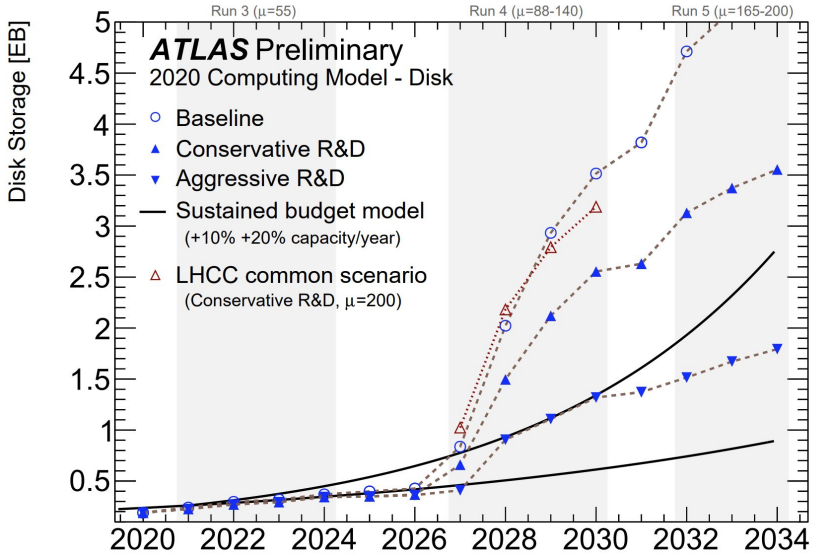
Overview of Storage Topics

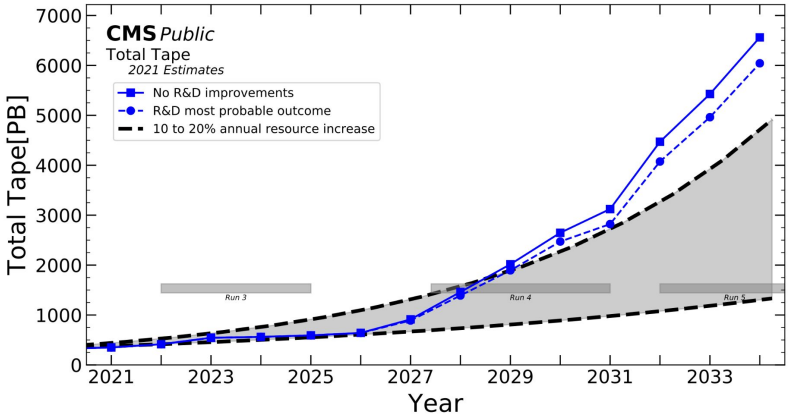
Bo Jayatilaka (FNAL), Carlos Maltzahn (UCSC), Peter Van Gemmeren (ANL)

CompF04 Topical Group Workshop
April 7, 2022

Future Challenge: Increased need for storage, e.g. HL-LHC

Figures: Top, ATLAS Estimated disk (at the Tier-1 and Tier-0) resources needed for the years 2020 to 2034. Bottom, CMS Estimated tape storage for the years 2021 to 2034





Future Challenge: Not ‘just’ storage

Processing for future HEP experiments, such as HL-LHC, faces similar challenges for CPU cycles:

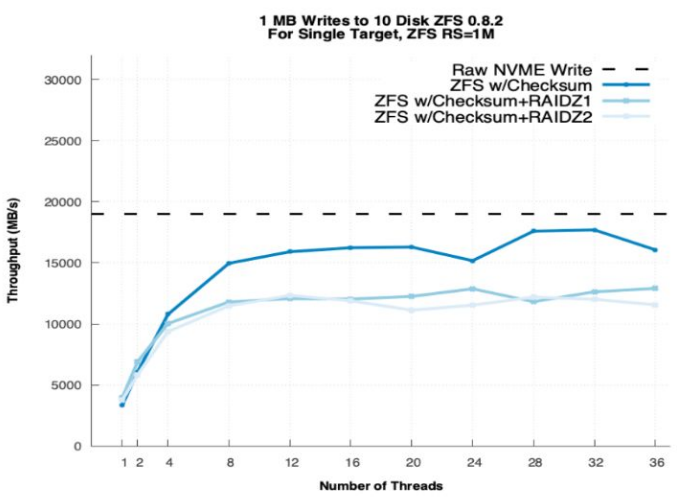
- Porting Workflows to High Performance Computer (HPC)
  - Requires storage and input/output on customized HPC storage systems and data offloading to compute accelerators

HL-LHC example: 10X data, 10X complexity

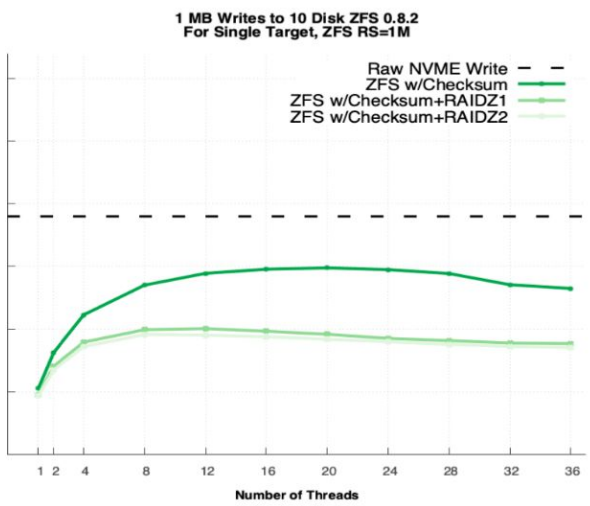
In 2030, LHC experiments will need:

- > O(100) PetaFlops in sustained compute performance
- O(10) Exabyte/year data throughput
- >O(1) Exabyte disk and tape storage

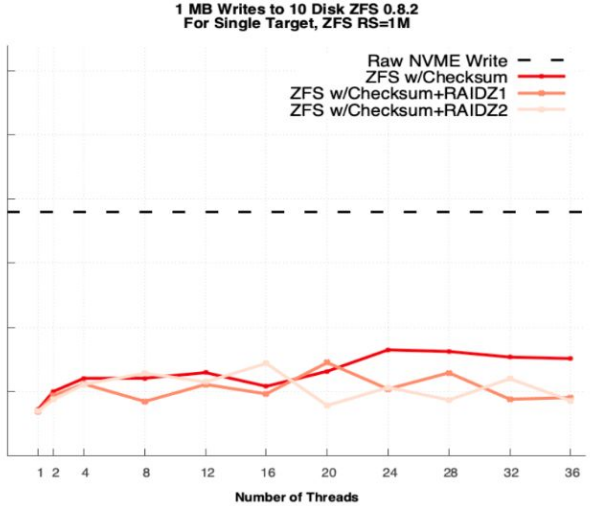
Future Challenge: Not ‘just’ storage



Intel Platinum (Dual Socket)



AMD EPYC (2nd Gen)



AMD EPYC (1st Gen)

Brad Settlemyer (LANL), “Accelerating File System and Data Services with Computational Storage,” Storage Developer Conferece (SDC’21), 9/28-29, 2021.

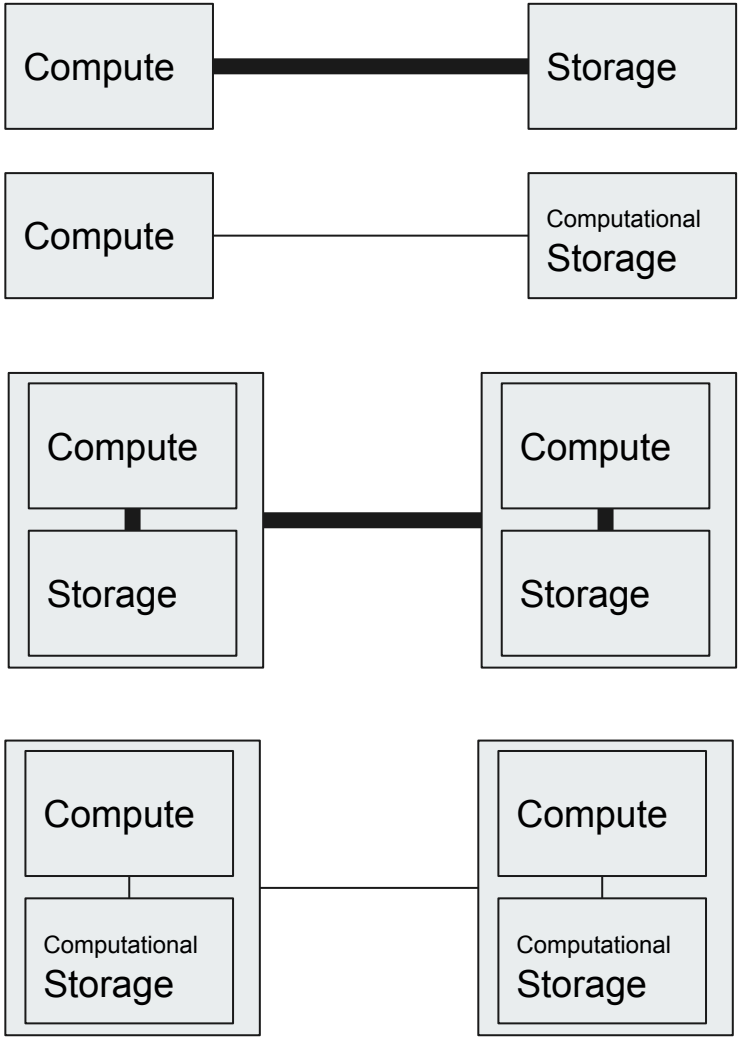
Storage Systems & Hardware?

Trends

- Hardware accelerators are becoming available within NICs, storage devices, PCIe devices, storage arrays
- Driven by disaggregation and faster storage devices (NVMs, NVMe Flash)
- Everyone wants common data management and access ecosystems, e.g. Apache Arrow and HDF5
- Standardization efforts (NVMe 2.0) and Open source ecosystems reveal internal strategies of hyperscalers:
- Leverage embedded processing and accelerators to reduce data movement while reducing latency

Storage Systems & Hardware?

Aligning *some of* data management and its context with data placement



Related work

HSF Community White Paper:

A Roadmap for HEP Software and Computing R&D for the 2020s, The HEP Software Foundation, Albrecht, J. et al., Comput Softw Big Sci (2019) 3, 7 https://doi.org/10.1007/s41781-018-0018-8.

Identified some DOMA/Storage R&D topics for the 2020s

- Sub-file granularity storage access and management
- Data organization and analysis used by other big data users (then Spark, now “Lake-housing”)
- Data placement and caching, including for ML applications
- Minimize infrastructure cost by coordinating tiered storage
- Globally minimize data access latency

Related work elsewhere

January 2022 DOE ASCR Workshop on the Management and Storage of Scientific Data

Workflows, database technologies, metadata and provenance

Management and Storage of Scientific Data:

- Interfaces for accessing data that resides on traditional persistent storage as well as memory devices;
- Storage-system architecture design that supports scientific workflows on varied hierarchical storage and networking devices;
- Devising metadata management infrastructure to support FAIR principles (Findability, Accessibility, Interoperability, and Reusability);
- Capturing provenance information about scientific data;
- Utilizing AI to learn I/O patterns of emerging workloads for efficient data management;
- Providing data management support for AI and complex workflows; and
- Understanding the overlap between traditional storage systems and I/O (SSIO) efforts and data management.

Storage Agenda for this workshop

Agenda: Thursday

We received one white paper submission:

1:05 PM **Data Storage for HEP Experiments in the Era of High-Performance Computing**

https://arxiv.org/abs/2203.07885

From: Amit Bashyal, Peter Van Gemmeren, Saba Sehrish, Kyle Knoepfel, Suren Byna, Qiao Kang

On Behalf of the HEP-CCE IOS Group (https://www.anl.gov/hep-cce)

Which will be presented by Amit Bashyal (ANL) today.

Note: Times are PST

Agenda: Friday

We have invited presentations form:

9:00 AM **Upcoming Storage Features in ROOT**

Presented by Philippe Canal (FNAL)

9:15 AM **Computational Storage**

Presented by Vitorio Cargnini (Samsung)

9:30 AM **Distributed data management at JGI**

Presented by Kjiersten Fagnan (LBL)

9:45 AM **Data Context and Sharded Data - Erasure Coding for HDD Computational Storage**

Presented by Philip Kufeldt (Seagate)

Note: Times are PST