Handling and Archiving of Magnetic Fusion Data at DIII–D[*]

J.F. VanderLaan, S. Miller, B.B. McHarg, Jr., P.A. Henline

General Atomics

P.O. Box 85608, San Diego, California 92186-9784

ABSTRACT

Recent modifications to the computer network at DIII–D enhance the collection and distribution of newly acquired and archived experimental data. Linked clients and servers route new data from diagnostic computers to centralized mass storage and distribute data on demand to local and remote workstations and computers. Capacity for data handling exceeds the upper limit of DIII–D Tokamak data production of about 4 GBytes per day. Network users have fast access to new data stored on line. An interactive program handles requests for restoration of data archived off line. Disk management procedures retain selected data on line in preference to other data. Redundancy of all components on the archiving path from the network to magnetic media has prevented loss of data. Older data are rearchived as dictated by limited media life.

INTRODUCTION

Research into the properties of high-temperature plasmas is being conducted in the DIII–D tokamak at General Atomics under contract with the U.S. Department of Energy. A prime goal of these efforts is to provide data to guide the design of future fusion devices. Upgrades to DIII–D data acquisition computer systems described in [1–3] have provided the capacity to handle increased data flow as diagnostic requirements have continued to grow. This report surveys the current status of hardware and software which move data from acquisition nodes to central disks and to magnetic tape, and then back on line when requested for analysis.

Plasma discharges, or "shots," occur at approximately ten-minute intervals during DIII–D operation. Over 35,000 shots have been run since the DIII–D upgrade was completed in 1986, and the uncompressed size of all data generated per shot has climbed from five to over 120 megabytes during the ensuing decade. The total accumulation of DIII–D shot data, archived to magnetic tape, exceeds by an order of magnitude the online disk capacity of the computer system. Expansion of local and wide area networks has made possible quicker access to data by a growing number of on-site and remote users.

Fig. 1 shows a generic block diagram of the evolving computer system at DIII–D. Three categories of data flow through this system: (1) new experimental data being assembled into shot files and archived, (2) data from shot files on disk provided to users on demand, and (3) data restored from tape at the User Service Center (USC) and transferred to disk for user access. Shot file data on disk at either the DIII–D or USC site is accessible to users at all nodes across the network.

The primary function of the computers at the DIII–D site is to provide immediate control, data acquisition and data analysis during plasma operations. Data analysis performed at DIII–D provides immediate feedback to experimenters for use in controlling plasma parameters during the operations day. Computers at rhe USC site are available for user access to old and new shot data for analysis requiring intensive computation. The distinction between the two sites is lessening as more nodes are added at the DIII–D site.

By comparison with the computer configuration of two years ago, the following changes have taken place:

1) The AEG Unix computer now acquires a large majority of all CAMAC data (50 to 60 megabytes, compressed to less than 20) which was previously transferred over a Hyperchannel link from the MODCOMP to VAXS.

2) All shot data is compressed at the node which acquires it; 40–50 megabytes of compressed data (reduced from 110–120 megabytes before compression) per shot is transferred over ethernet. The advantage of local compression is that shot data traffic flow on the ethernet has risen by only 60 to 70 percent in two years despite the removal of the Hyperchannel link and the substantial increase in overall raw data acquisition over that time.

3) Additional Unix nodes and X terminals have been installed throughout the local area network (LAN). These machines have added significantly to the traffic on all segments of the LAN.

4) The MODCOMP computer now generates only a few kilobytes of information on shot timing and neutral beams; this data is passed to the AEG computer via the Real/Max shared memory. This is a temporary configuration which should be eliminated in the next 6 to 9 months.

MANAGEMENT OF NEW SHOT DATA

The flow of new data from distributed acquisition nodes to the central VAX disks is regulated by a set of interacting codes collectively called the Interprocess Communication System (IPCS). Various VAXes, MicroVAXes and Unix nodes are located on the ethernet segments at DIII–D. They do not all share the same networking protocol. Digital
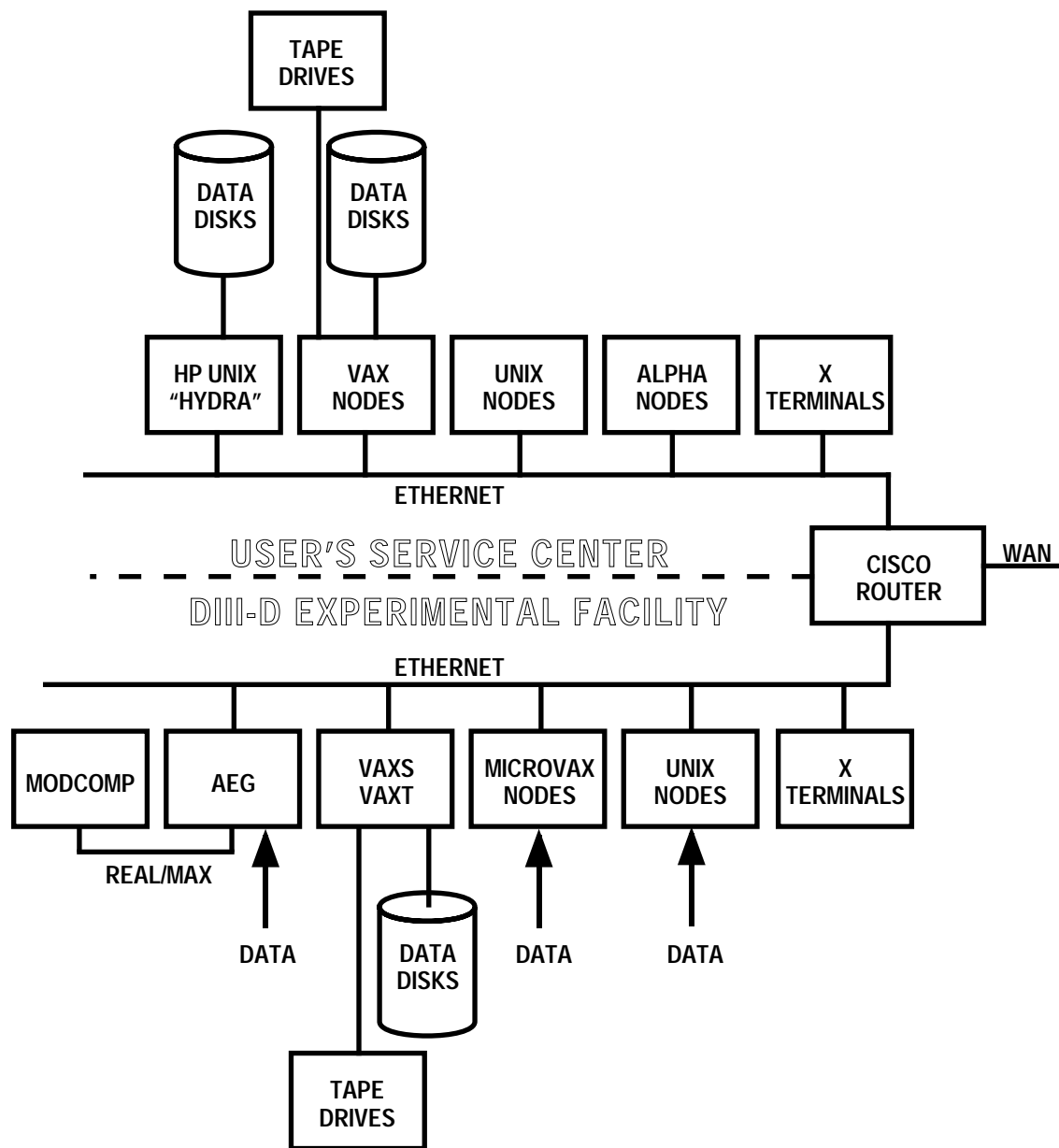
Fig.1. DIII–D data acquisition computer system (mid-1995).

Equipment Corporation's VAXcluster software links VAXS, VAXT and most MicroVAXes; most Unix nodes run TCP/IP, although others run DECnet, both of which are also available on VAXS and VAXT. IPCS provides a method for processes on the network to easily communicate with one another despite the variety of protocols. Messages are sent to a process by name, and IPCS finds the process and delivers the message. IPCS runs on both VAX/VMS and Unix systems. Each computer running IPCS has a library of IPCS routines and two processes: grandpa and netmailer.

The IPCS grandpa process creates various resources used by IPCS and then may start other processes. It then remains in memory and may be used to start and stop certain processes on the system.

The IPCS netmailer process (IPCSMAILER) is the only IPCS process that deals with network protocols. When the receiver of an IPCS message cannot be found on the local computer, the message is passed to the IPCSMAILER process. The IPCSMAILER maintains a cache of processes and their location in the network and forwards messages to the IPCSMAILER on the appropriate computer. If a process name is not in the cache, the IPCSMAILER coordinates a "name search" operation, asking the other IPCSMAILERs if they know about the receiving process.

The shot index server (SHOTIDXSVR) runs on VAXS and VAXT and reads and writes the shot index file. Two different versions of the shot index file exist. The DIII–D file contains a record for each shot file resident on DIII–D (but not User

Service Center) data disks; the record is removed when the shot file is deleted. The larger file at the User Service Center keeps a permanent record for each shot and is updated daily with new entries found in the DIII–D file. Keywords in shot index file records monitor the status of a shot file's acquisition, compression, movement to central disk storage and archiving on magnetic tape. SHOTIDXSVR security mechanisms ensure that only certain processes are allowed to write to the shot index file.

DATAMONITOR runs on both Unix and VAX nodes. It requests creation of shot index file entries for new data and compression of the files after acquisition is complete. It deletes an uncompressed shot data file from a local disk after compression is finished and later deletes the compressed file after it has been copied to the central disk and to one magnetic tape.

RPRESSIT runs on the same nodes as DATAMONITOR. It compresses shot files in response to messages sent by DATAMONITOR.

DATAMOVER runs on VAXS and VAXT. It monitors the shot index file, watching for files flagged by DATAMONITOR as compressed. It copies these files to the central data disk using ftp or VMS copy commands.

Other non-IPCS tasks also run on VAXS or VAXT and perform file management:

TSKMGR starts VAX and MicroVAX processes which control VAXCluster acquisition and analysis. FILMGR copies non-shot-data-file pointname files from VAX, MicroVAX and Unix nodes for compression and archiving. PRESSIT uses the same compression algorithm as RPRESSIT, but runs on VAXS and VAXT.

ARCHIVE scans the DIII–D shot index file for shot data files which are compressed and resident on the VAX data disks. Two copies of ARCHIVE run simultaneously; each shot file is written to two 8 mm data tape cartridges, one of which is stored at the User Service Center and the other at the DIII–D experiment site.

DATMGR runs on VAXS and maintains free space on the VAX disks. DATMGR deletes uncompressed shot data files after PRESSIT has compressed them. The handling of a compressed file depends on a judgement of the value of the shot by experimenters. A shot's files may be put on "hold" for temporary retention on disk, or may be marked "keep" if regarded as valuable for analysis. A permanent "Keepit" list of these Tokamak shots is maintained. Compressed files which have neither designation are eligible for deletion after both tape copies have been made, but actual deletion of the oldest eligible files takes place only when space is needed for new shots. In any case, the shot data can be retrieved if required. DIII–D shot index file entries are periodically scanned for transfer of new information to the USC shot index file. The COPYD3 code periodically moves "keep" shots to the USC VAX data disk via ethernet.

USER ACCESS TO SHOT DATA FILES

Provisions for user access to on-line shot files has been described in detail in [2]. Users at any node shown in Fig. 1 may access data stored on line at any node that generates data or on the VAX or Hydra disks; data is accessible over the wide-area network as well. As noted previously, the collection of computers on the ethernet at both GA sites use VAXCluster, DECnet or TCP/IP networking protocols, so that it is not guaranteed that a user's node can directly communicate with the node at which the requested data is stored. In addition, the user is expected to know only the shot number and pointname he wants; the data search logic must find the correct data file of the set which exists for the shot. The PTDATA and PTSERVER programs perform the search functions to fulfill the user's request for data.

The PTDATA code is a library routine available at all VAX/VMS and Unix nodes; PTSERVER runs at all nodes at which data may be stored on line. Tables on User Service Center computers supply shot file identification and location for requested pointnames. PTDATA logic attempts to locate data across the network at the lowest possible cost, weighing factors such as the possibility of finding data locally rather than over a remote link, the location at which a previous search was successfully completed, the speed of different network protocols, and the need to use a gateway node (VAXS or VAXT) to translate between unlike protocols.

RESTORATION OF ARCHIVED DATA

Fig. 2 illustrates the path by which shot data flows from the DIII–D data disks to those at the User Service Center. A total of 47 GBytes of centralized storage is distributed among the DIII–D VAXes, the USC VAXes and the USC HP T500 "Hydra". Keepit shot files are transmitted from DIII–D to the USC over fiber-optic ethernet; other shot files are archived on magnetic tape cartridges and may be restored to the USC VAX data disk if requested.

The selection of shots to be restored and retained on USC data disks is governed by user requests. A user can run the SHOTS code on any node to obtain a list of shot numbers which are currently on line on either USC or DIII–D data disks. If the desired data is not on disk, a request for restoration of data from tape can be queued by running D3SHOTS on a USC VAX.

D3SHOTS is one of several codes which access data sets created by S1032, a VAX/VMS database system supplied by PRAXIS International, Inc. S1032 permits interactive and programmed access to data sets and provides security against unauthorized access. Several data sets are used in management of USC data disks; these are (1) a copy of the USC shot index file, updated nightly; (2) a directory of tape numbers and shot sizes; (3) the queue of shots requested by D3SHOTS users; and (4) a compilation of statistics on shot restoration activity. A sorted version of the D3SHOTS output shows the computer operator which shot data files to copy from tape to the USC VAX data disk.
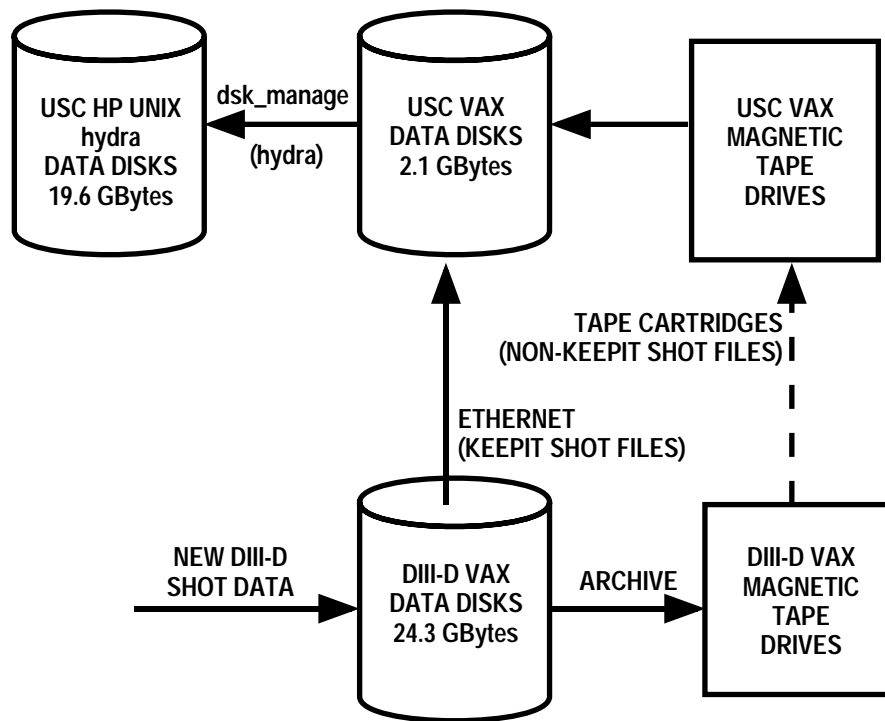
Fig. 2. Data flow from DIII–D site to USC site.

The hydra code "dsk_manage" controls movement of these files from the VAX data disk to the hydra data disks and the subsequent removal of files from the five hydra disks. At present "dsk_manage" runs three times daily to handle the volume of shots being restored from tape. The code moves all shot files from the VAX disk to the hydra drives. In order to determine which shots to delete from the hydra disks, "dsk_manage" reads a configuration list which normally guides it to maintain 10% free space on each disk, delete the least-recently read files, but never delete files resident less than 24 hours, and retain Keepit shots in preference to others. The maximum number of Keepit shots is 200.

Older shot files are periodically restored to disk and recopied to new tape in order to prevent data loss because of media degradation.

CONCLUSIONS

DIII–D shot data file handling has changed vastly over the past decade in order to accommodate advances in device and diagnostic capability and the resultant requirements for greater data collection rate and volume. Distributed shot file creation and compression, use of networking for all file transfers, and installation of faster Unix processors and larger data disks have been instrumental in ensuring that experimental data is reliably available, both immediately after creation and in the long term.

REFERENCES

[1] P.A. Henline, "Use of open systems for control, analysis and data acquisition of the DIII–D Tokamak," in Proc. 5th IEEE/NPSS Symp. on Fusion Engineering, pp. 127, 1993.
[2] B.B. McHarg, Jr., "Access to DIII–D data located in multiple files and multiple locations," in Proc. 5th IEEE/NPSS Symp. on Fusion Engineering, pp. 123, 1993.
[3] B.B. McHarg, Jr., "Conversion of the central DIII–D data acquisition System to a Unix based platform," this conference.
[4] G.G. Preckshot and D.N. Butner, "A simple, efficient interprocess communication system for dissimilar networked computers," Energy and Technology Review, Lawrence Livermore National Laboratory Report, UCRL-52000-89-10, October 1989.