[76.01] How Many Terabytes Was That? Archiving and Serving Solar Space Data Without Losing Your Shirt

Joseph B. Gurman
Laboratory for Astronomy and Solar Physics, NASA Goddard Space Flight Center

Abstract

Even solar missions of modest size in the next decade will produce terabytes (10^{12} bytes) of data. The Solar Data Analysis Center is already dealing with mission archives of similar volumes, and is serving the entire archives to the community over the Internet.

We examine present and near-term archiving strategies and media, and conclude rather surprisingly that online storage on network-attached RAID arrays is the most cost-effective, as well as the most usable, archiving method likely to be available over the next decade for keeping and serving scientifically useful data for a period of 10 years or more.

I.Why worry about about Terabyte datasets?

Both TRACE and SOHO currently produce over 100 Gbyte of scientifically useful data per year. Here, I use “scientifically useful” to refer to reformatted data in consensus (e.g. FITS) or de facto (e.g. Yohkoh) standard formats supported by a widely available body of access and analysis software, such as Solarsoft (Freeland and Handy 1998, *Solar Phys.*, 182, 497).

Solar missions missions planned and proposed for the next decade, some quite modest in cost compared to SOHO, will produce considerably larger data archives.

Table 1. Expected raw archive sizes, solar physics missions

with current formatted archive sizes for comparison

Mission	Launch	Tbyte/year	Tbyte (baseline)	Tbyte (5 years)
HESSI	2000	0.50	1.00	2.50
ASCE	2003	0.16	0.32	0.80
Solar-B	2004	2.00	6.00	10.00
STEREO	2004	0.40	0.80	1.00
Yohkoh	1991	0.05	0.15	0.25
SOHO	1995	0.30	0.60	1.50
TRACE	1998	0.16	0.24	0.80

Why worry about about Terabyte datasets? (continued)

Notes on Table 1:

ASCE is currently in pre-"downselect" Phase A and may be chosen for a 2003 or 2004 launch.

The SOHO figures exclude MDI high-rate data, the volume of which equals that from all other instruments.

The actual sizes of most new archives do not include decompression, but TRACE data are stored in compressed form in binary table FITS files. In practice, this could mean archives a factor of 3 - 10 larger (using lossless or h-compression, respectively).

There are nearly 7 years of *Yohkoh* data in the public archive already.

Thus it is possible that solar space data of active research interest a decade from now will total some 20 Tbyte. This is small compared to the datasets envisioned for earth sciences (which will accumulate at a rate of ~ 1 Tbyte day^{-1}, but significant if served to the Internet, as some mission databases currently are.

The service methods in use now are offline (*e.g. Yohkoh*), near-line (*SOHO* MDI), and online (*SOHO*, TRACE), but all include online catalogs. For a given storage hardware implementation, a scientifically useful *online* archive is clearly less expensive for a given storage strategy than methods involving human servicing of data requests, even if the humans involved are students. More likely, they are programmers or system administrators whose skills are currently valued highly in the commercial marketplace.

Why worry about about Terabyte datasets? (continued)

Even though we can safely assume that the costs of any storage hardware still in production several years hence will only drop, archiving costs should still be considered in the mission planning phase, lest

I therefore examine several methods for serving data sets of order 1 - 10 Tbyte in size, in order to determine the least expensive and most cost-effective methods.

II. Assumptions

For the purposes of this brief study, we assume that:

1. The real cost , including benefits, overhead, &c., of a person-year of effort ("full-time equivalent" = FTE) is $100K for an entry-level programmer and $150K for a base programmer or system administrator.
2. Access to the Internet is likely to continue to be "free" (i.e., not directly costed to projects) at universities and federally funded institutions for the foreseeable future, and will allow the transfer of tens of Gbyte of data per day with little impact.
3. Solar physicists are patient enough to wait ~ 1 day for data to be spooled online from nearline or offline storage and then download, but would prefer direct, online access if possible. (I.e., we're just as addicted to immediate gratification as every other Web user.)
4. NASA will continue to pay for access to archival data for up to 10 years after the end of a mission if the scientific demand warrants, and the putative Sun-Earth Connections Data System can afford it.

III. Strategies

Given these assumptions, we examine several possible strategies (currently in use or feasible) for serving archival mission data:

1. **Tapes on the wall** - offline storage, with human servicing of requests generated *via* a Web form, and tracking of which data reside on which tape through either commercial database management software or homegrown utilities

2. **Nearline/offline/attached** - Fast, high capacity tape storage with disk space attached to an interactive server for staging; tracking as above

3. **Attached Write-Once (WO)-ROM** - Large (~500 slot) CD-ROM jukeboxes with multiple drives, attached to an interactive host; tracking as above. Will also attempt to extrapolate DVD-ROM library costs from this model.

4. **Network-served WO-ROM** - CD-ROM's in network-attached drives served by an NFS "appliance" with no interactive operating system

5. **Attached RAID** - Data are striped across multiple hard drives and the RAID controller is attached to an interactive server, as in (3), above.

6. **Network-served RAID** - RAID arrays as in (5), served to the network by a non-interactive "appliance," as in (4).

Note that approaches (3) - (6) should not require human intervention for serving requests of reasonable size, but sys admin issues are different.

IV. Comparison

We do not consider the original archive population cost, which will be higher with individually generated tapes or CD-ROM's than with RAID because of the need for some human interaction. Similarly, we do not consider offline/offsite backup costs, as they are similar for all methods, though the incremental cost is lower for tape- and CD-ROM-based approaches.

We assume that any interactive server with multiple users with multiple applications will require a minimum of 0.3 sys admin FTE yr^{-1}. Real costs will probably be higher, because no one wants to keep the same interactive server for more than 3 - 5 years, much less 10.

Any currently unforeseen costs ("the next great thing") are assumed to be identical in all cases, and commercially popular technologies such as CD-ROM are assumed to be sufficiently stable that obtaining replacement drives will not be a problem in ten years.

Any approach based on tapes must include the development of a minimal database identifying the location of each discrete unit of data on a discreet tape.

IV. Comparison (continued)

We can then break down the life-cycle (10-year) costs of each strategy, assuming an archive sizes of 1 and 10 Tbyte in each case. For the larger archive, we will assume technology currently within one year of commercial availability will be applicable.

Table 2. Costs ($K) for various archive approaches, 1 Tbyte

Component	Tapes on wall	Near/off tape	Att. WO-ROM	Net. WO-ROM	Att. RAID	Net. RAID
media	1	1	65	65	214	260
production	250	500	250	250	0	0
storage	2	150	20	650	0	0
sys admin	155	450	450	75	450	75
database	187	187	0	0	0	0
Total	595	1288	785	1040	664	335

Notes:

Sys admin costs for tape solutions include 0.25 FTE/yr for data techs to swap tapes.

Current WO-ROM technology is CD-R.

Nearline/offline tape is based on the SOHO-MDI archive at Stanford University.

Attached WO-ROM is based on the SOHO-LASCO archive at the Naval Research Laboratory.

Network-served RAID is based on the SOHO and TRACE archives at the SDAC.

RAID solutions are based on 18 Gbyte UW or FC drives.

IV. Comparison (continued)

For the 10 Tbyte archive, we assume the general availability of storage technology currently available in limited quantities (*i.e.*., generally available in ~ 6 months): AIT II tapes, 5.2 Gbyte DVD-RAM jukeboxes, 50 Gbyte UW SCSI or fibre channel hard drives.

Table 2. Costs ($K) for various archive approaches, 10 Tbyte

Component	**Tapes on wall**	**Near/off tape**	**Att. WO-ROM**	**Net WO-ROM**	**Att. RAID**	**Net RAID**
media	10	10	1.5	1.5	214	400
production	250	500	250	250	0	0
storage	4	150	16	522	0	0
sys admin	155	450	450	75	450	75
database	187	187	0	0	0	0
Total	606	1297	717.5	848.5	664	475

Notes

All absolute costs cited here are probably pessimistic for 2005 - 2010, since storage costs continue to decline. Hard drives prices have declined the most steeply.

V. Conclusions

- Since 1993, when a comparison of this type was first carried out for the *SOHO* archive, network-attached RAID has consistently offered the most attractive price. This approach also offers many other attractive features, such as a logging file system ("snapshots"), and shares with attached RAID unusual data integrity and automatic failover (hot spares). Even lower-cost alternatives are possible with ABOD (not RAID) network-served devices, but they do not currently scale to more than 750 Gbyte (with 50 Gbyte drives).
- This approach also lowers initial archive creation cost (*via* simple mirroring from or direct writing by the PI team's reformatting system) and simplifies access for local area users (who can mount the system *via* NFS) even if they use diverse platforms (typical in a multiple-experiment team mission).
- Perhaps the most important features of network-attached RAID are its low system administration overhead, and absence of pressure to upgrade from competing applications (as on interactive servers). Since people costs are the first cut from archiving efforts, the lifetime of archives based on more labor-intensive models is liable to be curtailed.

References

G. Hurford *et al.*1999, *HESSI Project Data Management Plan (PDMP)*, U. of California, Berkeley

ASCE Web site (`http://cfa-www.harvard.edu/asce/spcraft.htm`)

Solar-B Announcement of Opportunity (`http://spacescience.nasa.gov/ao/98-oss-05/`)

STEREO Announcement of Opportunity (`http://spacescience.nasa.gov/ao/99-oss-01/`)

The author would also like to acknowledge informative dicsussions with R. Bogart, O.C. St. Cyr, and D. Wang, as well as with the representatives of various firms only too willing to sell storage solutions to government and universities.