The HPC Collaboration: CERN, SKAO, GÉANT, PRACE

New Computing Challenges

As you learned in the talk of Joseph Flix on Tuesday, the needs for computing in data intensive science are increasing dramatically

- The HL-LHC and SKAO face unprecedent computing challenges over the next 5-10 years

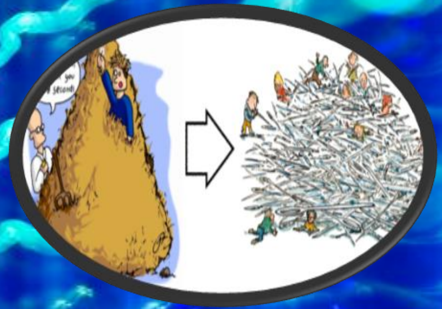
The needs for new resources are driving an extensive R&D effort

- Heterogenous hardware and High-Performance Computing (HPC)
  - There are big national investments in HPC
  - Improvements in heterogenous hardware are driving increases in capacity
- New methods and opportunities in AI/ML

Upgraded Accelerator
- Higher Luminosity

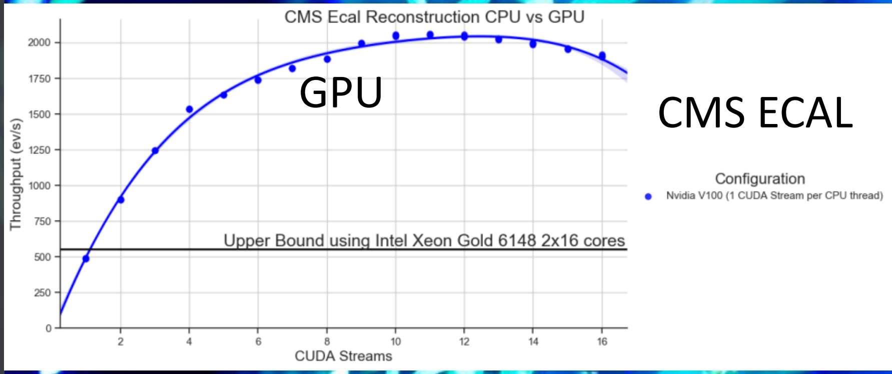
Upgraded Detectors
- Higher Granularity
- Higher Occupancy

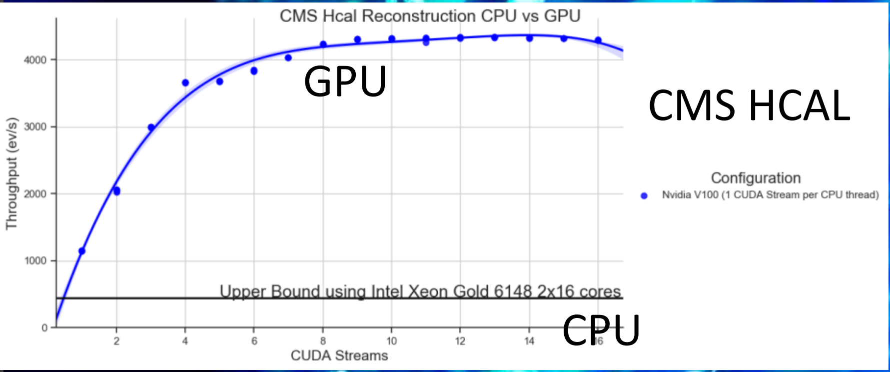
Changing Filtering Paradigms
- Higher Sensitivity
- Higher Data Rates



New Computing Challenges

- Explorations in heterogeneous hardware are one of the drivers of innovation in computing
  - Large improvements recently in processing performance have come from offloading work to GPUs
    - E.g. CMS HLT (**ECAL, HCAL)** with GPUs gives ~factor 10 in improvement in throughput (on NVIDIA V100)
- Reengineering the software is necessary
  - Results in performance gains, easier adoption of heterogeneous hardware and better maintainability
- Experiments are encouraged to make use of supercomputers (growing to exascale)
  - Common across US, EU, and Asia
  - HPC sites are early adopters of new technologies and sources of expertise





Motivation for an HPC Collaboration

Maria Girone
CERN openlab CTO

Changing Computing Landscape

General purpose CPU performance increases have slowed

Optimized heterogenous architectures have evolved faster, **HEP Is investing heavily in development to use new hardware resources**

- **GPUs** are the most common
- **FPGAs** currently used mostly in low latency applications
- **TPUs** and specialized ASICs are available

Accelerated Reconstruction And AI/ML

Low latency Online applications

		Accelerator				
		Intel	NVidia	AMD	FPGA	Other
CPU	Intel	Aurora	Cori Piz Daint Tsukuba Mare Nostrum		Tsukuba	
	AMD		Perlmutter JUWELS Booster	Frontier El Capitan LUMI		
	IBM		Summit Sierra Mare Nostrum			
	ARM		Wombat			Astra Fugaku

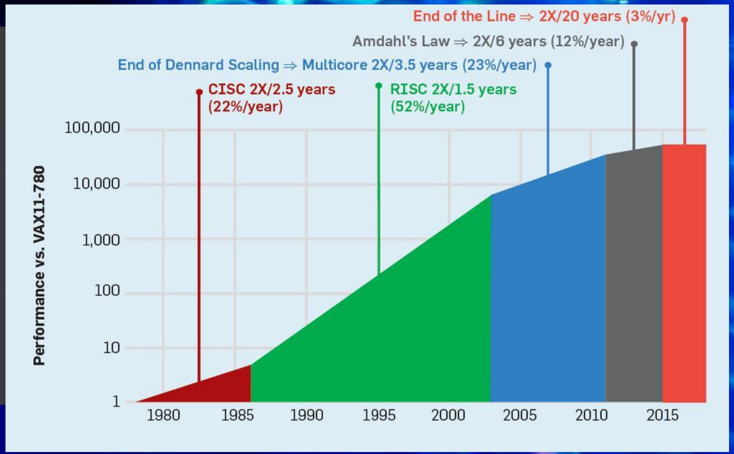
Amazon
Graviton2
Google Cloud TPU
Microsoft Azure
Intel DevCloud

General Purpose X86 processing resources

Code ported to Power

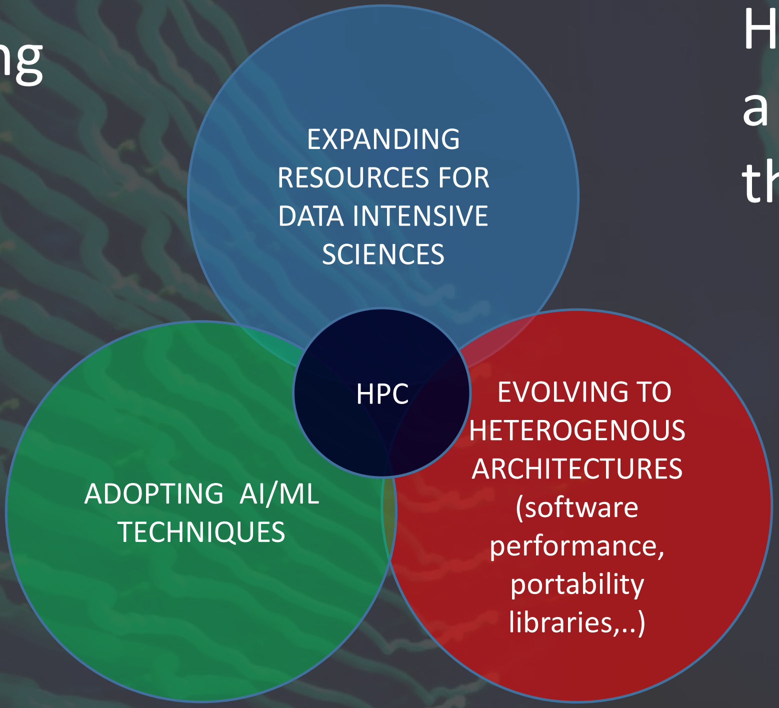
Low power highly parallelized

https://cacm.acm.org/magazines/2019/2/234352-a-new-golden-age-for-computer-architecture/fulltext



The current landscape is pushing HEP and other sciences to integrate HPC resources

HPC falls at the intersection of several important R&D areas



Engagement with the HPC Community can be a catalyst for progress

HPC Supercomputers will grow by a factor of 10 on the time scale of the HL-LHC

A thorough R&D program has been established

Unified programming models facilitate HPC adoption

High Performance Computing

Challenges

Software and Architectures

Supercomputers are early adopters of heterogenous architectures

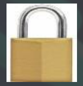
Benchmarking and Accounting

Performance on diverse architectures needs to be understood

Data Processing and Access

Enormous data volumes to stage, process, and export

Authorization and Authentication



Strict cyber security

Runtime Environments and Containers

Resources are shared, environment needs to be brought with the workload

Provisioning

Resources allocated for periods of time through allocations

Wide and Local Area Networking

Processing and storage resources are separate

Challenges in HPC Integration

The common challenges for HPC integration into LHC Computer were described in an engagement document

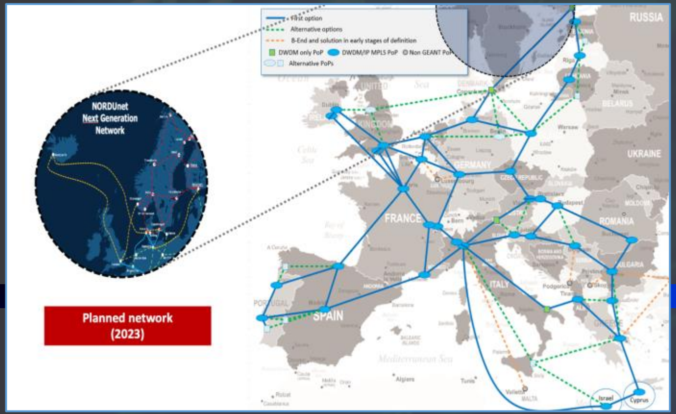
https://zenodo.org/record/3647548#.YBnA1y2cbVs

As we adapt

- Our consortium is ideally composed
  - HL-LHC and SKA have a burning physics need and in depth knowledge of the algorithms employed
  - PRACE provide considerable experience in the system adaptation of software environments
  - GEANT provides the infrastructure to take the computing to the many nodes that are needed to tackle the demand

[Signature Cerimony]

[From the HPC Collaboration Kick-off Workshop]



CERN, SKAO, GÉANT, PRACE Consortium

- Four areas of work have been identified as foundational. Progress will be evaluated by a series of common challenges and demonstrators
  - **Benchmarking**
  - **Data Access**
  - **Authentication and Authorization**
  - **Building a Common Center of Expertise**

- The next crucial step is to address the challenges through a **common program of work**
  - The **roadmap** is outlined on the next slides

The Four Pillars of the Collaboration

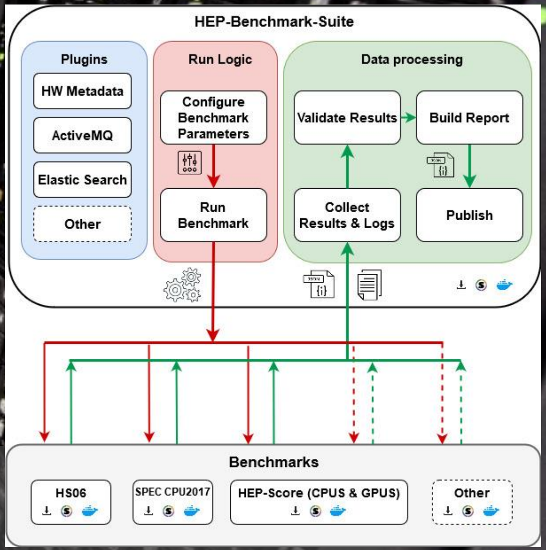
Benchmarking Activities

Open Question: How do we show we can effectively use HPC systems and account for contributions?

- PRACE-CERN-GÉANT-SKAO collaboration brings opportunity to expand capabilities using tools already developed for HPC sites by each community:
  - Unified European Applications Benchmark Suite (UEABS)- 13 workloads for HPC
- CERN is evolving the approach to benchmarking in HEP to embrace HPC:
  - Builds on experience from WLCG computing environment tools
  - Developed with secure, self-contained workload images (Singularity)
  - Assumes no privileges, no docker, limited/restricted node connectivity

PRACE Unified European Applications Benchmarking Suite

Application	Test Case	Piz Daint-HSW	JUWELS	MareNostrum4	Sisu	Hazel Hen	Irene-SKL	Piz Daint-P100	Marconi-SKL	DAVIDE-P8	Marconi-BDW	DAVIDE	Dibona	Frioul	Marconi-KNL	SDV	Irene-KNL
Alya	A		1.00	0.93				0.36					0.14		0.65	0.35	
	B		1.00	0.80				0.79							0.33		
	C											0.67		1.00			
Code_Saturne	A	0.91	1.00	0.69		0.96	0.68	0.14				0.14	0.18	0.19	0.18		0.15
	B	1.00	0.55	0.39		0.69	0.48	0.13							0.18		0.15
CP2K	A	1.00	0.57					0.20					0.27			0.32	
	B	1.00	0.95					0.18		0.85		0.06	0.53	0.06		0.30	
	C	1.00	0.89					0.36		0.15			0.31	0.31		0.31	
GADGET	A		1.00	0.84									0.65		0.19		
	B		1.00	0.95									0.71		0.23		
GPAW	S		1.00	0.69	0.94			0.14		0.74		0.22		0.10			
	M		0.44	0.34	0.32					0.06				1.00			
	L		1.00	0.77	0.80					0.13							
GROMACS	B		1.00	0.68		0.73	0.83	0.54				0.25		0.25	0.22	0.10	0.18
NAMD	B		1.00	0.47		0.95	0.62	0.41				0.61		0.15	0.14	0.12	
NEMO	G		0.89	1.00		0.82											
PFARM	1		1.00	0.84		0.38	0.82	0.69				0.14	0.13	0.33	0.32	0.29	0.24
	2		1.00	0.83		0.46	0.79	0.60				0.13	0.16	0.27	0.26	0.28	0.24
QCD	1		0.99	1.00			0.81	0.89				0.92	0.70	0.30	0.01	0.07	0.48
	2v1		1.00	0.99			0.75	0.90				0.72	0.52	0.27	0.24	0.04	
	2v2		0.60	0.59			0.56	1.00							0.31	0.12	
Quantum Espresso	S		0.96	0.50				1.00		0.67	0.76	0.57		0.24		0.18	
	L		1.00						0.56	0.98	0.21	0.51		0.17	0.18		
SPECFEM3D	C		0.35	0.32			0.34	1.00				0.08	0.03	0.10	0.01	0.18	0.04
	A		0.42	0.37		0.23	0.38	1.00					0.05	0.10	0.12		0.12
	B		0.37	0.45			0.42	1.00							0.04		0.15

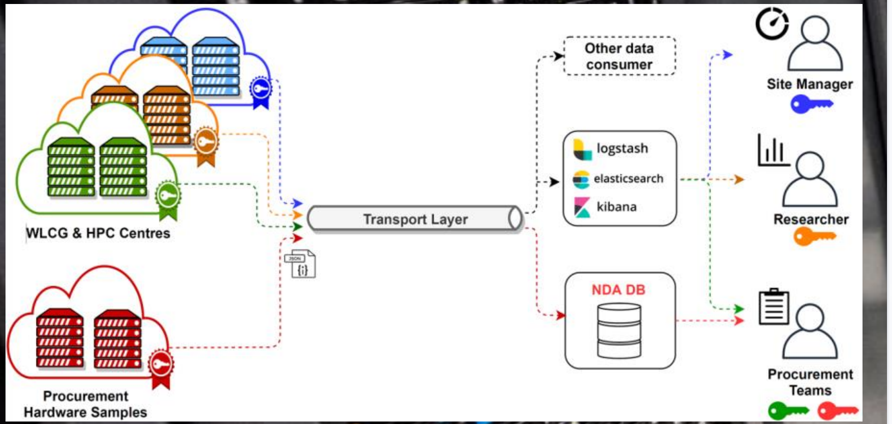


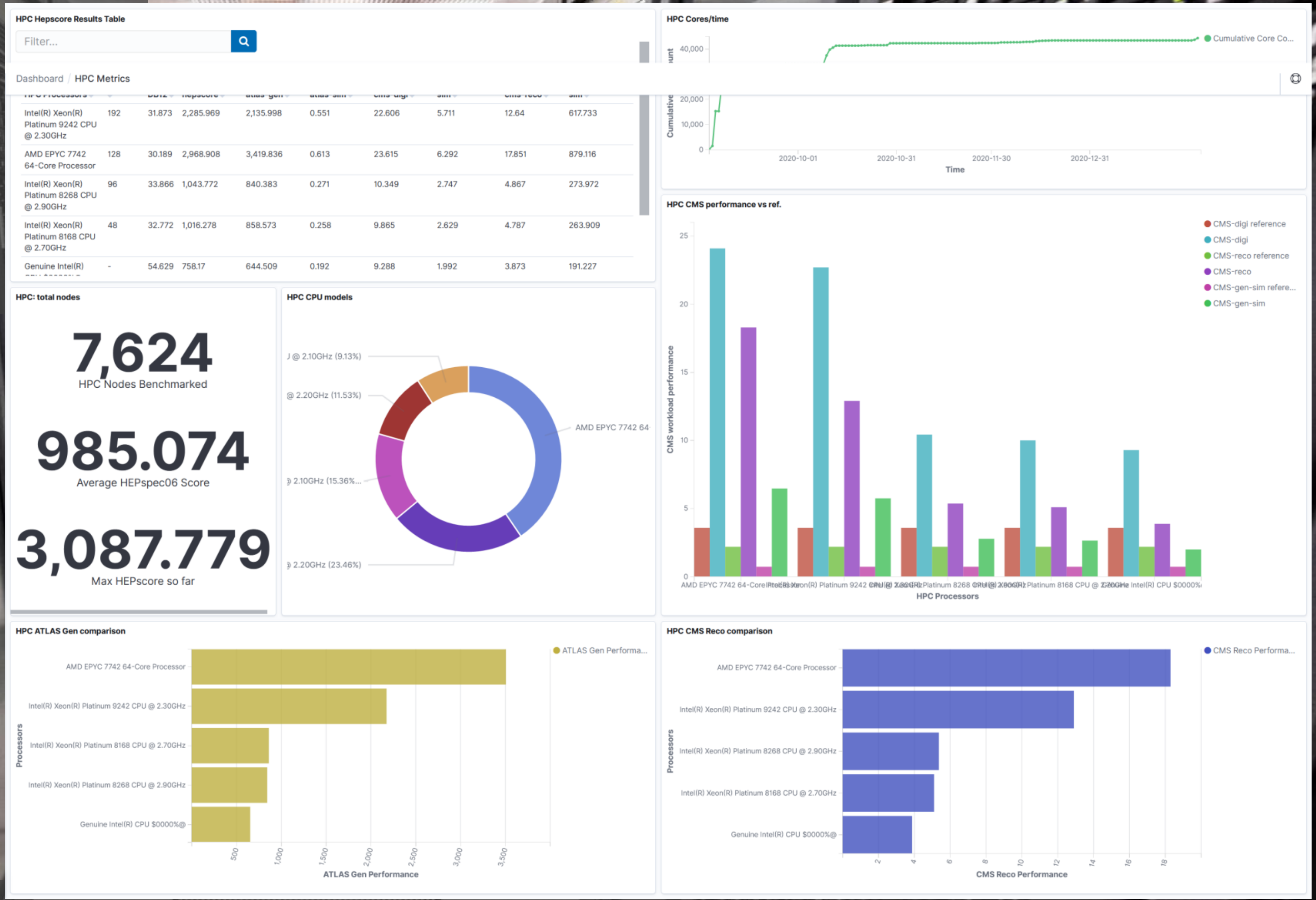
Benchmarking Demonstrator

Benchmarking Heterogenous architectures

- Multi-architecture as workflows become available (ARM, IBM Power)
- GPU accelerators (NVIDIA, AMD)

Automated collection and aggregation





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

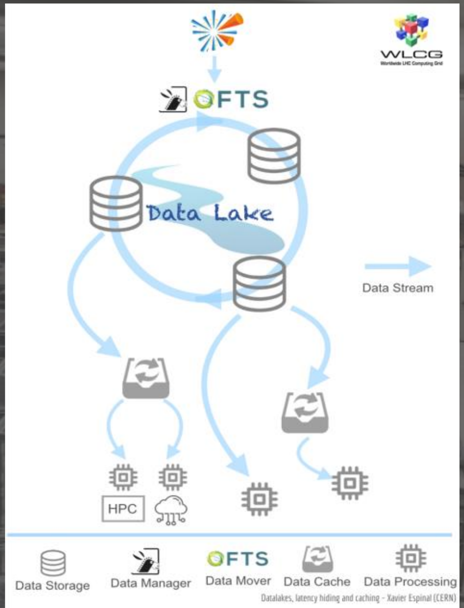
HEP Workflows on HPC

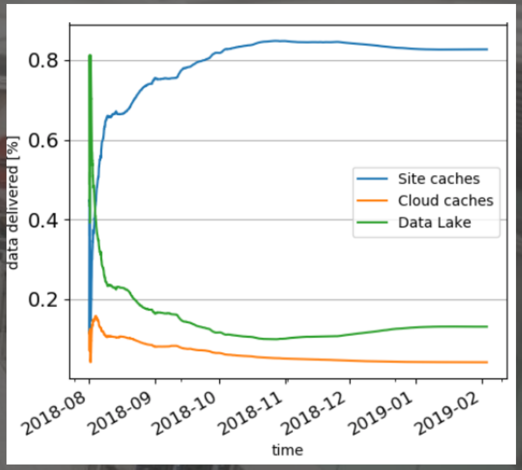
Courtesy of D. Southwick

Open Questions: How do we bring data to process on supercomputers?. Can we use HPC sites to produce simulation and reconstruction datasets at exascale?

WLCG *Data Lake* model separates storage and processing functionality. HPC will be a part of the *Data Lake* model

- Relies on caching and networking
- EuroHPC will have significant WAN connectivity and disk space

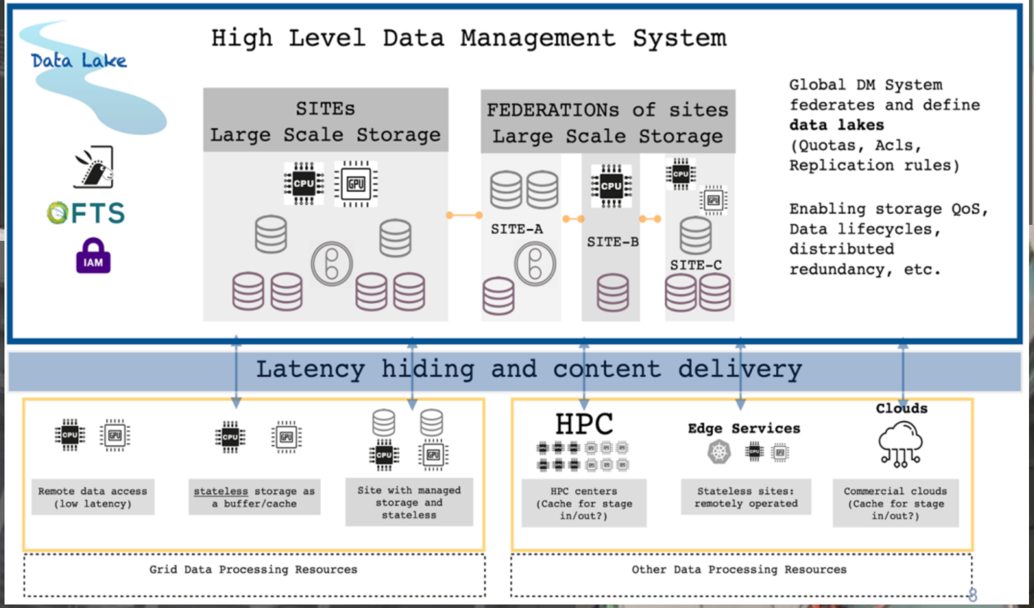




Emulation of cache delivery vs. time including regional caches

Technical Activities

Execute a series of data challenges to demonstrate the feasibility of the *Data Lake* model on a path to Exascale



Data Access – *Data Lakes*

- **With the help of Summer of HPC Students we started an IO Benchmarking project on HPC**
  - Evaluate and rank system performance.
  - Find limitations (more computing power isn't necessary more efficiency).

- **The goal of this project**:
  - Scale out I/O mock-ups using HPC benchmarks .
  - Evaluate and visualize performance metrics (e.g., bandwidth utilization).
  - Report performance under heavy dataflow load.

metric

Courtesy of V. Khristenko

Studying Data Access

❑CSCS Grand Tavé[1]

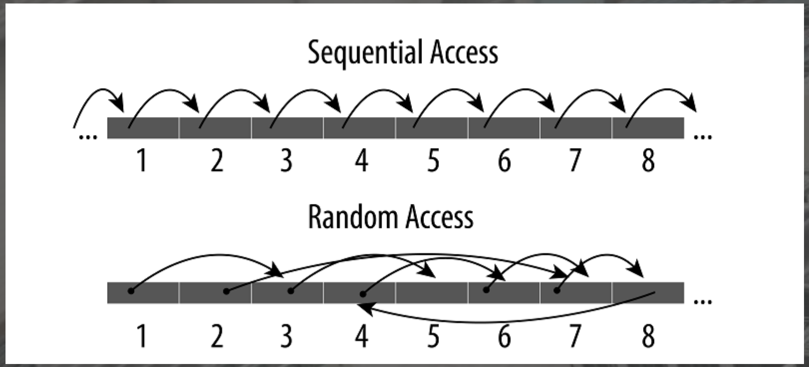
Specifications	
Model	Cray XC40 KNL
Compute Nodes	64 cores Intel(R) Xeon Phi @ 1.30GHz
Memory Capacity per Compute Node	96 GB, 16 GB HBM
Login Nodes	8 cores Intel(R) Xeon(R) @ 2.60GHz
Memory Capacity per Login Node	256 GB
Theoretical Peak Performance	436.63 TFlops
Max number nodes	164
Scratch capacity	/scratch/snx2000 904 TB

The $SCRATCH space (/scratch/snx2000/$USER) is connected via Infiniband interconnect to the system.

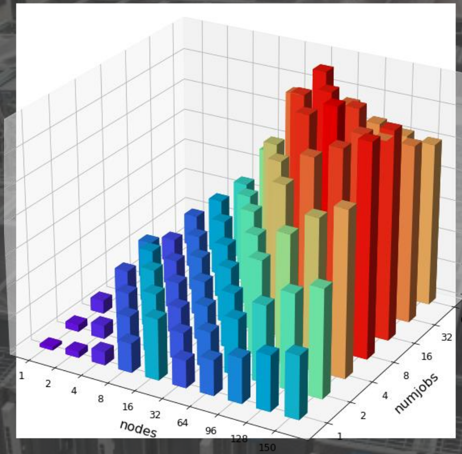
Courtesy of V. Khristenko

HPC Data Access Testbed

FIO (Flexible I/O tester) is an open-source synthetic benchmark tool
Generates various I/O type workloads (sequential/random reads and writes, etc.)

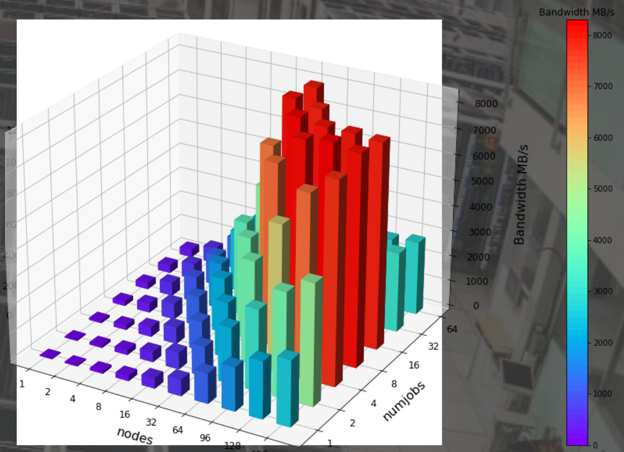


BS=16K
Peak: 1.37 GB/s



RANDOM

BS = 16K
Peak: 8.3 GB/s



SEQUENTIAL

Courtesy of V. Khristenko

Total Bandwidth vs Nodes

HEP Data in HPC

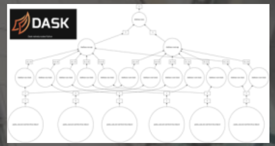
In addition to studying the file systems and storage performance, HEP is investigating how to improve the IO performance of the stored data files

HEP data is primarily stored as files, optimized for highly parallel HTC

- **ROOT** is the HEP analysis framework
- **ROOT** defines columnar data layout tailored for HEP: extreme throughput compared to alternatives
- https://root.cern

ROOT Challenges

- Maximize throughput I/O and optimize for HPC
- Optimize **persistent** data layout to facilitate conversion for CPU, GPU, SIMD (LLAMA), read patterns, and storage backend
- Ongoing R&D, bringing >4GB/s from off-the-shelf desktop to HPC
- **Scaling:** multi-threaded (>200 cores), distributed backends (dask / spark /...)



Courtesy of A. Naumann

ROOT

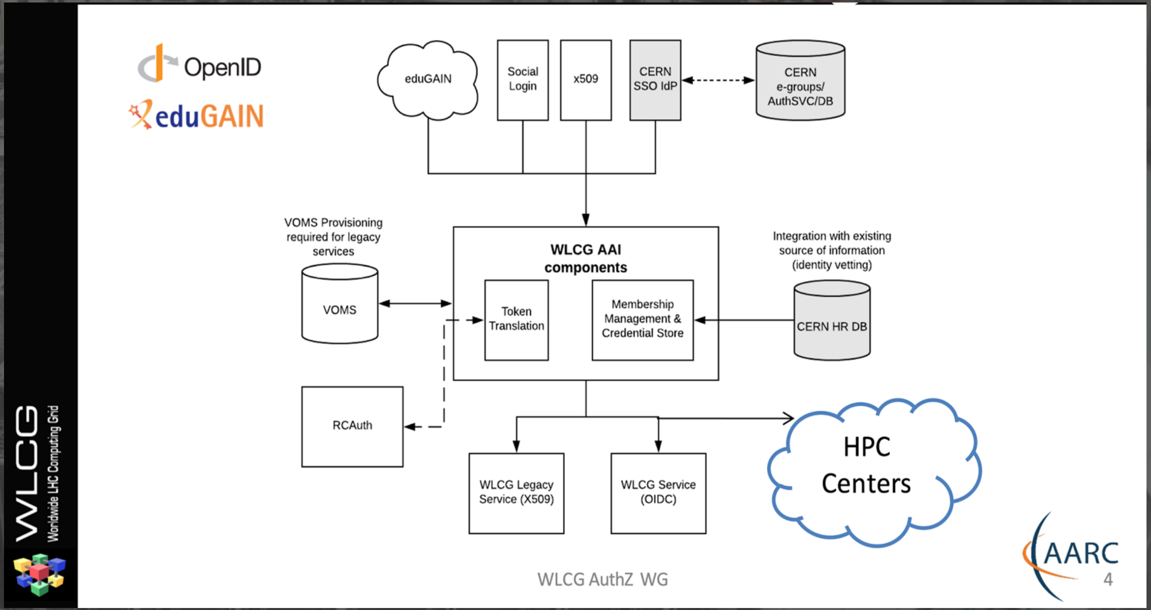
Data Analysis Framework

Open Questions

- Can we accommodate the strict cybersecurity requirement of HPC while enabling access by large international scientific collaborations?
- Can we capitalize on the efforts to modernize AAI infrastructure on the grid infrastructure?

Technical Program

- **Take advantage of growing experience in federated identity management (e.g. from AARC and GN4-3 EU Projects)**
- **Test OAuth2 token-based finer grained authorization on HPC, trusting token issuers that belong to scientific collaborations**
    - Token issuer for integration tests available



Authentication and Authorization Infrastructure

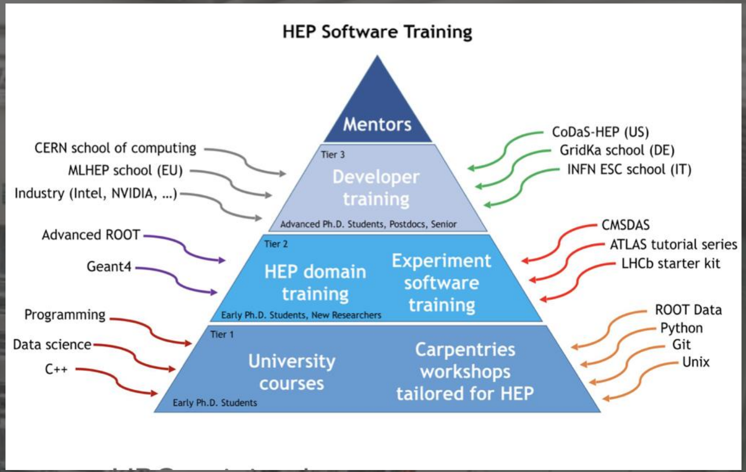
Open questions

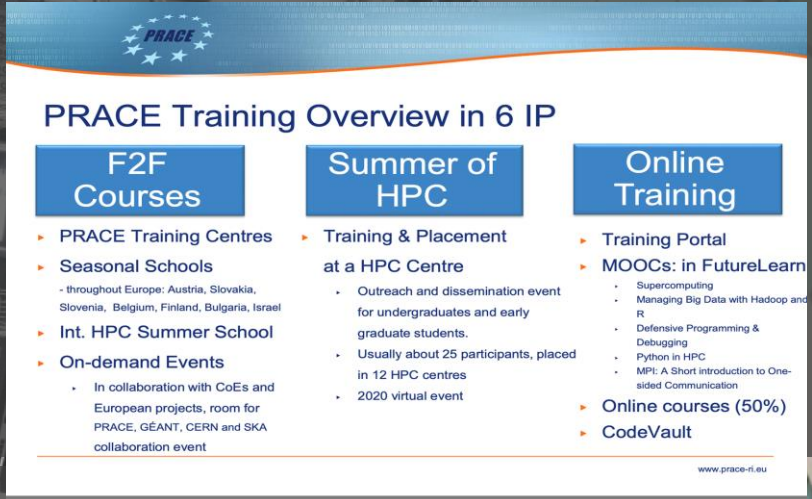
- How to build a common Center of Expertise
- How to make training relevant, scalable, and sustainable in HEP
  - Community investment in software is large, turnover is high, huge user and developer base

Training Program

- Dedicated events throughout 2021 on accelerator programming and performance tuning
  - Large expertise in PRACE

- Joint summer internship programs between PRACE Summer of HPC/CERN openlab in Q2 2021, focused on demonstrators



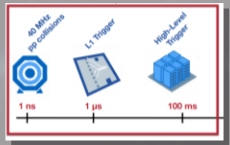


Building a Center of Expertise

Proven CERN capability

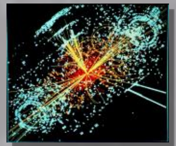
Use case specific

Fast ML



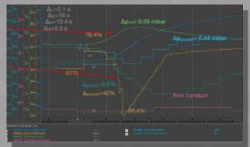
Ultra-fast on-edge inference under strict latency constraints

Anomaly detection



Object identification, classification, anomaly detection in big and noisy data sets

Industrial controls

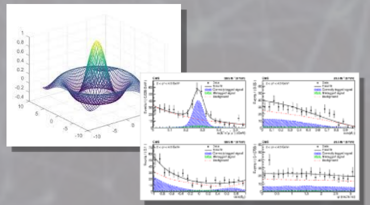


Machine efficiency and predictive maintenance with industrial control systems

Distributed computing

Optimization of distributed computing, storage, and networks; fast I/O for large files

Large scale, science grade data analytics and visualization

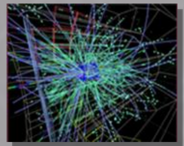


Cross use case

- Optimization and evaluation for science-grade precision of large data sets using advanced data analytics
- Data visualization, interactive plotting (e.g., statistical visualizations, uncertainties, distributions), model visualization
- Large-scale, quality-controlled CERN data as testbed/ benchmark (e.g., single data set with 100m examples, >1TB)

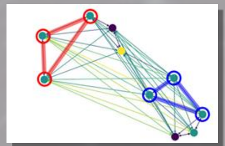
In development, opportunity for joint R&D

Simulation

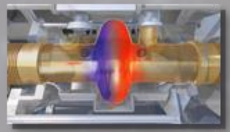


Simulation and reconstruction with generative DL for efficient computation

Graphs



Exploring Graph NNs for high-multiplicity problems with non-linear distances

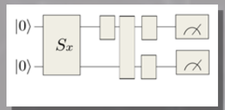


Determining optimal machine design and component configuration

ML in Robotics

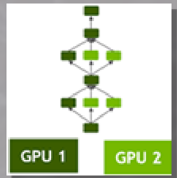
Remote maintenance and safety with autonomous robots and computer vision

Quantum ML



Research quantum algorithms to solve pattern recognition, classification and generation problems

Computing parallelization

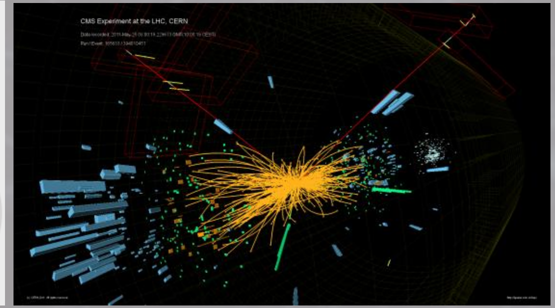


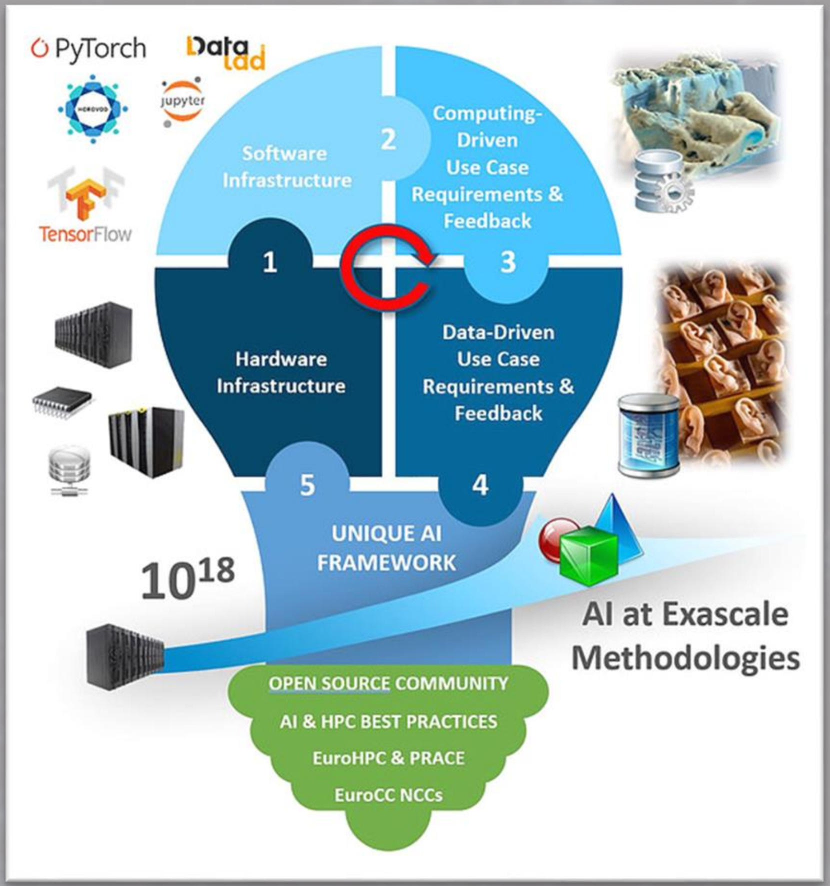
Training and optimization of complex NNs on parallelized GPU infrastructure

Progress on AI/ML Capabilities

Launched in January the RAISE Center of Excellence enabled researchers from science and industry to develop novel, scalable Artificial Intelligence technologies towards Exascale along representative use-cases from Engineering and Natural Sciences

CERN is leading the leading the data driven use-cases





AI/ML Projects

RAISE Center of Excellence

Testbeds are critical resources for development progress

- Representative hardware architectures, network structures and scale

- Some of the PRACE Tier-0 testbeds will be accessible in Q1 2021 through the collaboration for executing the working groups demonstrators
  - Benchmarking
  - Data access
  - AAI

Testbeds

Collaborations

HPC Collaboration

CERN, GÉANT, PRACE, SKA Square Kilometre Array

RAISE Center of Excellence

DEEP Projects

ROOT Data Analysis Framework

EGI-ACE

WLCG Worldwide LHC Computing Grid

HPC LAB

NTNU | Norwegian University of Science and Technology

FLATIRON INSTITUTE

SDSC SAN DIEGO SUPERCOMPUTER CENTER

intel

IBM

CERN openlab

E4 COMPUTER ENGINEERING

Micron

Maria Girone, CERN openlab Technical Workshop 2021

Participation on the path to exascale

Towards a Successful Collaboration

- All stakeholders have engaged and are expected to contribute to the fundamental tasks of **effectively** HPC systems, securely accessing them and intensively accessing data
  - Additional contributions are very welcome
- The next months are crucial to demonstrate collaborative use of HPC systems
  - Access to testbeds proving technical benefits
- Leveraging experience from PRACE and GÉANT as centers of expertise is valuable to our community
  - We will provide application knowledge for a joint program on software optimisation on HPC systems