HEP HPC Integration

CERN Pilot program

David Southwick
Viktor Khristenko
Maria Girone

Dissemination level: Public

Disclosing Party: CERN

Recipient Party: EGI-ACE HPC integration workshop

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.

Background & Motivation

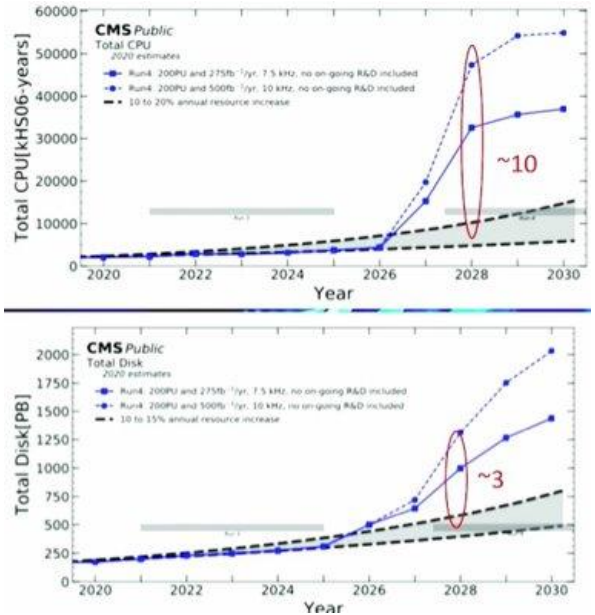
Big Data challenges in High Energy Physics

HL-LHC will produce more *computationally complex* physics events, with a *larger size* per event, and at *higher frequency:*

- Approaching the limits of what can be squeezed out of traditional CPUs for High Energy Physics (HEP) workloads
- Foreseen gap in computing resources far exceeds procurement feasibility by an order of magnitude

All avenues are being explored to enable processing and storage of full dataset:

- Aggressive event compression
- Aggressive event filtering
- Aggressive infrastructure expansion - **including HPC** compute sites (and **heterogeneous accelerators**)



Computing resource gap, CMS experiment projections

Challenges

Approaching HPC as High Throughput Computing

Common challenges for HPC integration are driving the technical program of demonstrators:

- HPC Benchmarking
- Data Access demonstrator

Demonstrate I/O can effectively use local storage and cached data delivery at the scale of HPC

- Working with WLCG DOMA and Datalake prototypes
- Working with AAI WLCG working group

Pilot Concepts

CERN use case

Utilize heterogeneous compute resources & accelerators

- All experiments currently working to exploit accelerators (GPU/FPGA) and arches
- Environments need to be packaged & mobile for shared computing

Benchmarking heterogeneous resources

- Understanding and accounting compute accelerators and other architectures

Data processing & access

- Enormous data volumes to stage, process, export from HPC sites
- Implicit authorization and authentication challenges
- Provisioning services for data management – both for dedicated storage sites with Data lake models and compute storage on HPC sites

Exploit synergies with other sciences!

HEP Benchmark Suite

A short history

HEP Benchmarking Suite: A benchmark orchestrator & reporting tool.
Provides an array of benchmarks, including HEPscore – the proposed solution for diverging HEPspec06 scores (over 15+ years use, EOL now)

- Designed for WLCG homogeneous compute environment
- Intended for procurement teams, site administrators
- First with VM containment, later nested docker images

None of these approaches are compatible with HPC:

- Collaboration with HEPiX Benchmarking Group to refactor & re-tool for **HPC** execution at scale!
- Enables R&D benchmarking; comparison across heterogenous architectures
- Easily extendable to other areas of science!

gitlab.cern.ch/hep-benchmarks/hep-benchmark-suite

HEP Score

Experiment workload benchmark orchestrator

- Modular python3 “microservice” approach
- Importable, Extendable, and **architecture agnostic**
- Executes set of containerized workloads (Singularity, Docker, Podman)
- Workloads decided by experiment experts & WLCG teams
- Detailed report delivered in JSON/YAML via AMQ/Elastic Search
- **Simple to extend to other sciences**

4 large LHC experiments represented

Experiment	Name	Description	Experiment license	Readiness	Pipeline status
Alice	gen-sim	link	GNU GPL v3	w.i.p.	pipeline passed
Atlas	gen	link	Apache v2	Y	pipeline passed
Atlas	sim	link	Apache v2	Y	pipeline passed
Atlas	digi-reco	link	Apache v2	w.i.p.	pipeline passed
CMS	gen-sim	link	Apache v2	Y	pipeline passed
CMS	digi	link	Apache v2	Y	pipeline passed
CMS	reco	link	Apache v2	Y	pipeline passed
LHCb	gen-sim	link	GNU GPL v3	Y	pipeline passed
Belle2	gen-sim-reco	link	GNU GPL v3	Y	pipeline passed

https://gitlab.cern.ch/hep-benchmarks/hep-workloads

Benchmarking on HPC

Production workloads with HEP Benchmark Suite

Workload containers packaged as OCI-compatible docker/singularity images

- Multi-arch container workloads (x86_64, IBM Power, ARM, ...)
- Multi-GPU container workloads (Nvidia, AMD, Intel OneAPI)

Simple integration with SLURM & other job orchestrators

- Single dependency on Python3.6 + container service of your choice


```
# HEP Benchmark Suite requires singularity 3.5.3+, python3.
module load singularity python3
python3 -m pip install --user git+https://gitlab.cern.ch/hep-benchmarks/hep-benchmark-suite.git

echo "Running HEP Benchmark Suite on $SLURM_CPUS_ON_NODE Cores"
srun bmkrun --config default
```

Data Access

Exascale challenge

Upcoming run4 (2027) expects **1 Exabyte physics data processing in 100 days**
Goal is to stream & process 10 PB of physics data through a HPC site in a day: several hundreds of Gbit/s continuously. HEP experiments can not store all the produced data at a single site.

- Challenge of increasing complexity: start with 10-20% goal, demonstrate management of hundreds of TBs data
- Maintain compute efficiency with high data rate in/out from/to storage & stream

Lots of moving parts! Break down challenge into three areas:

1. Data in/egress from HPC center
2. Efficient usage of storage systems on site
3. Dynamic scaling interaction between (1) and (2)

Data Access (cont)

Exascale challenge

Data in/egress from HPC center

- Modernising transfer tools to be able to fill 100Gb/s wide-area networks

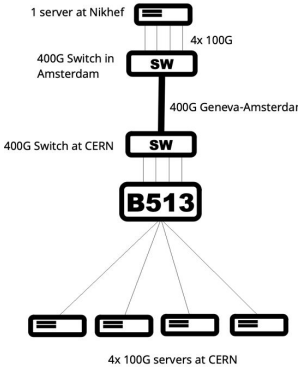
Efficient use of storage systems on site

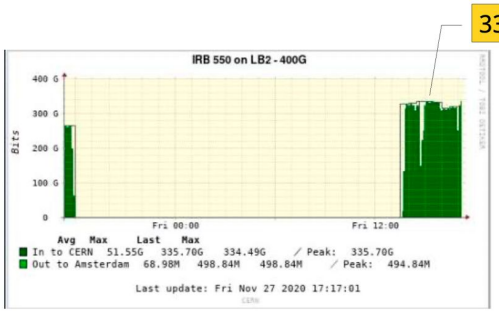
- Reducing the local footprint of data management components
- Exploring the amount of local storage needed at each scale

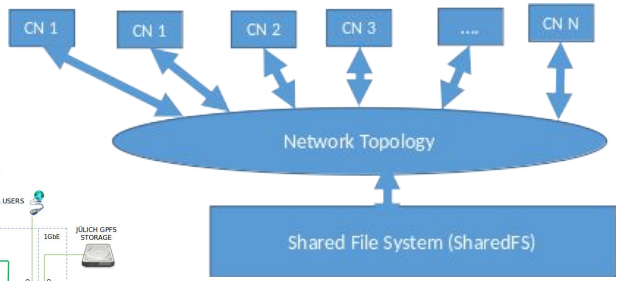
Local data delivery to the processing nodes

- Performance of local storage
- Local network structure

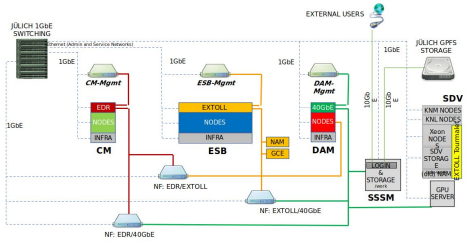
Pilot program to test all these elements







DEEP-EST Prototype – Schematic Network Overview



Collaborations

On the path to Exascale

NTNU | Norwegian University of Science and Technology

EGI-ACE

HPC Collaboration

CERN

GÉANT

PRACE

SKA

ROOT
Data Analysis Framework

intel

IBM

CERN openlab

E4
COMPUTER ENGINEERING

Micron

CINECA

SURF SARA

NVIDIA

RAISE
Center of Excellence

WLCG
Worldwide LHC Computing Grid

lrz

JÜLICH
Forschungszentrum

HPC LAB

FLATIRON INSTITUTE

DEEP
Projects

CSCS

SDSC SAN DIEGO SUPERCOMPUTER CENTER

Summary

CERN HPC Pilot

To meet a looming computational resource gap, CERN must evolve its computing platform to leverage heterogeneous computing and HPC systems

Developing benchmarking on HPC to enable:

- Profiling of workload performance critical to development and procurement
- Accounting of heterogeneous compute resources

Developing Data Access and

- Efficient ingress/egress via multiple channels, with increasing throughput
- Efficient usage of site storage systems, data staging and scaling
- Integrate with Authentication and Authorization efforts

Thank you!

Contact: **egi-ace-po@mailman.egi.eu**
Website: **www.egi.eu/projects/egi-ace**

EGI Foundation

@EGI_eInfra

EGI-ACE receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 101017567.