Efficient Media Migration for Temporally-Stored Data

Based on a practical implementation from CLASS' LTO-6 to LTO-8 migration

Ben Sautter
CLASS Archive Manager (Contractor)
What is CLASS?

Comprehensive Large Array-data Stewardship System

- An archive of NOAA environmental data from NOAA/DoD POES satellites, NOAA GOES satellites, JPSS, JASON, and other derived data products.
- Public users can sign-up for an account and search for datasets to order by date, type, satellite, etc...
- Orders are packaged and distributed asynchronously via HTTP and FTP, typically within 24 hours.
- Subscriptions are available for automatic ordering of new data.

www.class.noaa.gov
Hardware and Holdings

Data is received and duplicated in two independent NOAA locations: Boulder, CO and Asheville, NC
Each equipped with:

- SpectraLogic TFinity w/~7,000 slot partition for CLASS
- 36 - 48 LTO tape drives
- HPSS 7.5.3u6 - 1 core, 4 data movers

Data To Migrate:

- 384 million files
- 11.05 PBs
- 11 file families
- 3,230 LTO-6 tapes
Ingest
- CLASS is an archive that receives data from providers continuously and writes it within hours of when it was recorded.
- Tapes will naturally tend to store data for a small time range ("temporally")
- We very rarely delete and data often remains permanently unchanged.

Dissemination
- CLASS users order data for a time period (e.g. May-June 2018)
- If data is spread across multiple tapes, order fulfillment takes longer as more tapes need to be mounted.
- Therefore, performing a media migration between tapes while keeping similar dates together, minimizes order fulfillment time.
Auto Migration

Running multiple repacks at once with auto settings, can result in data being scattered across destination tapes for longer staging times.

User Order: May-Aug
Temporal Migration

If instead we run multiple repacks tied to specific source and destination tapes, it can result in data being optimally organized for quick staging.
A key feature is the ability to have “repack” choose specific:

- Source tapes (-S)
- Destination tapes (-D)

Then we provide repack with a date-ordered list of source tapes and have it start filling up empty destination tapes one-by-one.

Source tapes can span between two destination tapes so that every tape gets completely filled.

Efficient. But slow.
Grouped Repack

If instead we grouped source and destination tapes together into batches, then multiple instances of repack can each run on their group in parallel.

The difficulty is with choosing the groups. Only whole tapes can be grouped together, so it has to be done efficiently.
Grouping Challenges

- LTO-6 = 2.5 TB
- LTO-8 = 12.0 TB
- (24) LTO-6 = (5) LTO-8 (ratio of 4.8:1)

- Some tapes may not be completely filled
- Some tapes may have had data deleted
- Compression ratios can vary over time and within data types
- Some had mixed dates from prior indiscriminate repacks

The Goal: is to determine how to maximize performance with many groups running in parallel, but minimizing unused space on the destination tapes - the ".8" of the "4.8" ratio.
For each file family's tapes:

- Determine midpoint date of each tape and order by that date.
- Determine compressed volume of each tape, which should theoretically be the same between these LTO generations.
- In date-order, start "virtually" filling one or more destination tapes, and attempt to hit a minimal amount of slack space.
- Choose the best combination.
Simple Method

Transfer as many source tapes as possible into a single destination tape without overflow.
Medium Difficulty Method

Transfer as many source tapes as possible into a set of "X" destination tapes.

Earliest

x19

Latest

x14

x29

x4

x6

x3

100%

0%

100%

0%

100%

0%
Hard Difficulty Method

Determine what combination of every grouping results in the least amount of overall slack space.

Earliest

Latest

100%
0%
100%
0%
Hard Algorithm

Consider every combination of source tapes and destination tapes.

Group #1

Group #2

Group #3

Group #4

Group # to Insanity and Beyond....
Results

CLASS settled on 162 groups to divide up:
- 11 file families
- 3,230 source tapes

Even if tapes were processed one-by-one instead of in groups, it is still possible that the last tape will not fill up completely.
- Worst case the slack would be $\leq 1.2\%$.
- CLASS achieved a slack space of approx $1.4\%$.

Migration completed in 164 days using 20 group repacks in parallel.
Avg Rates:
- 67.4 TBs/day
- 2.3 mil/day
- 19.7 tapes/day
Migration Challenges

Recommend 7.5.3u6 to implement this method, which corrects a number of issues:

- Not being able to use a list of destination tapes without filtering for EOM tapes.
- Repack fails to use only the destination tapes listed.
- Support for using destination tapes from default file family "0" was removed but restored.
- CRC usage between LTO-6 and LTO-8 when using E2EDI.
Contact info:
- Benjamin Sautter
  ben.sautter@noaa.gov

For more info on CLASS:
- https://www.class.noaa.gov