THE USE OF A VAX CLUSTER FOR
THE DIII–D DATA ACQUISITION SYSTEM

B.B. McHarg, Jr.
General Atomics
P.O. Box 85608
San Diego, California 92186-9784

Abstract: The DIII–D tokamak is a large fusion energy research experiment funded by the Department of Energy. The experiment currently collects nearly 40 Mbytes of data from each shot of the experiment. In the past, most of this data was acquired through the MODCOMP Classic data acquisition computers and then transferred to a DEC VAX computer system for permanent archiving and storage. A much smaller amount of data was acquired from a few MicroVAX based data acquisition systems. In the last two years, MicroVAX based systems have become the standard means for adding new diagnostic data and account for half the total data. There are now 17 VAX systems of various types at the DIII–D facility. As more diagnostics and data are added, it takes increasing amounts of time to merge the data into the central shot file. The system management of so many systems has become increasingly time consuming as well. To improve the efficiency of the overall data acquisition system, a mixed interconnect VAX cluster has been formed consisting of 16 VAX computers. In the cluster, the software protocol for passing data around the cluster is much more efficient than using DECnet. The cluster has also greatly simplified the procedure of backing up disks. Another big improvement is the use of a VAX console system which ties all the console ports of the computers into one central computer system which then manages the entire cluster.

Introduction

The DIII–D tokamak is a large fusion energy research experiment operated by General Atomics (GA) and funded by the Department of Energy (DOE). The experiment is dedicated to the continuing investigation of high beta and confinement issues in high temperature plasmas approaching fusion reactor-like conditions.

A typical operating day of the experiment produces 40 to 50 tokamak shots with each shot lasting several seconds and repeated every 10 min. Data from a shot is acquired, archived, and analyzed by the computer systems. An average shot now consists of over 40 Mbytes of data and the largest shot yet generated contained 50.5 Mbytes of data. The largest potential shot, if all diagnostics were operating at the same time, would be over 60 Mbytes. These numbers are continuing to grow as new diagnostics are added to the experiment, faster time scales are used to diagnose plasma behavior, and plasma discharge times are lengthened.

Originally, most data was acquired by a data acquisition system consisting of MODCOMP Classic computers from Modular Computer Systems [1-5]. At the start of the DIII–D experiment, the computer systems were upgraded to include VAX computers from Digital Equipment Corporation (DEC). Two VAX computers were joined into a VAX cluster which permits a high degree of resource sharing between the systems [6-7]. Small MicroVAX II computers then began to be used to acquire data for certain special plasma diagnostics. These systems communicated only by network communications (DECnet) running over an Ethernet cable. The central VAX cluster systems acquired data from the MicroVAX systems using DECnet [8-10].

However, the amount of data acquired by these systems was small.

In the last five years, the number of auxiliary MicroVAX data acquisition systems has grown from 2 in 1986 to 15 today, and that number is expected to increase. The amount of data collected has grown from a few percent to now comprising about half the total data acquired. In order to centralize and simplify system management and to increase the overall efficiency of these systems, the overall efficiency of data collection, and the overall efficiency of computer personnel, the systems have now all been merged into one large VAX cluster. This cluster is called a mixed interconnect cluster because it has two types of cluster connection mechanisms. One is the 70 Mbit/s CI bus used by the central nodes of the cluster and the other is the 10 Mbit/s Ethernet used by MicroVAX systems.

This paper will describe the configuration of the DIII–D mixed interconnect VAX cluster and how it has improved the efficiency of data collection, system management, and personnel efficiency. The use of a VAX console system for central console management will also be described.

Computer System Overview

Figure 1 illustrates a global overview of the computer configuration supporting the DIII–D project. Much of the data is still collected via CAMAC data acquisition hardware connected to the MODCOMP data acquisition computers. This data is then transferred to the VD3D VAX cluster system via a Network Systems Corporation Hyperchannel 50 Mbit/s link. CAMAC based data is also collected by the MicroVAX systems which are part of the VD3D VAX cluster.

The User Service Center VAX cluster, labeled VUSC, is located about 1 mi away from the DIII–D experimental facility. The two VAX clusters communicate using DECnet, although the User Service Center is on a separate Ethernet. VUSC is
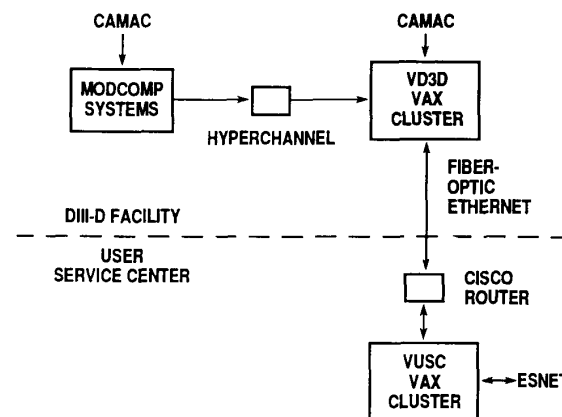

Fig. 1. Overview of DIII–D computer systems.

connected to the VD3D experimental facility by a fiber optic link and a cisco router from cisco Systems. The VUSC VAX cluster is heavily used for interactive examination of data during tokamak operations as well as nonoperational periods. It also supports all other needs of the fusion division at GA. The User Service Center is a primary node on the ESnet backbone which provides nationwide computer communication services for all Department of Energy facilities. The VUSC VAX cluster currently consists of one VAX 8650, one VAX 11/750, two VAX 4000-300, and two VAXstation 3100/76.

DIII-D VAX Computer Configuration

Figure 2 illustrates the connectivity of the VAX systems at DIII-D. These computer systems are partly located in the computer room at the DIII-D experimental facility and partly in the adjacent control room area. In addition, several of the systems are located within the facility but at locations that are remote from the computer/control room area.

Table 1 lists each node of the VAX cluster along with the type of computer and its purpose.

VAXS is the most important centralized node as it acquires all the data from both the MODCOMP systems and from the other VAX systems into a central shot data file. VAXS and VAXT also provide remote data file access to the User Service Center over DECnet. All of these systems are logically part of the same cluster. VAXS and VAXT cluster communication is over the CI bus while all other node's cluster communication is over the Ethernet to VAXS and VAXT. The HSC based disks are local to VAXS and VAXT but are available to the entire cluster via the cluster software. MVXS is not actually part of the cluster as it is the VAX console system which will be discussed later. MVXS can be booted into the cluster, if necessary, however. It is connected by DECnet running over the Ethernet. MVXW, MVXV, and MVXM are satellite nodes with no directly attached system disk. The system disk for the satellites is an HSC disk, the same disk as for VAXS and VAXT. All the other MicroVAX nodes do have a local system disk. Having a local system disk enables them to operate independently of any other nodes in the cluster should that be necessary. These other nodes can also be removed from the cluster and booted as stand-alone systems.

Disk Mounting in the Cluster

DEC's VAX cluster software permits disks on one node of the cluster to be mounted on another node. Thus, a disk will appear to be local to a computer even though it is physically connected to another computer. Figure 3 illustrates an example of how disks are physically and logically connected. The disk 1DUA0: is an HSC disk connected to VAXS. The disk
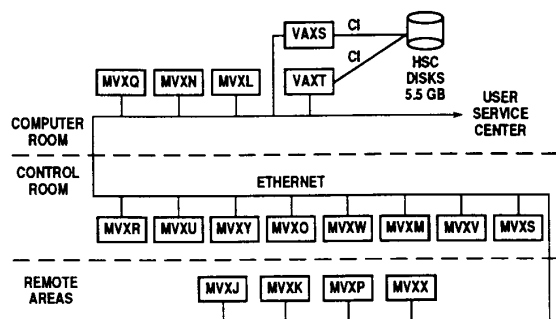

Fig. 2. DIII-D facility VAX computer systems.

Table 1

Node	Computer	Purpose
MVXJ	MicroVAX II	Lithium Beam diagnostic
MVXK	MicroVAX II	Multispatial divertor spectrometer diagnostic
MVXL	MicroVAX 3400	Fast Reciprocating Langmuir Probe diagnostic
MVXM	VAXstation 3100/38	Between shot auto and interactive analysis
MVXN	MicroVAX 3300	Spatial divertor spectrometer diagnostic
MVXO	VAXstation II/GPX	110 GHz ECH control system
MVXP	VAXstation II/GPX	60 GHz ECH control system
MVXQ	MicroVAX 3400	Thomson Scattering diagnostic
MVXR	MicroVAX II	Microwave Reflectometer diagnostic
MVXS	VAXstation 3100/30	VAX console system
MVXU	MicroVAX II	SPRED-spectrometer diagnostic
MVXV	VAXstation 3200	Between shot MFIT (MHD) analysis
MVXW	MicroVAX II	Between shot automatic analysis
MVXX	MicroVAX II	Charge Exchange Recombination diagnostic
MVXY	MicroVAX II	E parallel B charge exchange diagnostic
VAXS	VAX 6410	Data collection and management from MODCOMP and MicroVAX systems, between shot auto and interactive analysis
VAXT	VAX 8650	Data archiving, between shot auto and interactive analysis

2DIA0: is connected to MVXQ. However, because of the cluster software, each computer sees both disks connected locally. The cluster communication software is much more efficient for transferring data over the Ethernet than is DECnet software which has considerably more overhead. This has helped to speed up data transfer from the MicroVAX systems to VAXS by a factor of 2.

The central disks connected to the HSC storage controller are mounted on all nodes of the cluster. However, disks that are local to a MicroVAX are mounted only on the local Micro-VAX and on VAXS and VAXT. In general, it is not necessary for one MicroVAX system to access another. However, this can be accomplished using DECnet if required. By mounting disks throughout the cluster, disks on other systems can be manipulated directly without having to log on to those systems. This is also very useful for batch jobs that run overnight that perform certain disk cleanup and monitoring functions.

Disk Backups

One of the most time-consuming functions in a computer system is the backing up of disks to tape to maintain data integrity and to allow for the restoring of old files if necessary. With the MicroVAX systems occupying a variety of locations around
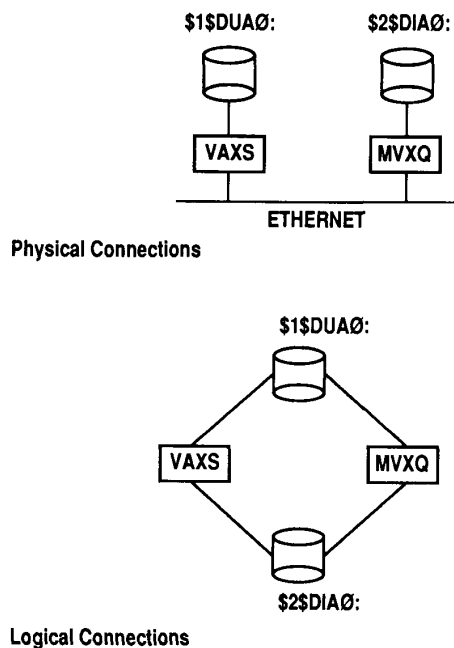
1DUAØ: **2DIAØ:**

VAXS   **MVXQ**

ETHERNET

Physical Connections

1DUAØ:

VAXS   **MVXQ**

2DIAØ:

Logical Connections

Fig. 3. Physical and logical appearance of disks in the cluster.

the DIII–D facility, a considerable amount of time had to be spent going to each one to do disk backups to a tape drive local to the MicroVAX. Also, these systems typically have TK50 tape drives, which are very slow and can only hold about 95 Mbytes each. Since some of the MicroVAX systems have 600 Mbyte disks or multiple disks, many TK50 tapes were often required to save the disks. Tape handling and storage was thus very resource intensive in terms of both human time spent and storage space for tapes.

About three years ago, 8 mm tape drives started to be used for data archiving. These tapes can contain up to 2 Gbytes of data on a tape the size of an audio tape cartridge. The tape drive was on VAXT and could only access local disks, that is the HSC disks. With the mixed interconnect cluster and having disks mounted cluster wide for the MicroVAX systems, it became possible for VAXT to access and backup disks from the entire cluster. Utilizing this cluster configuration greatly simplified the process of backing up disks and significantly reduced the amount of space required for storage of the tapes. This also frees up a considerable amount of personnel time which can be used more productively.

Full disk backups are done weekly on each disk in the cluster. Several computer system disks are grouped together in a procedure so that their combined capacity will mostly, but not completely, fill an 8 mm tape. Currently, there are four groups of this type. Each night, one group of systems is saved to tape by a batch procedure. The procedure knows which computers are in a given group and what disks are connected to each computer. The time required for the saves is not critical since they are being done after hours when the systems and the Ethernet are less busy.

Since full disk saves are done onto 8 mm tape which is attached to VAXT, a problem can occur if a MicroVAX system loses its local system disk, and the disk has to be restored. To

manage this problem, a dummy node was created called MVXT. Any MicroVAX can be booted into the cluster under this dummy node as a satellite. Once booted into the cluster, it is then possible to mount the node's local disk onto VAXT and then do a restore of the disk from 8 mm tape. Once the restore is complete, then that node may be booted normally from its own system disk which has just been restored.

Incremental backups save only disk files that have changed since they were last backed up. This type of save is still done to nine-track 6250 bpi tape and is done daily, Monday through Friday on either VAXS or VAXT. As with the full saves, the incremental saves are done centrally rather than on each individual system. A set of 31 tapes is used, one for each day of the month, thus the set is recycled monthly. A procedure is used which knows about all the nodes in the cluster and all the disks on each node.

VAX Console System

As illustrated previously, the various VAX systems are in different locations throughout the DIII–D facility. Each computer has a console port which normally is connected to a terminal which is used for booting the computer and also displays various system messages as events occur. To boot all 17 systems for example after a power outage and to be sure that they have booted correctly can take a considerable amount of time when one has to run around to each computer. This problem has been greatly alleviated by using DEC's VAX console system software. This software makes node MVXS the console for all of the other VAX systems. Thus, all systems can be halted, booted, or diagnosed from one location.

Figure 4 illustrates the VAX console connections. The console port of each computer is connected to a terminal server (DECserver 200) rather than a terminal. The terminal server is on the Ethernet as well as MVXS. The MVXS operating system defines pseudo terminals that correspond to the terminal server ports and, hence, to the console ports of each computer. With the VAX console software, the user can connect to a specific console port and, therefore, become the console for a specific computer. The software also utilizes DECwindows and, consequently, multiple windows can be brought up, each corresponding to one of the MicroVAX computers.

Another added benefit of the console system is that it logs everything that comes over the console line from all computers. Thus if a problem occurs, it is possible to go back and review the logs to possibly determine the problem. This logging is also useful for software installation by keeping a record of the installation. The console system also keeps a list of important events that appear on the console lines; these can be used to alert operations personnel of problems when they happen. This is particularly important with regard to the remotely located

CONSOLE MONITOR   **MVXJ**   **OTHER NODES**

MVXK   **CONSOLE LINES**

MVXS
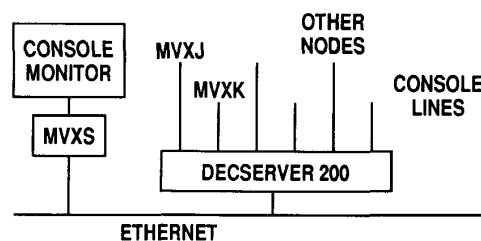
DECSERVER 200

ETHERNET

Fig. 4. VAX console system connections.

computer systems, since if one should halt for some reason, it will be noticed immediately on the console system rather than at some later time after which data may have been lost.

The console system also has a CDROM compact disk reader attached to it. .DEC now distributes all of their software on CD media. For operating system upgrades it is very convenient to boot the console system into the VAX cluster. The CDROM disk is then available to the cluster and significantly speeds up the process of system upgrades and other software installation.

Conclusions

As new plasma diagnostics are added to the DIII–D tokamak, the data produced by them is most likely to be acquired through MicroVAX based computer systems. The increasing number of these systems has led to the need for greater efficiency in acquiring data into the central database and to the need for better overall management of the systems. Adding the MicroVAX nodes to the central VAX cluster has helped to achieve these goals. The greater efficiency of reading data from MicroVAX systems via the cluster software over Ethernet prevents data acquisition from being a limiting factor for the shot cycle. The methods of doing disk backups have become very routine and insure data and system disk integrity and frees up personnel time for other work. The VAX console system has greatly simplified many aspects of system management and produces more immediate awareness of the status of the various computer systems.

Acknowledgment

References

[1] McHarg, B.B., Jr., "Real-Time Diagnostic Data Acquisition System for the Doublet III Tokamak Fusion Experiment," Proc. IEEE Transactions on Nuclear Science **NS-30** (1983) 5, 3787

[2] McHarg, B.B., Jr., "Hardware and Software Configuration of the Doublet III Diagnostic Data Acquisition Computer System," General Atomics Report GA-A17031 (1983).

[3] McHarg, B.B., Jr., "Expansion and Redundancy of the Doublet III Data Acquisition Computer System," Proc. 9th Symp. on Fusion Engineering (1981) 1011.

[4] Glad, A.S., et al., "Acquisition, Archiving, and Analysis of Doublet III Diagnostic Data on a Distributed Computer System," Proc. 11th Symp. Fusion Technology, Paper BP2-23, Oxford (1980).

[5] Glad, A.S., et al., "Real-Time Acquisition, Processing, and Archiving of Doublet III Diagnostic Data Employing Table Driven Software," Proc. 8th Symp. on Fusion Engineering (1979) 1964.

[6] McHarg, B.B., Jr., "Distributed Functionality and Configuration of the DIII–D Data Acquisition and Analysis System," Proc. 11th Symp. on Fusion Engineering (1985) 594.

[7] McHarg, B.B., Jr., "Big Dee Upgrade of the Doublet III Diagnostic Data Acquisition Computer System," Proc. 10th Symp. on Fusion Engineering (1983).

[8] McHarg, B.B., Jr., "Data File Management in the DIII–D Data Acquisition and Analysis Computer Systems," Proc. 13th Symp. on Fusion Engineering (1989) 185.

[9] McHarg, B.B., Jr., "The Use of Networking in the DIII–D Data Acquisition and Analysis Computer Systems," Proc. 12th Symp. on Fusion Engineering (1987) 636.

[10] Henline, P.A., "The DIII–D VAX Computer System," General Atomics Report GA-A19594 (1989).