLY DIVERSE GROUP OF SCIENTISTS

- *International research effort*
- *Lack of common data formats for easy data exchange*
  - *Experimental raw data files*
  - *Simulation results*
  - *MDSplus data system being adopted by more facilities worldwide*
- *Database technology remains with the relational paradigm*
  - *How would object relational queries aid research for large 2D and 3D datasets?*
- *What network speeds will be required for remote analysis?*

Fusion Researchers Come to San Diego From All Over The World



Algeria
Argentina
Australia
Austria
Bangladesh
Belgium
Brazil

Bulgaria
Canada
China
Czechoslovakia
Denmark
England
Egypt

Finland
France
Germany
Greece
Hungary
India
Indonesia

Iran
Israel
Italy
Japan
Mexico
Netherlands
North Korea

Norway
Pakistan
Phillipines
Romania
Russia
South Africa

South Korea
Spain
Sweden
Switzerland
Taiwan
Thailand

Ukraine
United Arab Emirates
United Kingdom
USA
Vietnam
Yugoslavia

CONCLUSIONS

- *Ability to manage and visualize larger datasets can be beneficial*
  - *More detailed experimental measurements*
  - *More detailed plasma simulations*
- *Greater CPU power will be required to do real time analysis*
  - *Real time meaning during and between pulses*
  - *Will lead to better utilization of experimental time*
- *An ability to easily share data among a geographically diverse group of research scientists will benefit worldwide fusion research*
  - *Database technology will become critical*