

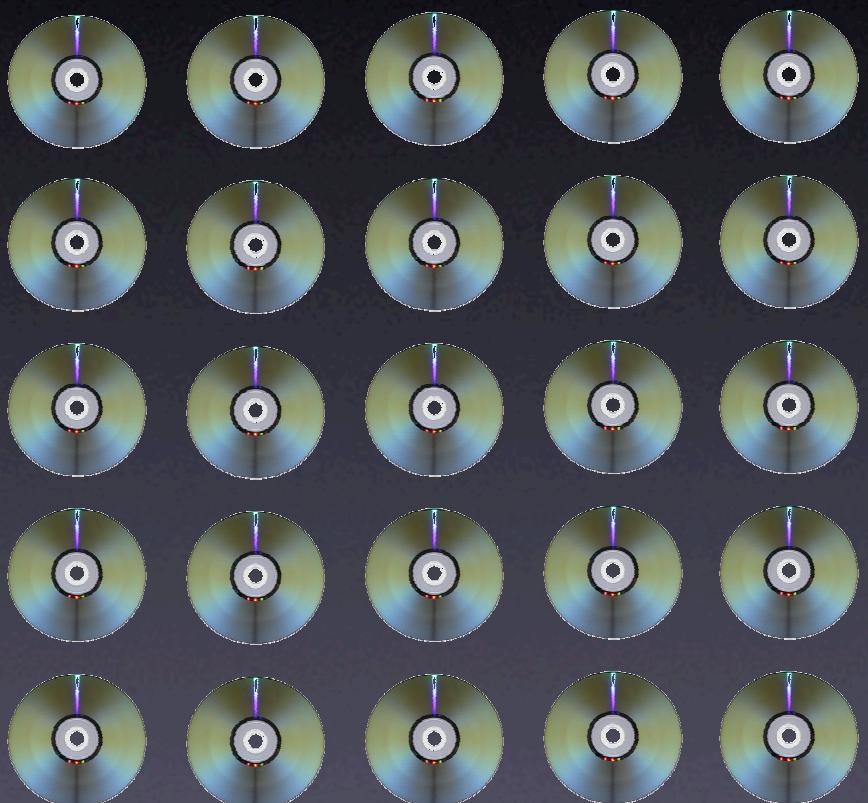
media-prep.pl

A tool for preparing large data sets for archival.

End of Cruise



50+ Gb
Data Store



media-prep.pl

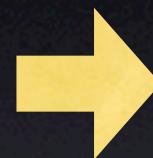
...sorts files by date and divides them into volumes.



50+ Gb
Data Store



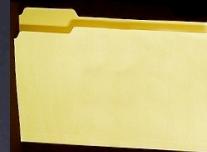
Sort
files by
date



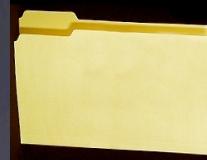
HLY0601



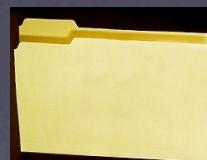
4.5Gb



4.5Gb



4.5Gb



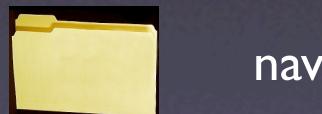
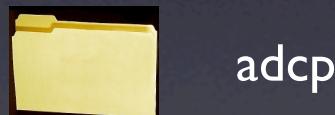
4.5Gb

Divide
into
volumes

media-prep.pl

...creates an identical directory structure in each volume.

Source Directory Tree



vol-1

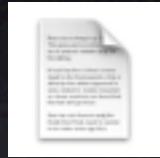


vol-2



media-prep.pl

...generates an index for each volume and a master index for the archive.

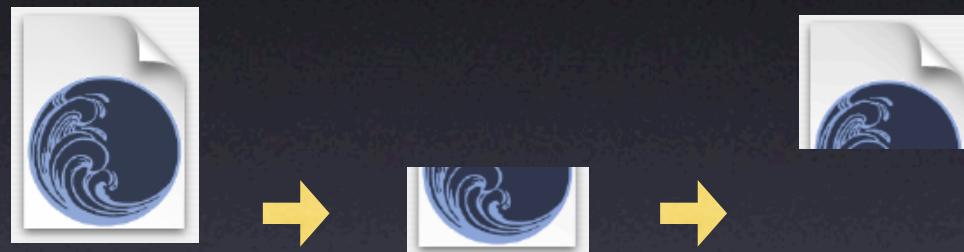


vol-1.index

```
# HLY-04-04 --- vol-1.index
#
# For each file in this volume, this index provides the filename, its last
# modification date-time stamp (mtime), the results of cksum
# (e.g. a 32-bit checksum and a block count).
# DATE          CKSUM  BLKCNT FILENAME
Thu Sep 2 23:59:14 2004      361685543 449116 /Volumes/Healy-340409420/HLY-04-04/media-vol1/scs/Datalog/gyro/
Gyro_20040902-182509.RAW
Thu Sep 2 23:59:14 2004      624661052 974255 /Volumes/Healy-340409420/HLY-04-04/media-vol1/scs/Datalog/
PCODE_Bridge_gga/PCODE-Bridge-GGA_20040902-182509.RAW
```

media-prep.pl

supports iterative and incremental runs by
...only copying fractions of files not already copied.



...resorts and reshuffles the files when additional files
are added to the source archive with old dates.



media-prep.burn

...allows one to burn media as volumes fill up.

HLY060I



vol-1



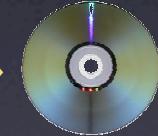
4.5Gb



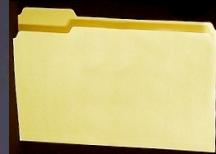
vol-2



4.5Gb



vol-3



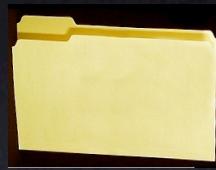
4.5Gb

vol-4



4.5Gb

HLY060I



vol-1



4.5Gb



vol-2



4.5Gb



vol-3



4.5Gb

vol-4



4.5Gb

[http://www.ldeo.columbia.edu/~vschmidt/
scripts/media-prep.html](http://www.ldeo.columbia.edu/~vschmidt/scripts/media-prep.html)