Tape Evolution
pre-GDB report

Alastair Dewhurst
Pre-GDB on Tape Evolution

- Previous meeting in February 2021.
- Well attended:
  - 20 – 30 in person
  - 40+ people on zoom
- Talk from industry
- ATLAS developments
- Updates from the storage developers.
- Site reports
Tape Fundamentals

▪ Not everyone is an expert in tape.

▪ Principles:
  ▪ Tape systems will always prioritize writing data to tape.
  ▪ The tape drives which read/write the data to tape are the bottleneck in the system and need to be used efficiently.
  ▪ Tape systems are designed for long term storage with infrequent access, we need to use them wisely.

▪ A lot of discussions are about recalling data
Spectra Logic – Future of Tape

- IBM is the one remaining company that develops new tape technology.
  - We need to watch developments carefully.

- IBM recently released (August 2023) the new TS1170 drives and media.
  - 50TB per tape (increase from 20TB in the previous generation)
  - 400MB/s read/write (same performance as previous generation)
  - Increased environmental specifications to make this work.
    - Max of 50% humidity down from 80%.

- Price per TB for Tape continues to remain well ahead of HDD.
  - E.g. for LTO-8 media <$5 / TB
Storage Technology Roadmap

- Tape has a much larger surface area that HDD.
  - LTO-9 tape is 1,035 meters long and ½ inch wide – 20,374 square inches
  - HDD is 3.5 inch in diameter with 10 platters – 96 square inches
- This difference allows for a much higher capacity with standard magnetic recording technologies using tape while disk has already hit the superparamagnetic limit with conventional technologies.

Figure 1: Areal Density Trends. Hard Disk Drive, Tape Product and Tape Technology Roadmap

INSIC 2019 Technology Roadmap
INSIC Technology Roadmap 2019 - SM

Alastair Dewhurst, 8th November 2023
ATLAS Developments

- ATLAS Data Carousel has been in production since 2021.
  - New work focuses on DAOD on demand.
    - Is it better to archive or re-create rarely used DAOD?
  - Smart writing
    - Trying to co-locating files on the same tape that will be recalled together.
    - Demonstrator at FZK showing factor of 2 performance improvement.
  - Archive Metadata.
DAOD-on-demand HL-LHC demonstrator (2/2)

- Tests done so far on both scenarios, using data sample from the recent ATLAS data deletion campaign, at two Tier-1s (FZK and RAL).
- Preliminary results
  - Comparison of TTC (Time To Completion) among different scenarios

<table>
<thead>
<tr>
<th>Data type</th>
<th># datasets</th>
<th># files</th>
<th>Size (GB)</th>
<th>Action</th>
<th>&lt;TTC&gt; per dataset (h)</th>
<th>Source (tape) site</th>
<th>Time stamp</th>
</tr>
</thead>
<tbody>
<tr>
<td>AOD</td>
<td>13</td>
<td>31627</td>
<td>107047</td>
<td>Staging</td>
<td>19 +/- 9</td>
<td>FZK/RAL</td>
<td>July-Sep 2023</td>
</tr>
<tr>
<td>DAOD</td>
<td>11</td>
<td>1555</td>
<td>7284</td>
<td>Staging</td>
<td>3 +/- 4</td>
<td>FZK/RAL</td>
<td>July-Sep 2023</td>
</tr>
<tr>
<td>DAOD</td>
<td>5</td>
<td>1158</td>
<td>5459</td>
<td>recreation</td>
<td>7 +/- 3</td>
<td>N/A</td>
<td>July-Sep 2023</td>
</tr>
</tbody>
</table>

- Bulk mode tests are ongoing, of which the results will be used to estimate both the TTC at scale and the extra load on the tape resources.
There is ongoing work to allow additional metadata to be sent to tape services allowing them to optimize their service.

Details will be presented by Julien at the Data Challenge 24 workshop.

VOs will be able to supply additional metadata regarding the datasets they are writing.

This will allow sites to improve tape utilization.
CTA + EOS

Recent developments:
- http REST API available for VOs since Q1 2023.
- EOS 5 now available.
  - CERN will upgrade LHC experiments now heavy ion run complete.
  - cback – backup orchestrator using CTA.

Future plans:
- gRPC
- Move to Alma 9
- Addition of Archive metadata
- Improving repacking and monitoring
- Schedule separation and migration to PostgreSQL
dCache

- dCache = disk cache in front of tape.
- dCache can be used with a variety of backends:
  - CTA, HPSS, TSM, Enstore, DMF etc
- Seamless integration with dCache is merged into upstream CTA code.
StoRM

- Extensive talk from StoRM developers.

- StoRM Tape basics
  - GEMSS component
  - Current data life cycle within a tape-enabled storage area

- StoRM Tape REST API
  - The WLCG Tape REST API specification
  - NGINX and OPA deployment roles
  - OPA authorization example
  - Testing tools
  - Ongoing developments

Alastair Dewhurst, 8th November 2023
Site Reports
CERN

Hardware Inventory

6 Libraries:
- 4x IBM TS4500
- 2x Spectra Logic
- TFinity

184 Drives:
- 1:1 drive to Tape Server mapping
- New: IBM TS1170

2023 Data Taking — Heavy Ion

New CERN Record

Data archived to tape storage each month since 2008

35 Petabytes in July 2023

15.9 Petabytes in November 2018

26 Petabytes in November 2022

Alastair Dewhurst, 8th November 2023
INFN – An audience with the Pope

What can the Tecnopolo host?

The computing infrastructures

B1 B2 B3 B4 B5 C2

Each of the 6 “botti” (barrels) is ~5000m² of usable IT space

Same architect and design of the "Sala Nervi” in the Vatican

Alastair Dewhurst, 8th November 2023
Metropolitan Tape Area Network

- 2 libraries at CNAF
- 1 new library at the Tecnopole
- About 7 km of fiber to connect the 2 datacenters
  - yellow + red paths
- 2 fiber pairs dedicated to extend the fiberchannel TAN
  - Brocade optics for 10km distance

**Brocade® 32Gb/s LWL (10 km) SFP+**

Optimized, Certified Optical Transceivers for Extending Service Provider and Data Center Networks

**Overview**

Brocade enterprise data centers are undergoing an infrastructure transformation, requiring higher speeds, greater scalability, and higher levels of performance and reliability to better meet the demands of business. As speed and performance needs increase, optical transceivers—once considered a generic component of Fibre Channel switching technology—now become an integral part of overall system design.

The Brocade 50GB/s Long Wave-length Unit (10 km SFP+ or SFP) helps drive cost-effective, high performance, and high reliability solutions in data centers and service provider environments.
CTA Frontend/CTA CLI

Tier-1 EOS cluster

Facilities EOS cluster

Tier-1 CTA Tape Servers

Facilities CTA Tape Servers

XRootD Client

TPC proxies

CTA Catalogue

CTA object store

Data

Metadata

TS1160 Tape Drives

LT08 Tape Drives

LT09 Tape Drives

TS1160 Tape Drives

LT09 Tape Drives

Facilities SpectraTfinity tape library

Tier-1 Spectra Logic Tfinity tape library

Alastair Dewhurst, 8th November 2023
FZK – Putting the HP in HPSS!

Recalling files from HPSS

Main goal: recall files efficiently from tapes for O(50k) requests
- Best for tapes: mount only once and read from front to end
- Best for experiments: obtain files at stable rates of O(1GB/s)
- Experiments recall large fractions of datasets during recall activities
  → Optimize based on these boundary conditions:
    - full aggregate recall (FAR) in HPSS
      → faster reading of files on a tape from the same aggregate
    - recommended access order (RAO) in HPSS
      → multiple aggregates are recalled in most efficient order from a tape
    - number of used drives per experiment configurable
      → remaining flexible w.r.t. the load on HPSS

Deployed in an adapted dCache ENDIT-Provider and dedicated ENDIT-HPSS interface
  → technical details to be published in CHEP 2023 proceedings

Alastair Dewhurst, 8th November 2023
Data Center Migration

- Tape operations split between data centers
- Bldg 515 - Original “legacy” data center
  - Hosts data primarily from before run 3
  - 3 ATLAS Oracle SL-8500 libraries
  - ~11K LTO-7, 6K LTO-6 tapes with ATLAS data
- Bldg 725 - New, energy efficient and highly available data center
  - Hosts data from Run 3
  - HPSS core server
  - ATLAS HPSS disk cache
  - ATLAS IBM TS-4500 libraries
  - LTO-8 tapes containing new data
  - ATLAS LTO-8 tape drives

* High Performance Storage System - [hpss-collaboration.org](http://hpss-collaboration.org)
Tape Summary (2021)

Alastair Dewhurst, 10th March 2021
Tape Summary (2023)

Alastair Dewhurst, 10th March 2021
Questions?