AT. 39.2/014

The Australia Telescope Data Archiving Problem

S.W. Amy, D. McConnell & R.M. Wark

28 April, 1993

1 Introduction

Compact Array data are copied to exabyte tapes for both archive and export. Certain difficulties with exabyte usage have been experienced by Observatory staff and visitors. This reports investigations of this problem undertaken at Narrabri over the period from 1 April to 8 April, 1993. The aim of this work was to characterize the problem and to make recommendations involving either solving or minising the problem.

2 System Architecture

In this section we present a brief overview of the current on-line system architecture. This differs markedly from the original ATCA on-line computer design.

Data is written (in real-time as RPFITS files) to one of two identical SCSI disk drives attached to the Correlator Control Computer (SANCHO). The observer or duty astronomer typically chooses at the beginning of the observation which disk has the most empty space and assigns this as the "DATA WRITE" disk. At the conclusion of the observations a personal copy of the data is made for export. This is done using a DCL command procedure which uses the VMS `COPY` command to write data directly to Exabyte 8mm tape. Both the on-line data disk and the Exabyte drive are SCSI devices accessed by a quad-height Emulex UC-08 controller which emulates the DEC MSCP (mass-storage control protocol) and TMSCP (tape mass-storage control protocol) devices. This allows the use of standard VMS device drivers (`DUDRIVER.EXE` for the disk and `TUDRIVER.EXE` for the Exabyte tape drive). Tapes in this format can be read on any machine which is capable of reading ANSI labelled tapes and can be read directly by the *AIPS* task `ATLOD`. In most cases the copy is made during further use of SANCHO for observations. At a later stage Observatory staff will make a further copy (again using VMS `COPY`) for the ATCA archive.

SANCHO is a MicroVAX II which has a number of time critical tasks to perform in addition to the (lower-priority) data archiving, including high-level control of the ATCA correlator and delay units (through CACOR), processing of the visibility data using a Sky Warrior Array processor and real-time communications (using DECnet task-to-task communication over Ethernet) with the higher level array control task, CAOBS (which runs on NOEL, a MicroVAX 3100 Model 80). The process priority scheduling is done by assigning the various real-time processes VMS priority levels which should ensure data archiving does not interfere with array and correlator operations. The machine is relatively heavily loaded.

This scheme had been working reasonably reliably during 1992 but over the past few months copies had become less reliable. Many possible causes were suggested. Possible

causes which were considered to be reasonably likely but, due to time constraints, were not fully investigated are:

- copies were less reliable when the machine was under heavier CPU and I/O load (e.g. during spectral line observations),
- the operating system version (VMS 5.4-3) was different to the machines in the offline and online VMSClusters (VMS 5.5-1),
- the Q-bus was possibly suspect (see later),
- the brand and quality of tape that was used seemed to be a significant factor, particularly in the later stages.

In addition to the Exabyte drive on SANCHO, the observatory has two Exabyte drives connected to VAXstations in the off-line cluster and a further two drives connected to Sun SPARCstations used for off-line data analysis. The Exabyte drive connected to DESK2 (a VAXstation 4000VLC) would successfully read tapes but recently the drive would not write tapes. The second drive was connected to WSENG1 (also a VAXstation 4000VLC) and was being using regularly with VMS `BACKUP` successfully. Both these workstations have inbuilt SCSI controllers.

3 Tests and Results

Each of the three drives were tested in a variety of ways. Initial tests concentrated on the two drives on the off-line cluster as scheduled observing was in progress.

3.1 The DESK2 Exabyte

As mentioned, this drive (connected to a VAXstation 4000VLC) failed to write with either VMS `COPY` or VMS `BACKUP` but would read tapes successfully. These results were confirmed and in addition, tapes could not be initialized. As an additional test, the drive was moved to another machine in the off-line cluster, a VAXstation 3100 (JOVIAL) which also has an inbuilt SCSI controller. The drive failed in the same manner.

To determine if the problem was related to either the VMS operating system or controller hardware the unit was moved to a DECstation 3100 running Ultrix and although existing tapes (written using `tar`) could be read, new tapes could not be written using the standard `tar` utility. At this point it was decided that this unit was definitely faulty and should be considered for repair or replacement.

3.2 The WSENG1 Exabyte

Although this drive seemed to be functioning as expected there had been the suggestion that this drive would only write VMS `BACKUP` save-sets and could not be used using VMS `COPY`. Our tests confirmed this, although there is a suitable workaround that can be used to write VMS `COPY` tapes on this drive.

Providing, the *first* file on the tape is written using VMS `BACKUP`, the tape can then be used for VMS `COPY`. As an example consider the following sequence of commands below assuming the tape drive is `MKA600:`.

```
$ CREATE IGNORE_THIS.TXT
Ignore this file
<CTRL-Z>
$ INITIALIZE MKA600: TEST
$ MOUNT/FOREIGN MKA600:
$ BACKUP/LOG/LABEL=TEST IGNORE_THIS.TXT MKA600:IGNORE_THIS.BCK
$ DISMOUNT/NOUNLOAD MKA600:
$ MOUNT MKA600: TEST
$ COPY/LOG <filename>.<ext> MKA600:
$ DISMOUNT MKA600:
```

Apart from this *feature* which may be due to the interaction of one or more of the inbuilt SCSI controller, VMS device driver, firmware revision of the Exabyte drive and operating system version the drive works reliably. This may still be an outstanding problem which may need further investigation.

3.3 The SANCHO Exabyte

As discussed this is the most critical Exabyte of the three and probably the most used as it writes two copies of all ATCA data using VMS COPY. Once scheduled observing had ended at the commencement of the April maintainance shutdown, the failures were observed easily with both types of tape thus ruling out the hypothesis that tape brand was significant. The first priority was to copy the last set of data to tape for the observatory archive. This proved *impossible* using the SANCHO Exabyte and as a last resort the data were copied using DECnet to LEON and archived (manually) using the DESK2 drive. The copying failed whether the source was from a SCSI disk (i.e. that attached to the same controller) or from the internal DEC disk/controller combination.

Initial tests concentrated on checking cabling and connections but nothing in this area looked as if it was the likely culprit. It was then realized that the Emulex UC-08 SCSI controller had on-board diagnostics which were capable of testing many aspects of the drive, including tape movement, end-of-file mark processing, reading and writing. These tests are run from the VAX monitor and are therefore independent of VMS. The tests suggested that there was nothing wrong in either the controller or tape drive as no failures were reported after 12 or so iterations.

Next, it was decided that as the drive/controller seemed to be functioning correctly, the problem may lie in an intermittent fault in one of the other boards in SANCHO. Sequentially the following boards were removed and the Q-bus rearranged to satisfy the electrical requirements: TK50 tape controller, Ethernet controller, Correlator/Q-bus interface, Sky Warrior array processor. After each board was removed, the machine was rebooted and tested using VMS COPY. On every occasion the copy failed which suggested that the on-board diagnostics should not necessarily be trusted!

Next, SYSGEN was invoked to check that no devices had clashing CSR or Vector addresses and as expected everything appeared normal. As a last resort, it was decided to swap Exabyte drives, with the one from WSENG1 being removed from its enclosure and exchanged for the unit in SANCHO after appropriate jumpers were changed. Surprisingly this proved successful and many attempts could not make a VMS COPY actually fail thus suggesting it was a physical drive problem from the outset.

Hindsight is a wonderful thing and many could (justifiably) say that swapping drive units should have been one of the first things attempted. For many reasons we considered the problems to be elsewhere.

4 Conclusions

Following the tests the two faulty Exabyte drives (the original SANCHO unit and the DESK2 drive) were sent to XSI for evaluation and possible repair. XSI reported that the DESK2 unit would fail to write file-marks successfully, consistent with the failures observed on DESK2 and under Ultrix. The old SANCHO unit had a high number of write errors due to possible head wear which is consistent with a drive under heavy usage. It may also explain why VMS `BACKUP` could still use this drive whereas `COPY` failed – VMS `BACKUP` does a reasonable amount of error checking and can recover from errors using a CRC which is computed at the time of writing whereas `COPY` does a raw copy with little checking. Both units are being repaired for a fixed cost.

Serious difficulties were encountered in reassembling of SANCHO. Some of the boards had to be re-seated several times before the array performed normally. Particular difficulties were experienced with the correlator/Q-bus interface board (built in-house). We suspect that either the Q-bus backplane a board has an intermittent fault. The SANCHO system hardware is **very fragile**. We suggest that a spare correlator/Q-bus interface board be verified and kept on hand and that an alternative backplane/cardcage assembly be identified for use in case of a complete SANCHO failure.