

Internet Archive

Frontiers of Frugality

John C. Gonzalez

Director of Engineering and Accessibility

jcg@archive.org



The Internet Archive:

Non-Profit Library

Universal Access to All Knowledge

Internet Archive: Lots of data

- 300 billion web pages (700M web-captures/week)
 - 14 million books and texts (plus 1000 books scanned/day)
 - 4 million audio recordings (including 160,000 live concerts)
 - 3.6 million videos (including 1.4 million TV News programs)
 - 1 million images
 - 100,000 software programs
- 45 PB of unique storage (mirrored in a ~100 PB cluster)
- 25,726 spinning disk drives

To Reprise our 8T SMR Experience...

- Introduced to our production cluster in January 2016
- NOTE: these drives are NOT recommended by Seagate for this app!
- BUT: Great \$/TB → ~\$32 list price (raw, formatted)
- Drive write speeds slow considerably when capacity reaches ~80%
- Took some work to get this model stable on our platform
 - Without workarounds, problems with AR17 firmware
 - Sorted out firmware issues with Seagate and
 - Developed some work-around configuration settings
- Stable & happy: 8640 SMR disks in service; +5400 in special project

Storage Excursions in 2017/2018

- Storage expansion of ~7PB unique (~14PB raw)
- Rolling upgrade of “standard 36-disk node” → 32 core 10Gbps
- EOL announced for the 8T SMR drives we were using (SGx002)
- Intended to shift to latest gen (SGx003) 8T SMR Archival drives
 - New version goes from 6 platters to 4; lighter and lower cost
 - NOT really recommended by Seagate for our application
 - Qualified 4 drives in the HDFS cluster (non-seq. writes), then...
 - Put 360 into production, and... We found a frugality frontier!

4 worked: What Could Go Wrong?

Q: When is a disk drive not a storage device?

A: When what you read back is not what you wrote!

- In caution, we introduced the new drives alongside the proven ones
 - 2 racks: each with 10 identical nodes; 5 with new model, 5 with legacy model
- We're an archive. We always do a high-sample rate of immediate file-level checksum read-verify after writing item files...
- On filling the drives, bad-checksum errors on write verification... Only on nodes with new drives. Start of file fine, then only "00's" after some $n \cdot 4K$ boundary
- Low frequency: about 1/node/30 minutes (remember 36 drives/node)
 - Not only that, if you waited a while, bad reads would magically self-correct!
 - Not only THAT, at about 5% of capacity, ALL drives would STOP the bad behavior
- **So... We called Seagate → very supportive, and wanting "data trail" of proof**

Snagging the Golden Anchovy

- Seagate sent a tech armed with bus-capture gear...
 - Could do simultaneous capture across 4 drives
 - Capture buffers are not super-big...limited history!
- But we have 36 drives on the bus, and very rare events!
- Our application will not run with < 36 drives... So we:
 - Set up a continuous, firehose FILL of a data node
 - Modified the application to HALT when a checksum fault is detected on any of 4 selected drives within the node.
 - And we set up data capture on those 4/36 nodes (32 others still in service)
- Loaded up factory-fresh drives and let it rip... Took a while to trigger!

Resolution and Future Directions

- We could successfully get fault data (once a week or so)
- Seagate analyzed and fixed the F/W problem
- We have decided to sit out the 8T V11 SMR drives
- Looking to upcoming 14T drives are on the horizon...
 - But... NOT attempting SMR: neither host nor drive-managed
 - Holding out for CMR (traditional) drive management
 - Working on an approach to mitigate long drive copy times
- So many other lessons of frugality: SSDs, Cooling, electrical, etc.

Internet Archive

Frontiers of Frugality

John C. Gonzalez

Director of Engineering and Accessibility

jcg@archive.org

IA Storage Principles (Reprise)

Transparency

Simplicity

Durability

Performance at Scale

Longevity of Access

Items = Directories on Disk

Disk = Unit of storage

Each disk is replicated

BOTH disks serve content

Evolve formats as needed

DENSITY OF STORAGE

36 Drives per data-node

For details, see this [blog post](#)

Detail on OLD Firmware issue...

- We only saw “hang” problems with firmware version AR17...
 - But earlier versions of firmware has other problems fixed in AR17
 - Hangs were frequent (but random) in Hadoop cluster... Rare in storage cluster
- Fix for hang problem for firmware version AR17:
 - All crashes we saw were when linux asked the disk to flush internally cached writes
 - → Disable write caching in the disk drive. That fixed it! (But this took quite a while to diagnose!)
- Population Details:
 - DATA: 'AR13'->2176 . 'AR15'->874 . 'AR17'->4114
 - WBGRP_HADOOP: 'AR13'->258 . 'AR17'->417 . 'RT17'->9
 - We now use RT17 ONLY in HADOOP applications