dCache tape pools

Tape technology and other stuff for the next 5 years

Tape technology (current)

- All numbers native/uncompressed
- LTO8/Ultrium (Nov 2017)
  - 12 TB per tape
  - 360 MB/s
- TS1160/3592/Jaguar (Dec 2018)
  - 20 TB per JE tape, 15 TB on JD tapes (ie. TS1155 generation)
  - 400 MB/s

Tape technology - media roadmap

- Tape tech (Fujifilm)
  - 50-60TB 2022-2023 timeframe
  - 100+TB using Strontium Ferrite tapes in the 2030s

Tape technology - LTO roadmap

- LTO/Ultium
  - LTO 9
    - Up to 24 TB per tape
    - Up to 700 MB/s
    - 2020 timeframe?
  - LTO 10
    - Up to 48TB
    - Up to 1100 MB/s
    - 2023 timeframe? (likely after TS1170)

Tape technology - 3592/Jaguar roadmap

- 3592/Jaguar
    - TS1165?
        - Maybe? TS1155 was not on any roadmap.
        - Would make sense for IBM to keep 3592 ahead of LTO.
    - TS1170
        - Up to 50 TB (JF tapes), 30 TB (JE tapes)
        - Up to 1000 MB/s
        - 2022 timeframe? (likely after LTO 9).

What are our requirements?

- Is bandwidth the limiting factor?
  - Read and Write tape pools have different characteristics
  - Tape technology
  - Experiment requirements/behavior
- How much buffer do we need?
  - If TSM server or library dies on friday evening, still be able to take data until monday morning
- Hardware purchased must perform during entire production period
  - Must have headroom to cope until it's replaced
- Hardware lifetime?
  - Not more than 5 years, but if we're unsure if what we're purchasing today will be good enough we might need to make funding people aware.

Will sites upgrade tape tech at all?

- If pledge/capacity needs increase, probably.
- If pledge/bandwidth needs says so, probably.
- You need to communicate with your tape admins when looking into the future!

LHC roadmap likely affects pledges/requirements

- Long stop 2 2019+2020 (24 months + 3 months beam commisioning)
- Run 3 2021 Q1 - 2023 Q4
- Long stop 3 2023 Q4 - 2026 Q3 (30 months + 3 months BC)
- Run 4 High Luminosity - 2026 Q4 - 2030 Q4

HDD vs SSD

- Cost issue
  - If you've got the funding, go with SSDs!
- Short-term (our current discussion)
  - HDDs cheapest option if we need the size
    - 150 MB/s reasonable bw estimate for single LFF 4T HDD bandwidth
- Long-term (high luminosity and beyond)
  - Hoping that dCache evolves so we can take data to disk pools and migrate from there to tape pools
  - Allows smaller/faster storage in tape pools (ie. SSDs)
  - Multiple 1 GB/s tape drives will be hard to do with HDDs
  - We will likely be expected to deliver the capacity and performance provided by future tape technology generations...

Requirements in numbers

- Disk size
  - MoU
    - Service degradation less than 20% -> fix next business day
    - If all sites perform good, losing one isn't critical
    - Aiming for at least 40 TB is likely prudent if purchasing HDD
  - Minimum size determined by tape bandwidth, being able to buffer data to get 30 minutes of tape activity reduces number of tape mounts
    - 1 GB/s for 30 minutes is approx 2TB
      - At least double that to take data while writing to tape
      - Say 5 TB as an absolute minimum?
- Disk bandwidth
  - 10GigE to LHC OPN => 1.25 GB/s
    - Limiting the tape pools to 10GigE is probably wise, even on machines with 25GigE
  - Catering for continuously migrating to tape, doubling gives us 2.5 GB/s

Discussion!