

REMOTE COLLABORATION AND DATA ACCESS AT THE DIII-D NATIONAL FUSION FACILITY

Presented by
D.P. SCHISSEL

*for the DIII-D National Team**

Presented to
Workshop on Remote Participation in Fusion Experiments

June 28–29, 1998
Prague, The Czech Republic

ACKNOWLEDGMENT

- **The Data Analysis Group**
 - Qian Peng, Jeff Schachter, and 30% of Cathy Campo from GA
 - Bill Davis and Ted Terpstra from PPPL
- **DIII-D Computer Staff**
 - Kristi Keith, Bill McHarg, and Carl Parker
- **The DIII-D User Community**
- **From LLNL – Tom Casper, Bill Meyer, and Jeff Moller**
- **From MIT/C-Mod – Martin Greenwald, Tom Fredian, and Josh Stillerman**

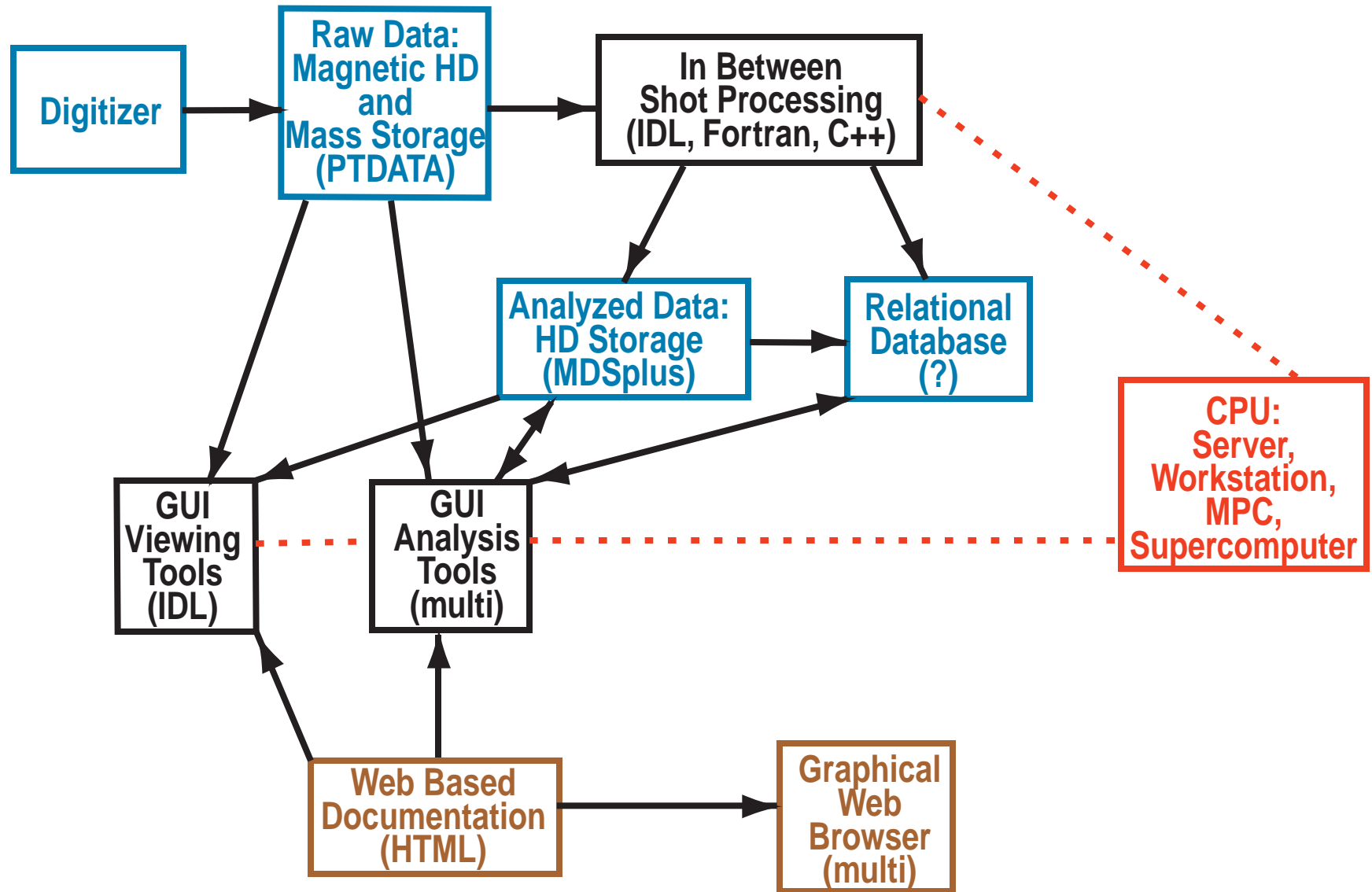


LONG TERM PLAN TO AID REMOTE COLLABORATORS

- DIII-D National Program is a multi-institutional collaboration
- Rapid 24 hour data retrieval for on-site and remote collaborators
- New data viewing & analysis tools aid the remote collaborator
- Web based documentation for computer codes and archived data
- Computer power for data analysis supplied to all collaborators

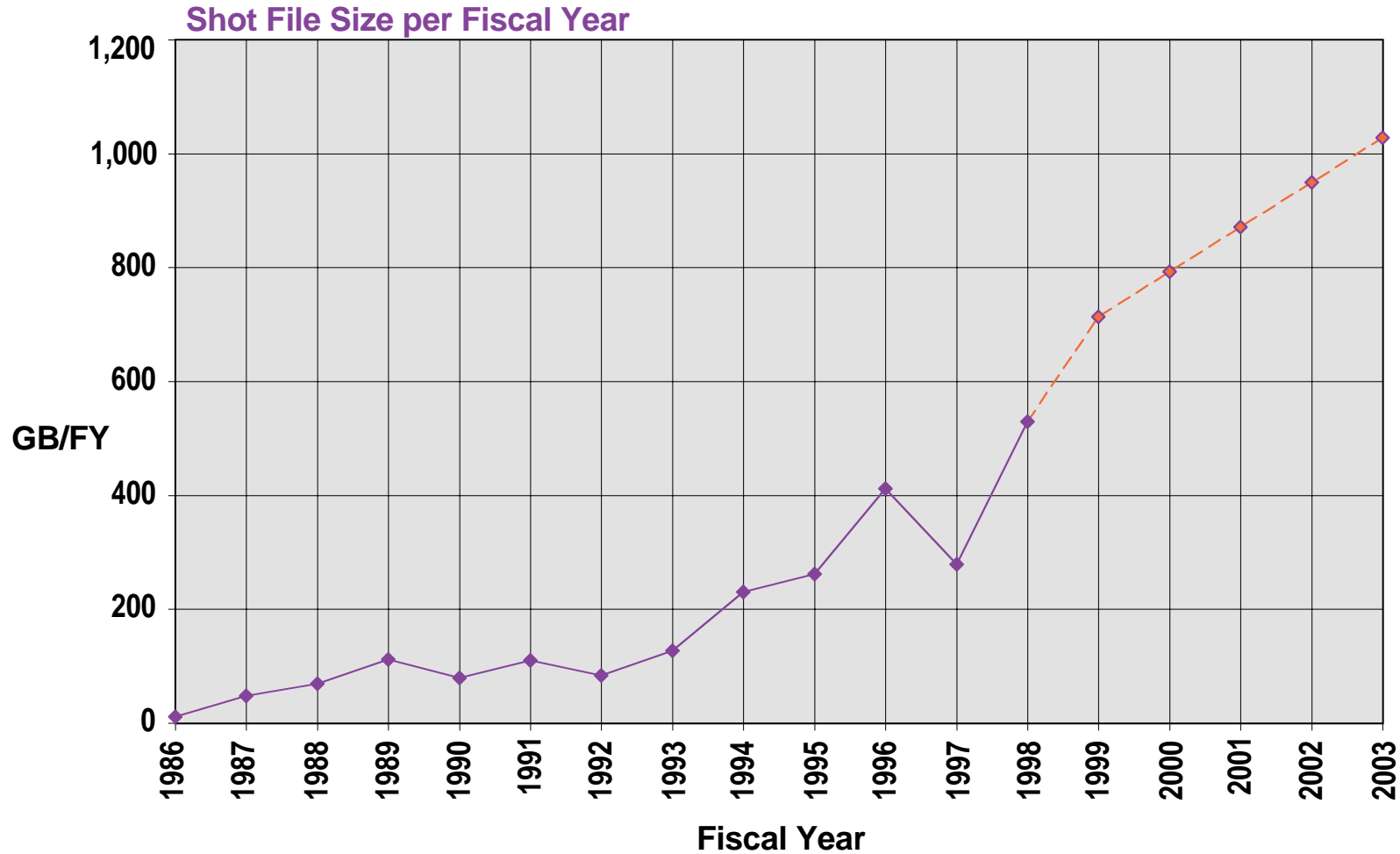


DIII-D DATA FLOW CHART



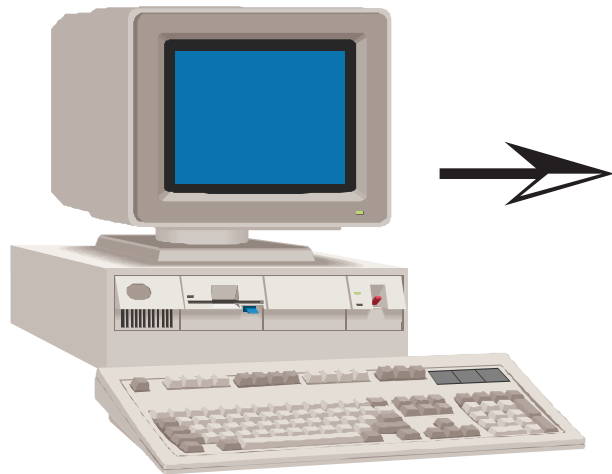
RAW SHOT DATA FILE SIZE CONTINUES TO INCREASE

- More diagnostics, faster digitization rates, potentially longer pulse lengths
- Assume 15 weeks in FY98 and 18 weeks thereafter



ALL DIII-D DATA AVAILABLE AROUND THE CLOCK

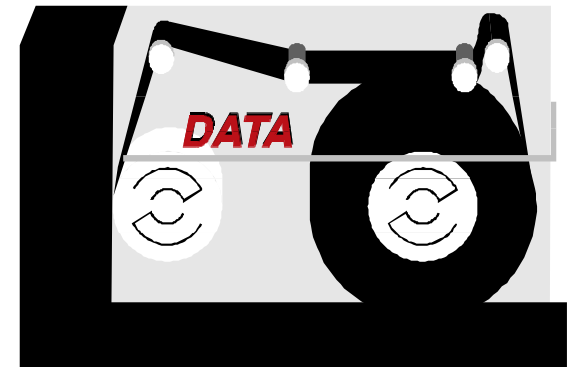
- All raw shot data automatically available (7x24) for data mining
 - 720 GB interactively (15–30 sec minimum), ≈9000 raw compressed shots
 - 2.4 TB on a 2–5 minute minimum time scale per request
 - Present needs are 1 TB with about 0.2 TB added per year
- All analyzed shot data on magnetic hard drive



Sun Ultra-1 Clone
100 GB Magnetic-Raid 5



HP Magneto-optical 600fx
620 GB from 238 platters
2.6 GB/platter, 6 drives



ATL DLT 7000 drives
2.4 TB from 68 bays
35 GB/tape, 4 drives

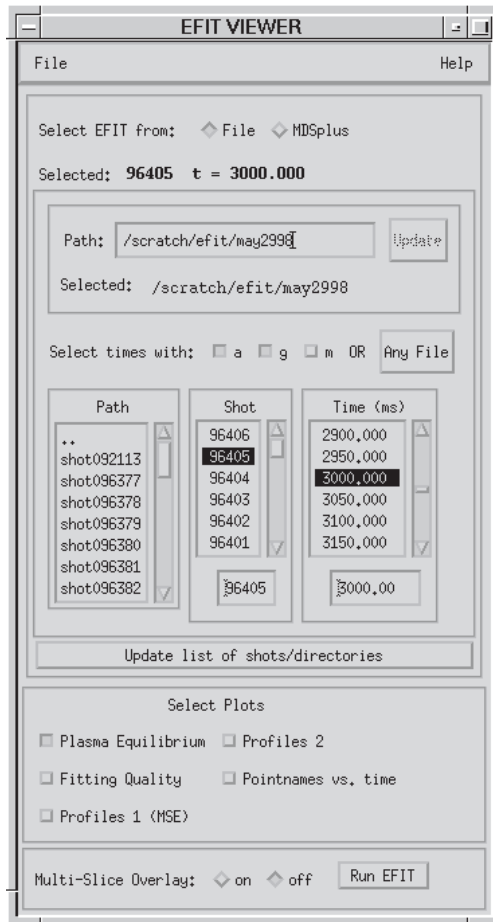
A NEW UNIX BASED RELATIONAL DATABASE WILL AID ON-SITE AND REMOTE COLLABORATORS MINE DIII-D DATA

- **Present relational database is old, has limited functionality, and does not operate in the Unix environment**
 - Scalar time-slice orientation moving towards profiles versus time
- **Database functionality requires easy interface**
 - Web access, relational queries from Fortran or IDL
 - Multi-platform clients
 - Allows mining of our 1 TB existing data and our anticipated 216 GB/year
- **Desired user functionality**
 - “Broad & shallow” containing basic data from essentially all DIII-D pulses
 - “Highly Analyzed” containing publication quality data
 - “Special Purpose” containing user specified information



EASY TO USE VIEWING & ANALYSIS TOOLS WILL HELP THE REMOTE COLLABORATOR TRANSITION INTO THE DIII-D ENVIRONMENT

EFIT Selection GUI



- Efficient GUI design with object oriented programming
 - Tools written in IDL, a commercial data package
- Created an object oriented IDL based graphics library
- Viewing tools created to examine EFITS, time histories, and plasma profiles
- Analysis tools perform interactive EFIT, kinetic EFIT, profile fitting, and a ONETWO GUI interface

VIEWING AND ANALYSIS TOOLS ARE BEING DISTRIBUTED TO REMOTE COLLABORATORS TO AID THEIR WORK

- **Viewing tools**
 - Source code management system (CVS)
- **Analysis tools**
 - EFIT: CVS or one tar file with sources
 - ONETWO: one tar file with sources
- **Benefits of IDL**
 - Rich language
 - Code can move among operating systems with little modification
- **Drawbacks of IDL**
 - Remote collaborators must purchase a license
- **Public domain software, Java and Python, are being investigated**
 - Presently they do not have all the functionality of IDL



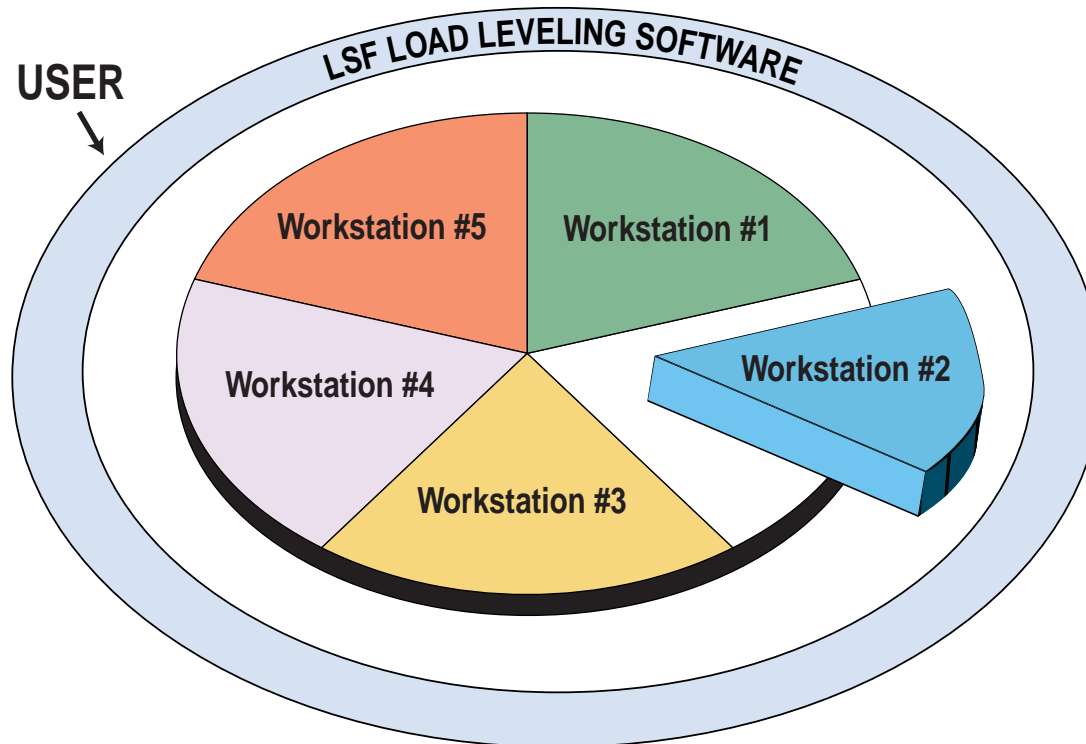
WEB BASED DOCUMENTATION BRINGS CRITICAL INFORMATION TO THE ON-SITE AN REMOTE COLLABORATOR

- **Computer code documentation**
 - By the scientific research staff
 - By the computer scientist staff
- **Data documentation**
 - Includes both raw digitizer data and analyzed data repositories
- **One source, multiple paths to reach the source**
 - Direct from Web browser
 - From within a tool



SUBSTANTIAL COMPUTATIONAL POWER IS PROVIDED TO ON-SITE AND REMOTE COLLABORATORS

- Unix environment with a large server and numerous workstations
- Control room EFITs operated in a distributed environment
- Load sharing software being tested for more efficient CPU usage



SUMMARY

- **Long term plan will aid the remote collaborator**
 - Easy and fast access to both raw and analyzed data
 - Tools with a GUI and ample documentation
- **Future work to investigate a massively parallel computing environment**
 - Run integrated modeling and simulation codes
 - Store results in MDSplus for easy comparison to experiment
 - New visualization tools to handle the increased data



DIII-D PROGRAM COLLABORATORS

National Laboratories	Universities	International Laboratories
ANL	Cal Tech	Academia Sinica (China)
INEL	Columbia U.	Cadarache (France)
LANL	Hampton U.	CCFM (Canada)
LLNL	Johns Hopkins U.	Culham (England)
ORNL	Lehigh	FOM (Netherlands)
PNL	MIT	Frascati (Italy)
PPPL	Moscow State U.	Ioffe (Russia)
SNLA	RPI	IPP (Germany)
SNLL	U. Maryland	JAERI (Japan)
	U. Texas	JET (EC)
	U. Washington	KAIST (Korea)
	U. Wisconsin	Keldysh Inst. (Russia)
<u>Industry Collabs</u>	UCB	KFA (Germany)
CompX	UCI	Kurchatov (Russia)
CPI (Varian)	UCLA	Lausanne (Switzerland)
GA	UCSD	NIFS (Japan)
Gycom		Troitsk (Russia)
Orincon		Southwestern Inst. (China)
		Tsukuba U. (Japan)

