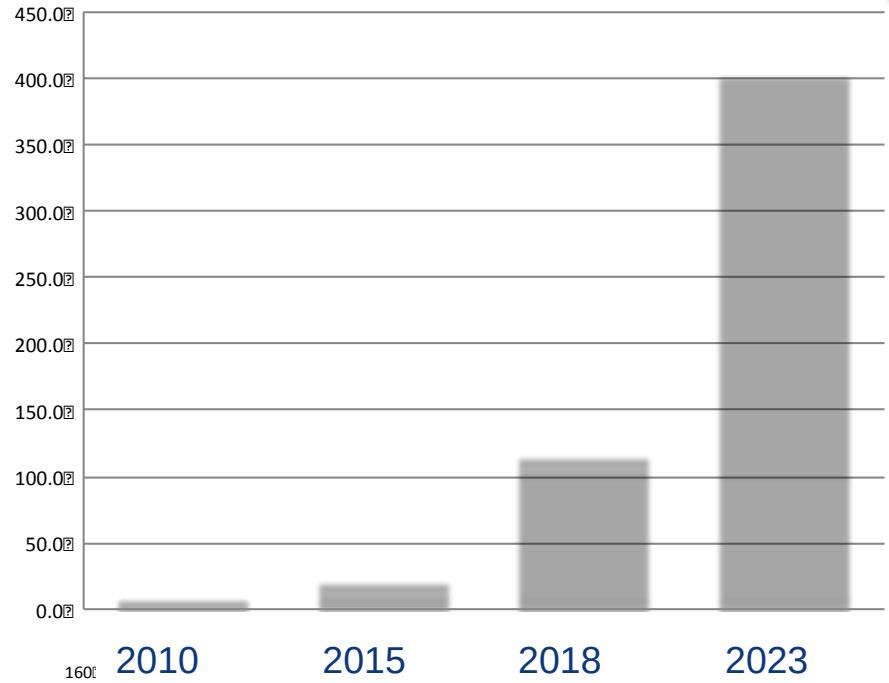
Ian Bird

Trigger, Online, Offline Computing Workshop

CERN, 5th September 2014

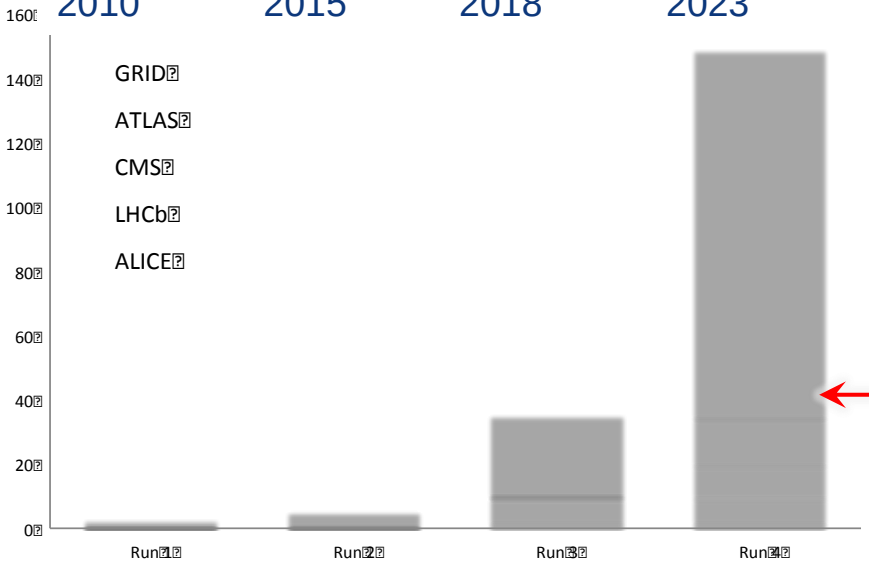
Resource Provisioning - Outlook

Scale of challenge …



Data: ~25 PB/yr → 400 PB/yr

10 Year Horizon



Compute: Growth > x50

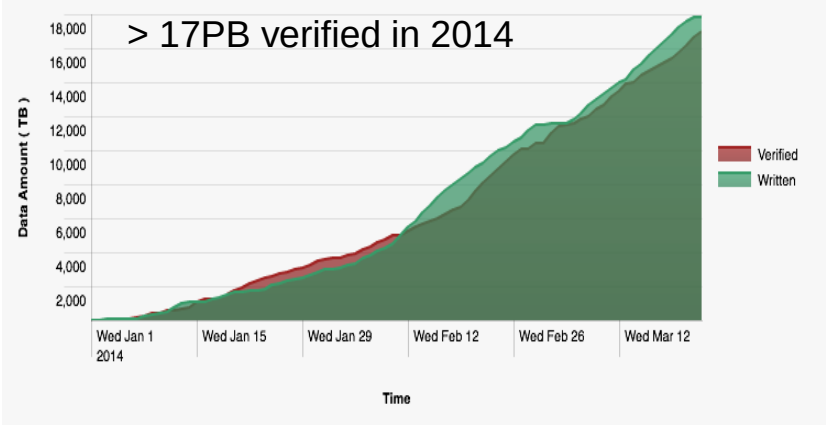
What we think is affordable unless we do something differently

Data – scale and challenges

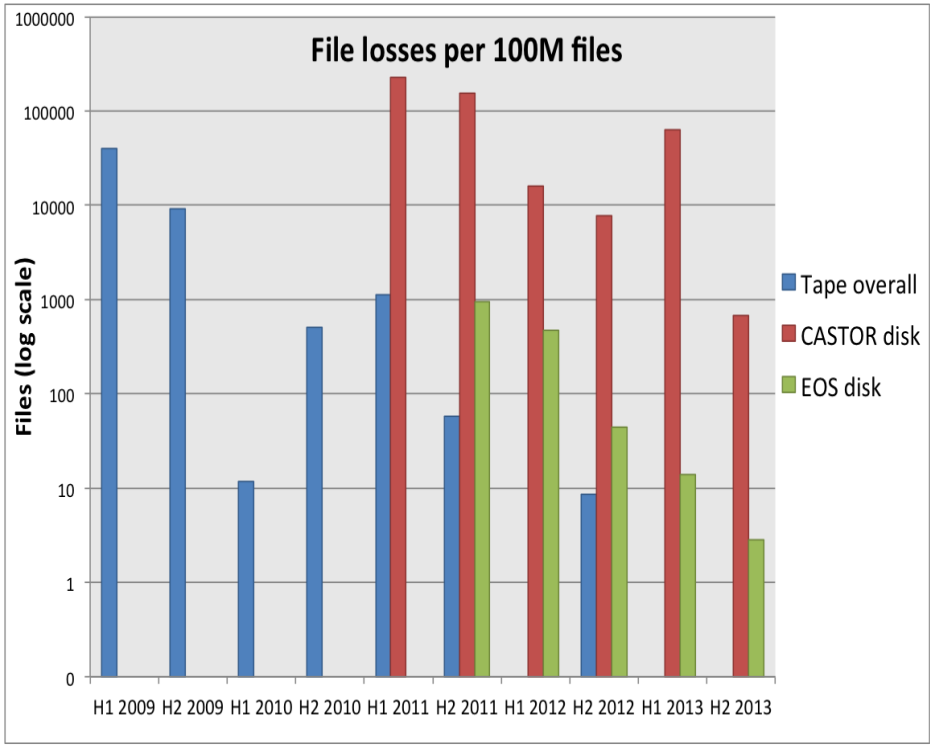
Following slides on tape from German Cancio (IT-DSS)

Reliability and "bit" preservation

- Systematic verification of archive data ongoing
  - "Cold" archive: Users only accessed ~20% of the data (2013)
  - All "historic" data verified between 2010-2013
  - All new and repacked data being verified as well



- Data reliability significantly improved over last 5 years
  - From annual bit loss rates of $O(10^{-12})$ (2009) to $O(10^{-16})$ (2012)
  - New drive generations + less strain (HSM mounts, TM "hitchback") + verification

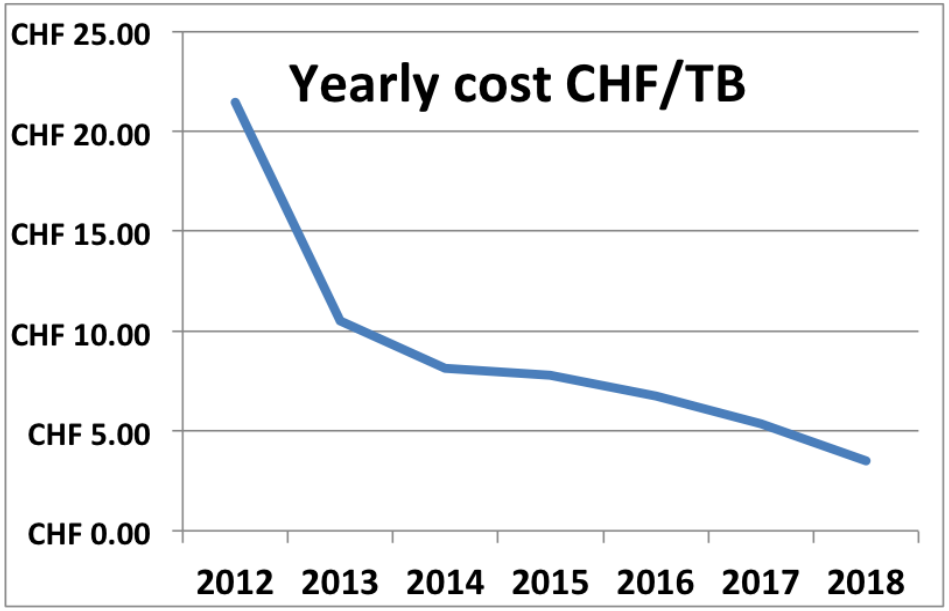


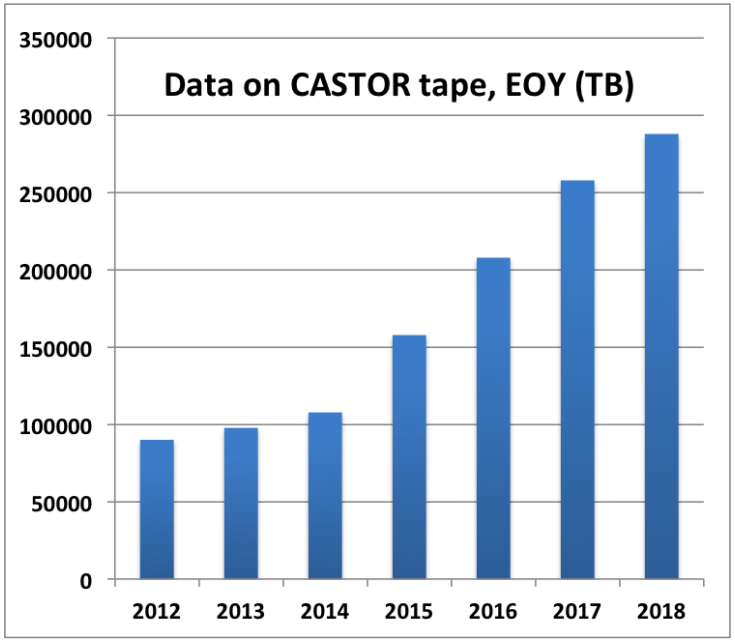
- Still, room for improvement
  - Vendor quoted bit error rates: $O(10^{-19})$
  - But, these only refer to media failures
  - Errors (eg bit flips) appearing in complete chain

an

Physics tape: cost outlook

- Capacity/cost planning kept for ~4y time window (currently, up to LS2 start in 2018)
  - Strategy: Dual-sourced enterprise media/drives; no LTO as not competitive
- Forecast
  - Assuming +50PB/year in 2015-17 (+30PB in 2018)
  - Includes HW, maintenance, media
  - Cost/year *usable* TB: 8.2CHF(2014).. 5.4CHF(2017)

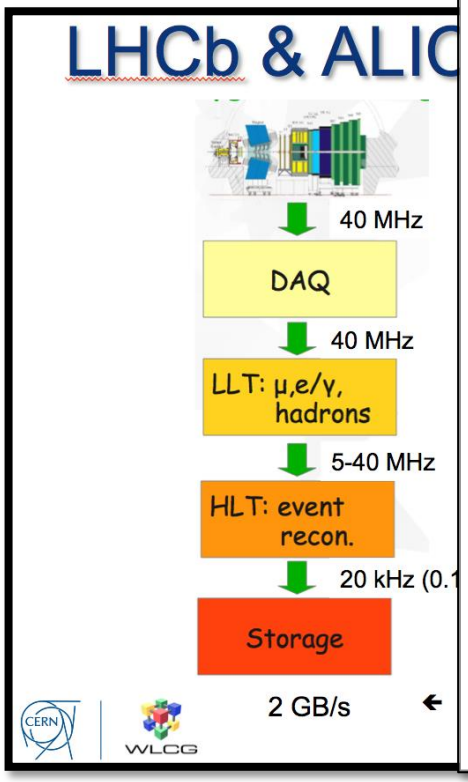


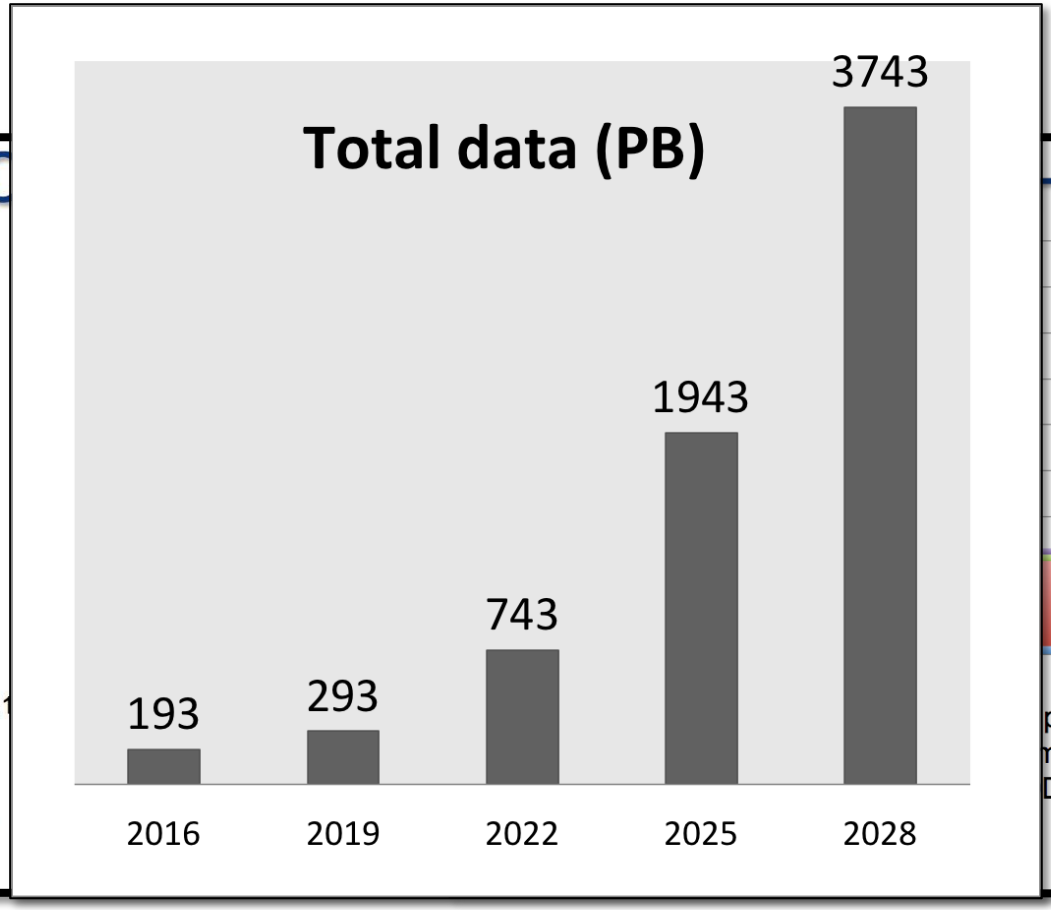


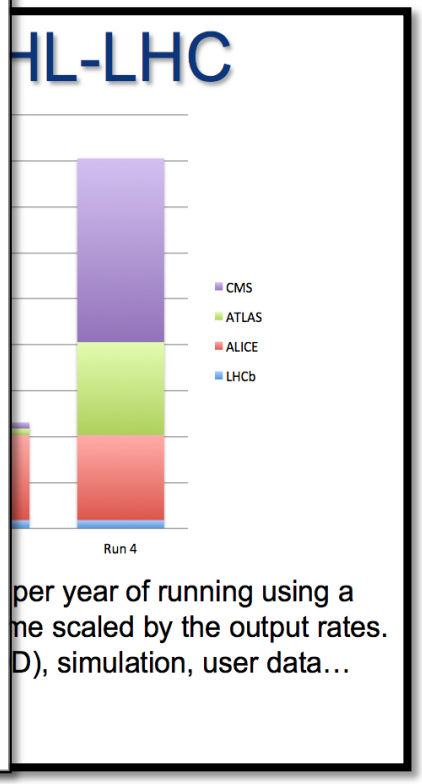
an

Longer term?

- Beyond 2018?
  - Run 3 (2020-2022): ~150PB/year
  - Run 4 (2023-2029): ~600PB/year
  - Peak rates of ~80GB/s



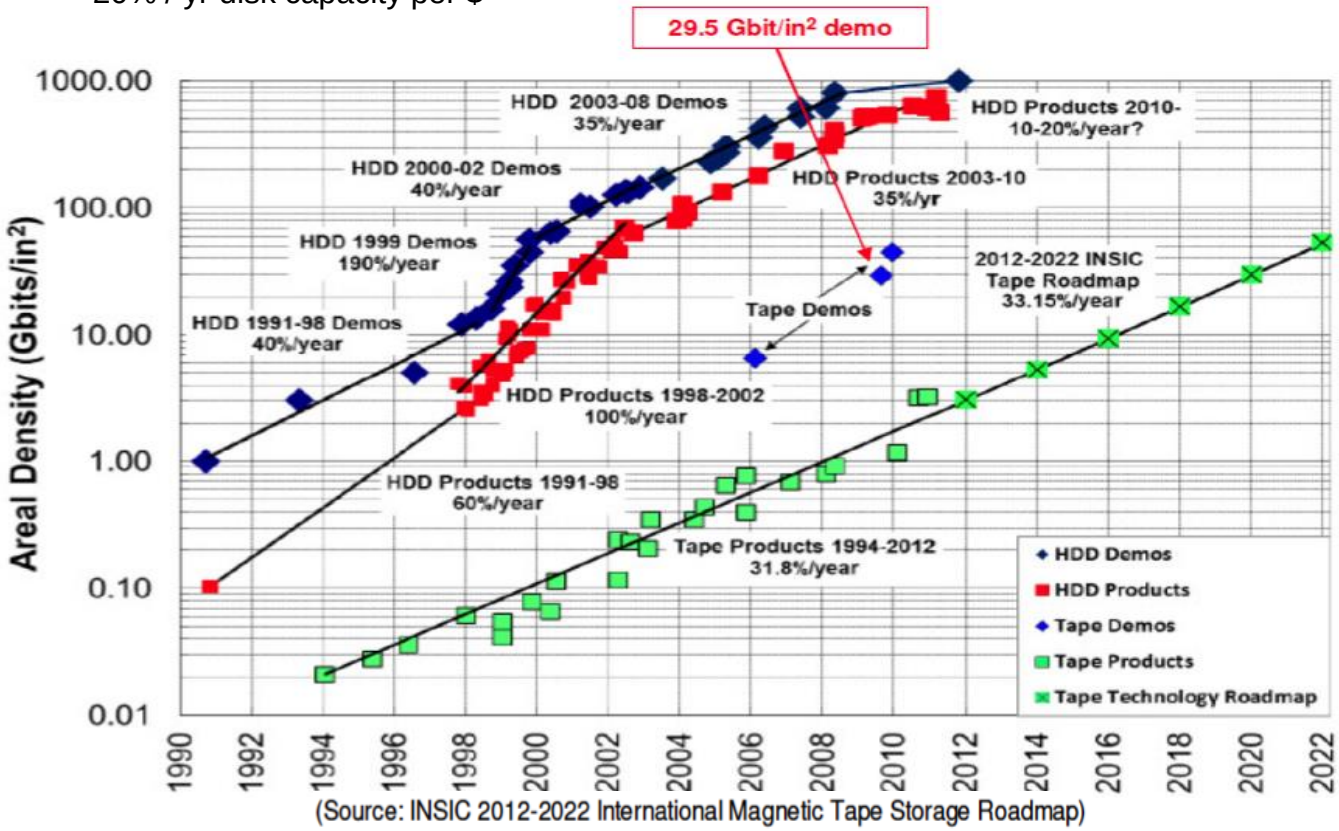




an

Longer term?

- Beyond 2018?
  - Run 3 (2020-2022): ~150PB/year
  - Run 4 (2023-2029): ~600PB/year
  - Peak rates of ~80GB/s
- Technology/market forecast (…risky for 15 years!)
  - INSIC Roadmap:
    - +30% / yr tape capacity per $ (+20%/yr I/O increase)
    - +20% / yr disk capacity per $

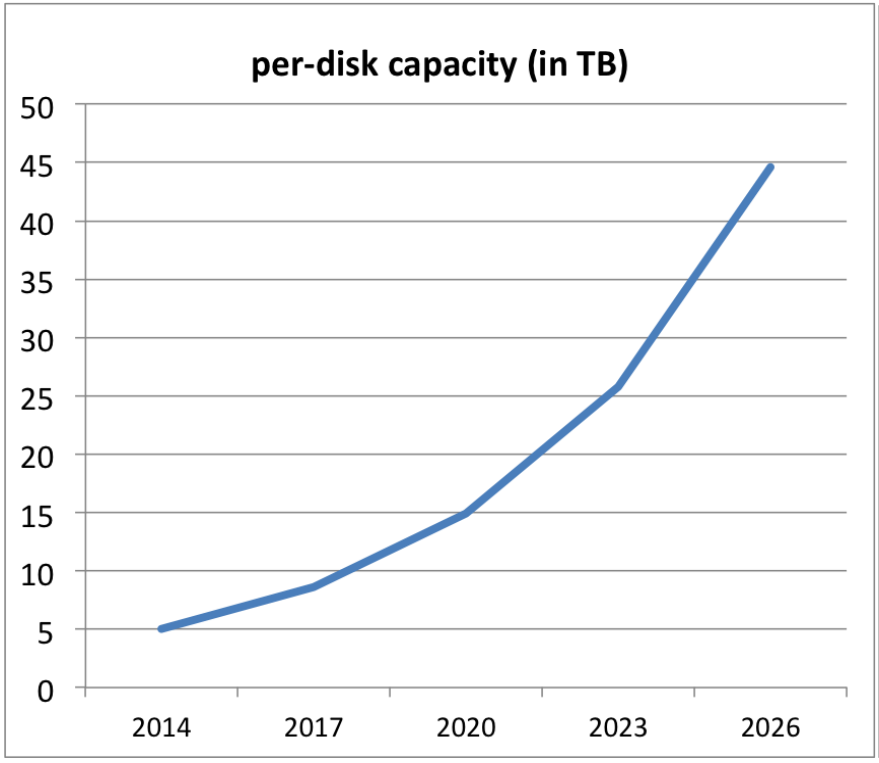


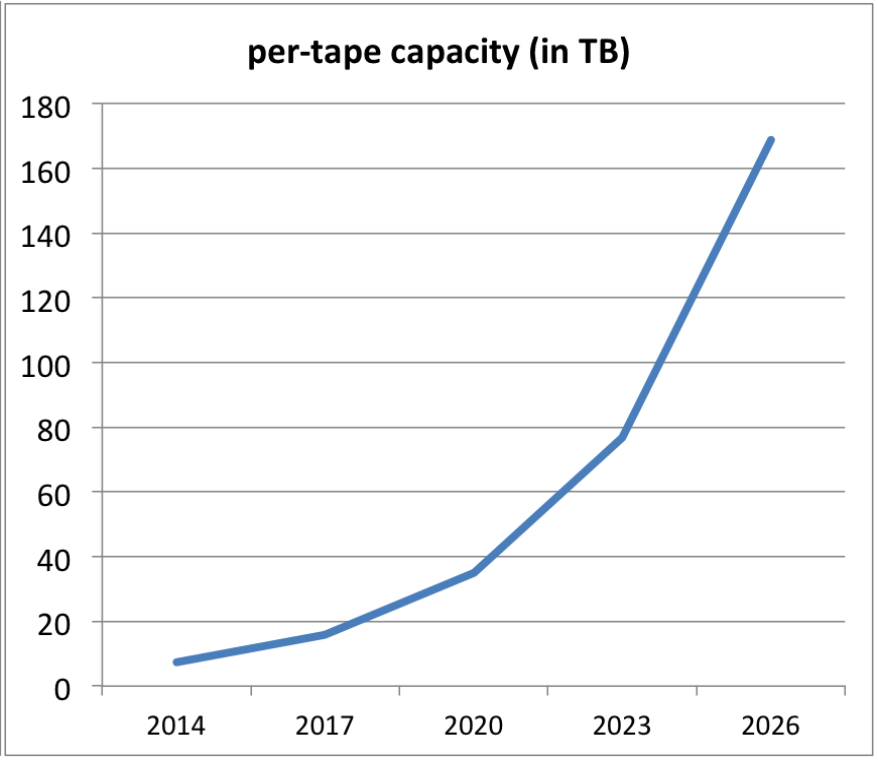
(Source: INSIC 2012-2022 International Magnetic Tape Storage Roadmap)

an

Longer term?

- Beyond 2018?
  - Run 3 (2020-2022): ~150PB/year
  - Run 4 (2023-2029): ~600PB/year
  - Peak rates of ~80GB/s
- Technology/market forecast (…risky for 15 years!)
  - INSIC Roadmap:
    - +30% / yr tape capacity per $ (+20%/yr I/O increase)
    - +20% / yr disk capacity per $

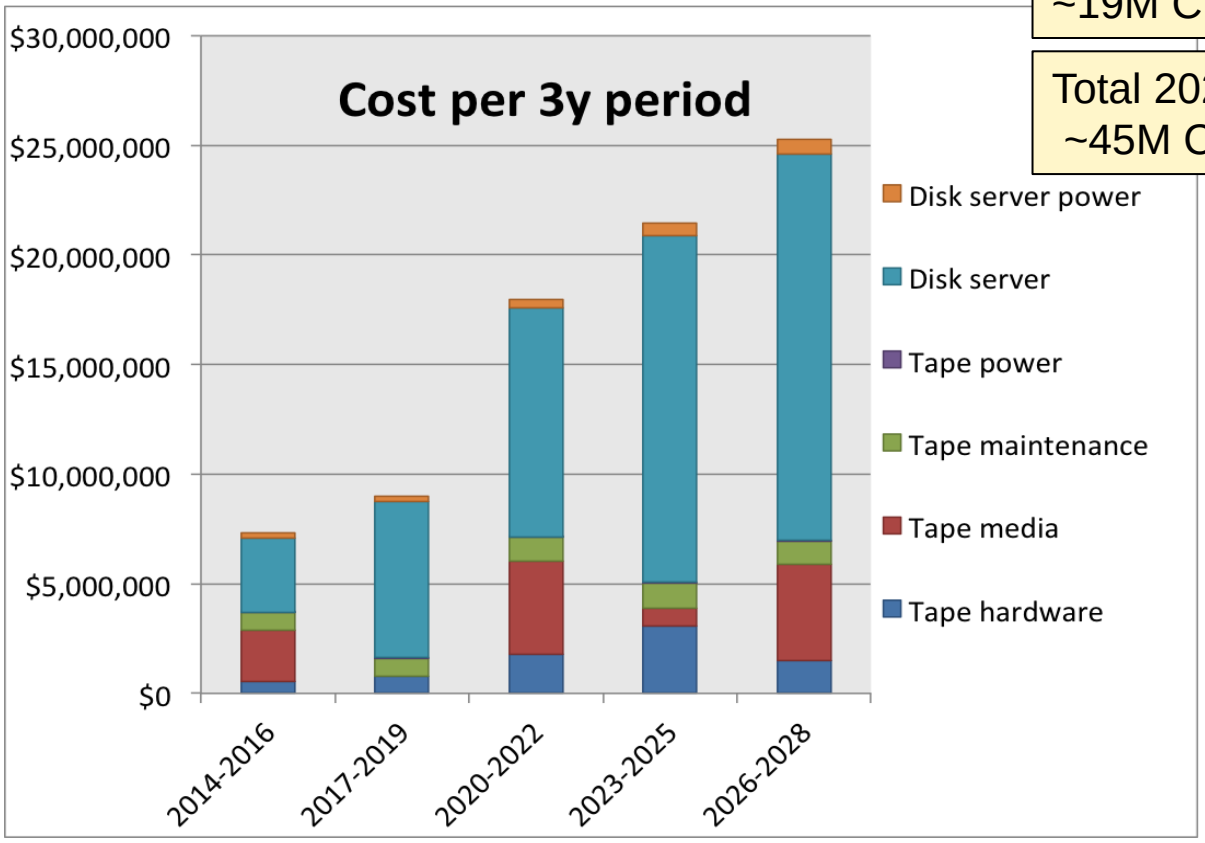




an

Longer term?

- Cost prediction - with many assumptions:
  - No paradigm change…!
  - 10% disk cache (with 20% redundancy overhead)
  - 3y cycle for disks and tape drives, and 6 years for reusable enterprise tape media (repack every 3y)
  - Tape libraries upgraded/replaced around 2020-2025
  - No inflation

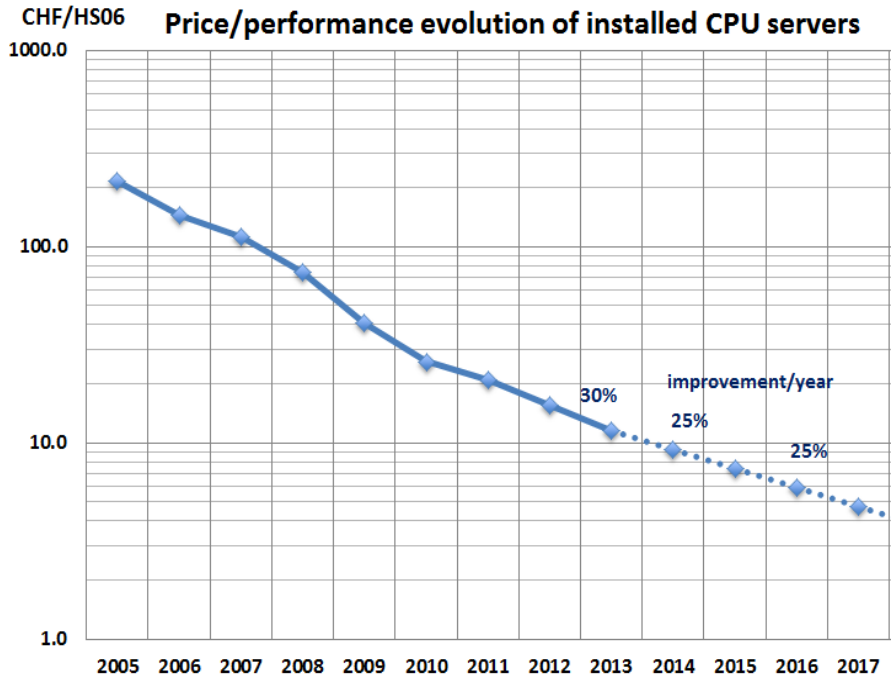


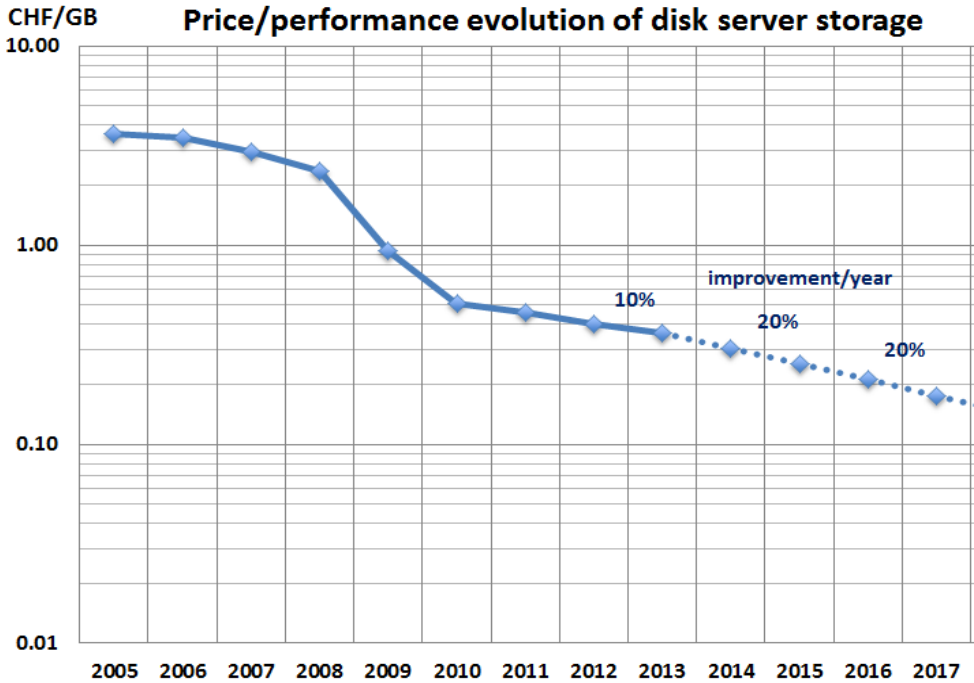
Total 2020-2028 tape:
~19M CHF (2.1M CHF / year)

Total 2020-2028 10% disk:
~45M CHF (5M CHF / year)

an

Technology outlook



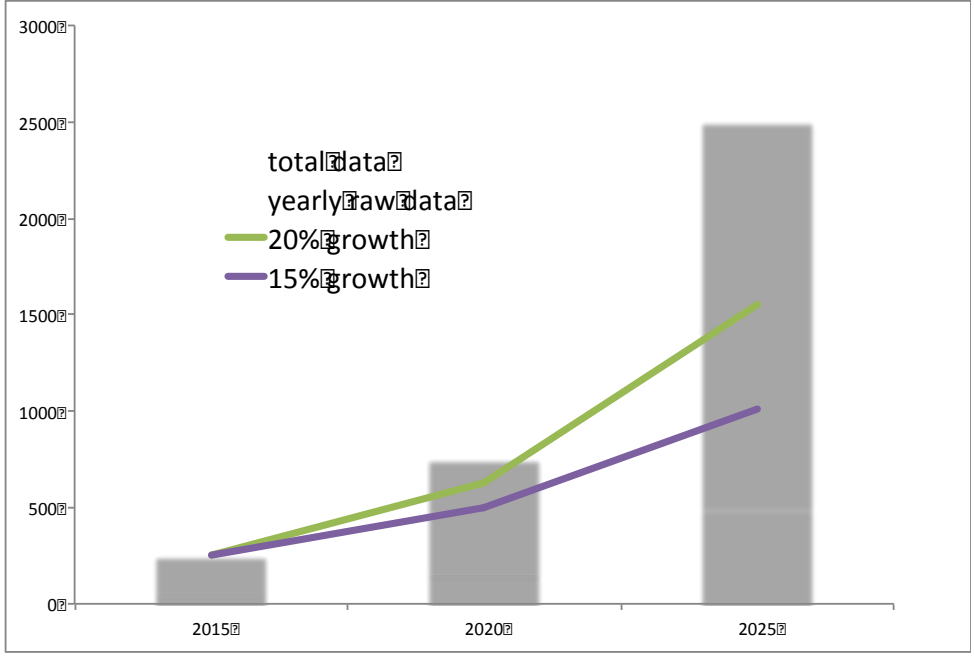


- *Effective* yearly growth: CPU 20%, Disk 15%, Tape 15%
- Assumes:
  - 75% budget additional capacity, 25% replacement
  - Other factors: infrastructure, network & increasing power costs

Storage costs

- Archives:
  - Expect that tape costs will be reasonable
  - Disk costs will increase
  - Overall – ~ doubling of cost for archiving and accessing the archived data

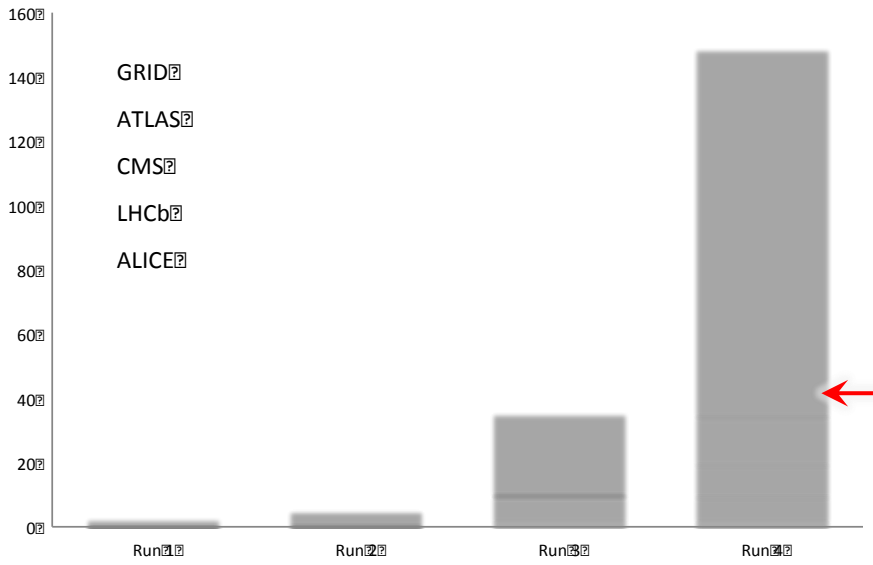
Active data - disk



- Assumes ratio of disk to yearly raw data is as currently requested for 2015
- Assumes flat budget annual growth remains at 15-20%
- In 2025 cost is >x2 too high

- Problem is compounded by technology:
  - Steady decrease in costs (/2 every 18 months) is faltering
  - New technologies “close”, but not clear how easy to use, or if they need to be used differently (more like tape)
  - Real growth is likely to slow – our 15% assumption may be too high!

CPU

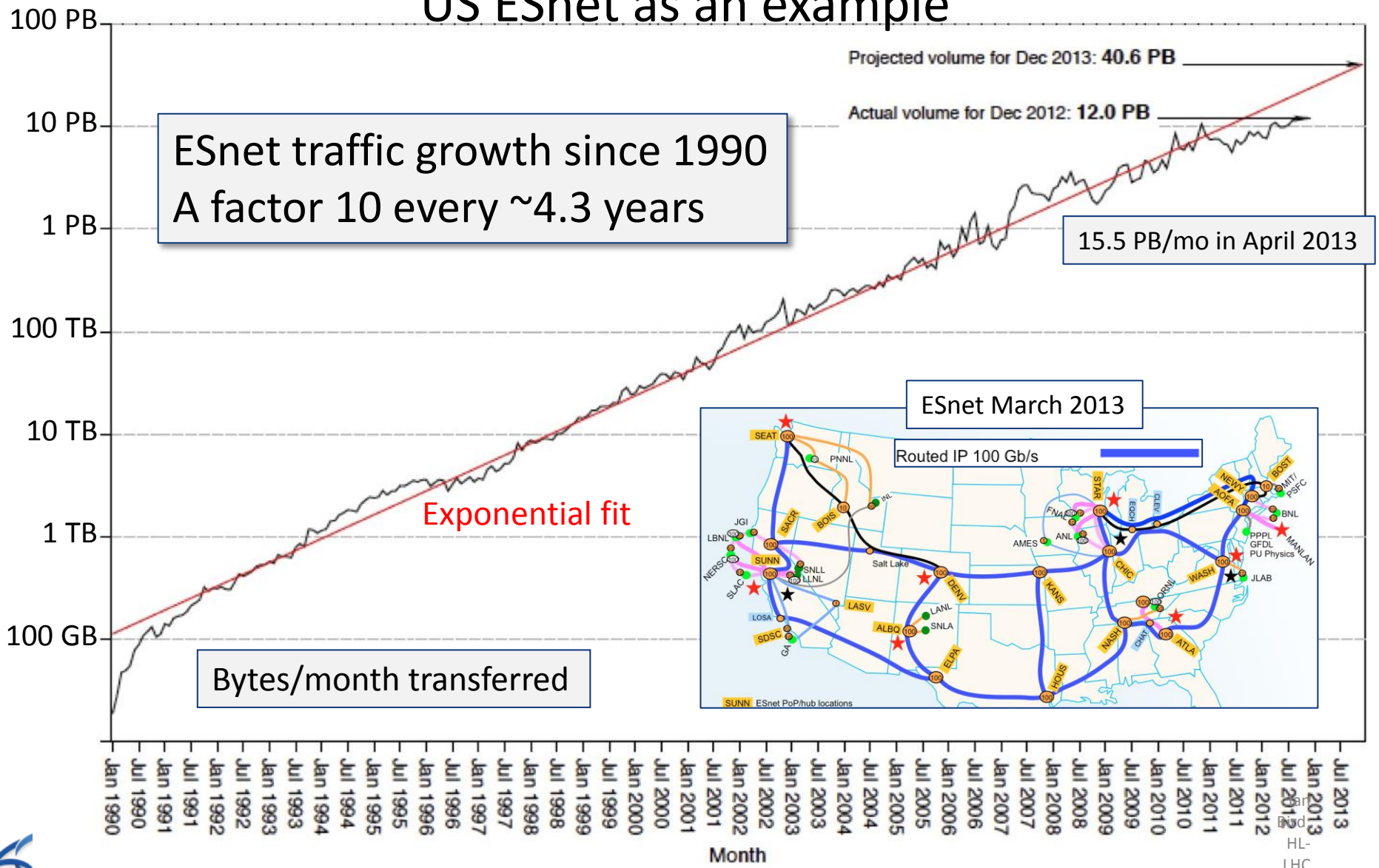


What we think is affordable unless we do something differently

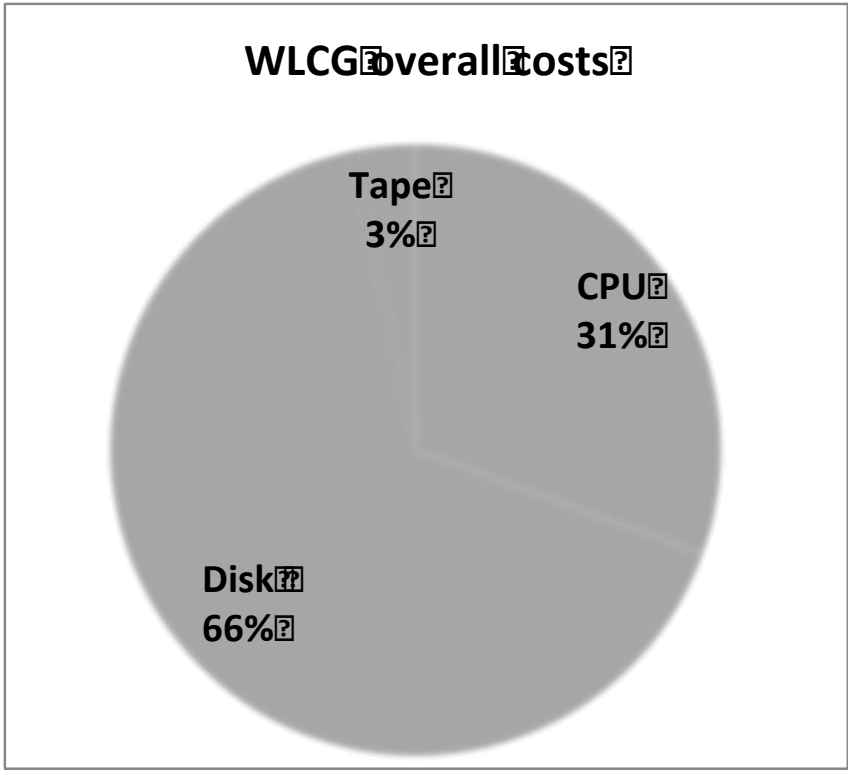
- CPU cost – assuming no change of model from today is x3 too expensive

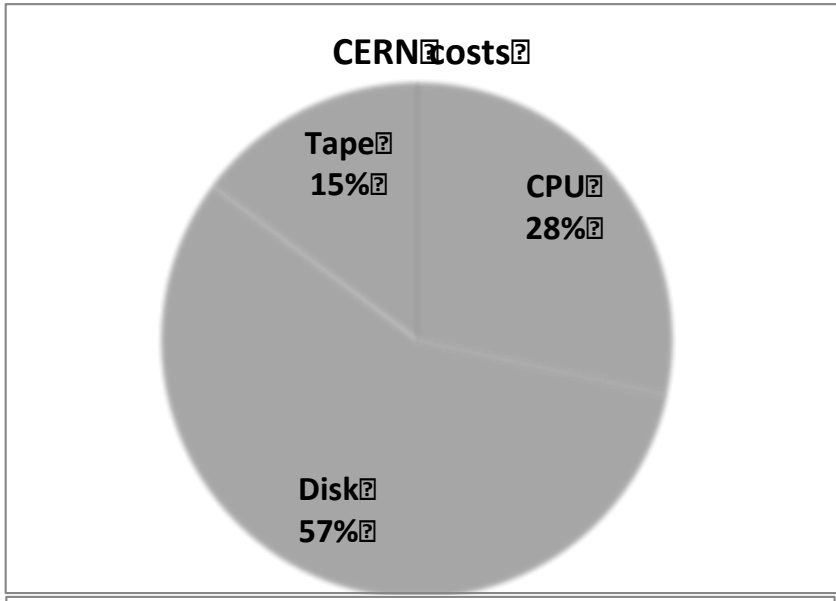
Networking growth has been dramatic

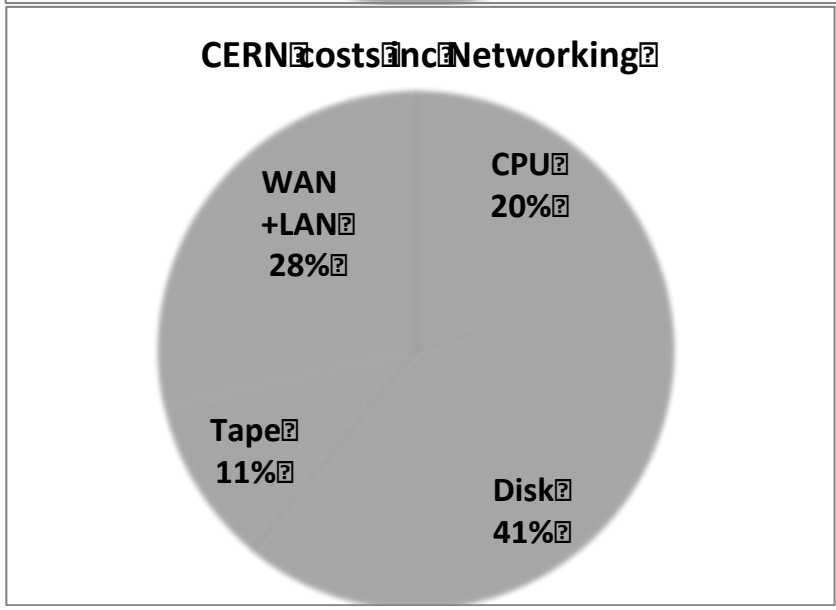
US ESnet as an example



Costs







Main cost driver is active storage – disk

Resource growth - summary

Assuming similar computing models as today:

- Networks
  - Technology growth will provide what we need;
    - Cost ? Affordable if today's trends continue
- Archive storage
  - Tape (robotics, drives, media) – cost similar to today for full anticipated HL-LHC data growth
  - Disk buffer cost will be much higher
- Active storage (data copies, caches, etc)
  - Costs factor 2-3 higher than flat budgets
- CPU
  - Costs factor 3-5 higher than flat budgets

➢ **Biggest impact on overall costs is disk storage**

Technical measures

1. Need to reduce disk usage significantly
2. Need to be more effective in use of CPU
   - But not so much that makes point 1 worse
3. Need to be able to use resources:

• Pledged • Opportunistic • Commercial	• Grid • Cloud • HPC • ???	Some of this only helps CPU, storage is not opportunistic (easily)

➢ Simple interfaces are needed

- 0 config; 0 installed software …

Funding related…

- Optimistic view is flat-cash budgets for pledged resources
- Must supplement with new resources
  - New countries, new collaborators, other opportunities (partnerships, …)
  - Unlikely to find factors more resources …
- Are there better ways to use available funding?
  - What are the FA's prepared to do?
  - These are driven by national issues – environment, scale, decisions, …

Open questions

- Why do we need to maintain (very) distributed computing
  - How far can some consolidation happen? Benefit with economies of scale.
  - Nationally? Internationally?
- Why do physicists need data locally? (this is a problem that is different for well connected and remote countries)
  - Better to process at or very near the storage
  - Reduce the number of copies of data to a minimum
- Can we imagine a model of (national) large centres forming a logical data processing hub which data does not leave? Outputs would be physics data sets.
- Access datasets stored in a “physics cloud”
- Could reduce from 200 centres to O(40) countries ?

Funding agencies

- Expect to see (NOW!) common efforts towards future economies
  - They will not accept 4x software stacks for much longer
  - All 4 experiments are losing computing effort
- Also expect HEP to collaborate with other large data sciences (at least) to ensure their investments are re-used
  - Astronomy, astro-particle, cosmology, photon science, life-sciences, etc, etc;
- Not only do we have to express our own commonalities – we are expected to seek them with other disciplines
- If we don't there is a real risk of further reduced funding

Summary

- Fulfilling LHC resource needs for 2025 is a real challenge
- Existing computing models will not scale
- Existing insular attitudes will not be acceptable
- We need to aggressively address these issues
  - And show that we are doing so