Astronomical Data Analysis Software and Systems IX
ASP Conference Series, Vol. 216, 2000
N. Manset, C. Veillet, and D. Crabtree, eds.

Using DVD-R for Storing Astronomical Archive Data (cont'd)

B. Pirenne

Space Telescope – European Coordinating Facility, European Southern Observatory, D-85748 Garching, Germany

Abstract. At last year's ADASS the status of the DVD technology and its prospects for storing astronomical data for the long term were presented.
This time, I want to review the current state of our (still) favorite successor of the CD-R for astronomical data storage: the DVD-R. The second generation with full 4.7 GB density per volume has finally appeared, the software to write the media is also available for a variety of computer platforms. The media price is going down to the level of 40USD per media, bringing the price/GB of medium size archives with the newest jukeboxes down to about 15USD/GB. Media producers are now also seriously considering the introduction of double-sided media, which would take this price further down, to about 11USD/GB, far less than last year's most optimistic estimates.

Our current experience with first generation media is now very positive, to the point that ESO is now implementing the production of DVD-R archive media directly at the observatory.

1. Introduction

To address the growing capacity needs of the Paranal, La Silla and HST archives at ESO (see Pirenne & Durand 1995), a new affordable solution had to be found. Coming from a CD-R storage system, it was only natural to try and move to the next available medium, the DVD-R.

This particular storage device has all the advantages of the CD-R: same form factor, a well defined recording and file format standard, the support of a whole range of companies producing readers (even if only one company is producing the recorder) and a large choice of jukeboxes. Also, thanks to the very affordable price of these readers, DVD-R could be seriously considered as a data distribution medium.

Other DVD technologies such as DVD-RAM are not currently as advanced in terms of capacity and do not offer the same form factor -a caddy is required- or the possibility to be read by any DVD player. Moreover in the case of such re-writable media, the durability of the recorded information is unknown.

After a long wait and various dissapointments (see Pirenne, Albrecht & Schilling 1998), the bet we had made on the DVD-R technology finally produced results. We managed to understand and overcome the media writing problems and the media reading problems; we have a solid software support as well as reliable jukeboxes. The various datasets copied on those media have been offered to the public successfully for now over 6 months through our web interfaces and data request handler.

In this short report, I will elaborate on some of the problems encountered and advocate DVD-R as a viable and cheap alternative to other data archive and distribution media.

2. Current Use Of DVD-R At ESO

2.1. Our Recipe

For those technically inclined, the configuration we have now in operation consists of the following items. The choice is based on our experience and is now working well.

- Solaris 2.6 based Ultra SPARC computers;
- Pioneer DVR-S201 recorders;
- Kodak Disc Transporters to automatically feed the recorder with blanks;
- ASM 1306 DVD jukebox (and planning to acquire the Pioneer DRM-7000 as well). The jukeboxes are controlled with the "HyperRom" commercial software from Tracer Technology Inc.;
- The public domain `mkisofs` software to prepare our ISO9660+RockRidge disk image used in conjunction with a `cdrecord`-based DVD recording software. Some home grown software and procedures automatize the production of media, now mostly unattended. The recording does not take place inside the jukebox;
- Stand alone readers proven to work on Solaris and Linux include the Panasonic LF-D101 (a SCSI DVD-RAM device), the Matsushita SR-8171 (an Atapi model) and the Pioneer DVD-303 (SCSI). Only the Pioneer model supports the v1.9 4.7GB disks;
- Blank media that work include the Pioneer 3.95GB and 4.7GB v1.9; the TDK 3.95GB and 4.7GB and the Mitsui 4.7GB media. The BASF 3.95 did not work.

2.2. What Did We Do so Far?

Since the end on 1998, a number of media have been recorded and introduced into our jukebox: about 80 DVD-R with second Digitized Sky Survey (DSS-2) data, 17 with DSS-1, 30 with NTT data, 10 with HST data, 35 with ESO 2p2 Wide-Field Imager (WFI) data. Those sets are offered since months and have caused little if any problem. We have started to accelerate the recording and will soon transfer the entire HST and WFI archive on the new media.

3. Facing The Data Rate

But the growth is really the issue. At the moment the ESO observatories are feeding the archive at a rate of about 100GB/week! 70% of the data comes from the ESO Wide Field Imager CCD Mosaic (WFI) and is still mostly on DLT tape cassettes.

The current plan calls for compressing all the raw data and for the recording of everything (including the WFI data) on the mountain directly on DVD-Rs. Such a system would deliver about 11 new media every week. This would be acceptable as it would fill roughly one jukebox per year.

This rate will however increase rapidly as more and more instrument are put on-line at the VLT. Therefore, in the long run, another solution will have to be found: higher density media, more segregation (seldom used data is migrated to tapes or kept outside jukeboxes), migration to yet another type of medium with higher density, etc.

4. DVD-R As A Distribution Medium

As was mentioned in the introduction, DVD-R has proven a potentially very good medium to distribute data from our archive center or directly from the observatory to the users or the PI. The media (in particular the 3.95GB one) are readable on a large variety of very cheap readers, well supported by quite a variety of Unix systems. MS Windows users can also make use of them. The choice of the more widely know ISO9660 system over the less popular UDF system is off course helping the issue, although we expect the generalization of UDF support in the future.

We are therefore going to propose PIs of ESO Programs in the future the choice of either CD-R or DVD-R to receive their service-mode data. Similarly, large archive data requests will be offered on DVD-R too.

5. Conclusion: How Expensive Is Say, A 10TB Archive?

Table 1 lists the various ways available today to keep a 10TB archive available [quasi] on-line. Prices are indicated for both the on-line and "shelved" solution. Random access time to a file is also provided.

Table 1. Various possibilities and price comparisons for keeping a 10TB archive on-line

Technology	Vol. cap.(GB)	vol. $/GB	JB slots	JB $/slot	seek (ms)	10TB (K$)
50 GB Mag disk	45.0	43.0	N/A	N/A	10	439
DVD-R 3.95GB	3.7	8.4	670	25.5	5000	157
DVD-R 4.7GB	4.5	9.5	670	25.5	5000	155
CD-R 650MB	0.6	1.4	500	19.3	25000	328
DLT7000	33.5	1.8	420	299	90000	111

Figure 1. The main ESO/ECF archive jukebox: an ASM 1306 jukebox with a capacity of 1083 slots and a 5 seconds turn-around time for the media.

References

Pirenne, B., & Durand, D. 1995, in "Information & On-Line Data in Astronomy", ed. D. Egret and M. A. Albrecht, Kluwer, 243

Pirenne, B., Albrecht, M., & Schilling, J. 1999 in ASP Conf. Ser., Vol. 172, Astronomical Data Analysis Software and Systems VIII, ed. David M. Mehringer, Raymond L. Plante, & Douglas A. Roberts (San Francisco: ASP), 269