16st of March 2023

CERN openlab

Heterogeneous Architectures Testbed at CERN

Joaquim Santos (CERN)
Krzysztof Mastyna (CERN)
Luca Atzori (CERN)

What we do

- Reception, installation, configuration and benchmarking of new hardware
- Project and User Support: access grant and system administration
- Support for Intel tools on CVMFS
- Hardware and Software maintenance
- Organisation and support for workshops, training courses...
  - Intel Software Tools training, 2022

Some numbers

- 100+ different users and 281 accounts
- ~ 95 systems: mostly bare-metal nodes + some VMs
- Different operating systems in use:
  - Mainly CentOS Stream 8, RHEL 8 and CentOS 7
  - RHEL 9 and CS9 being gradually adopted
- SNOW FE support (requests and incidents):
  - 280+ tickets handled in 2022
  - 30+ tickets in 2023 and counting...

Hardware: CPU Systems

- x86 (Intel Xeon)

  Sapphire Rapids, **Ice Lake**, **Cascade Lake**, **Skylake**, **Broadwell**, **Haswell**, **Ivy Bridge**, **Sandy Bridge**

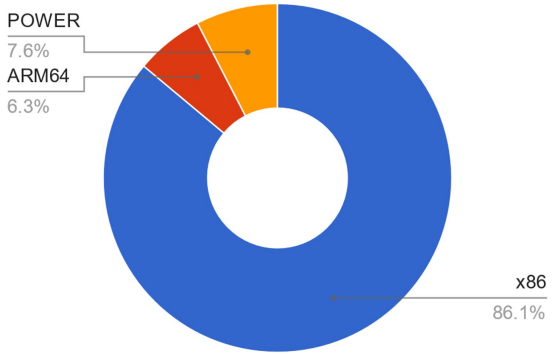
FULLY IN USE	
FEW AVAILABLE	
AVAILABLE	

- ARM64

  Cavium ThunderX and **ThunderX2**

- IBM Power

  POWER8

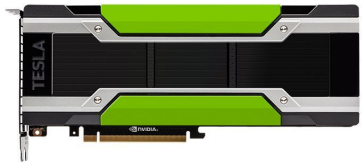


Main use-cases	
EP-SFT	CMS
ATLAS	HEP Benchmarking team

Hardware: Accelerators

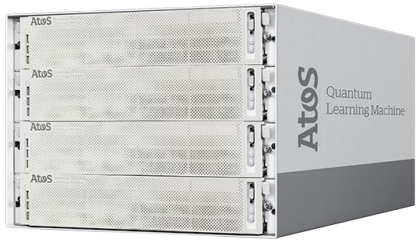
- openlab ML projects - NVIDIA Tesla V100, Tesla P100 and NVIDIA T4
- oneAPI - Intel ATS-P
- openlab QTI - NVIDIA Tesla V100S
- Older accelerators – AMD Vega 10, Alveo U200, Altera Arria 10

Main use-cases	
openlab and QTI	CMS
ATLAS	LHCb/Allen



Hardware: other technologies

- ATOS QLM – Quantum Learning Machine
- NVIDIA Blue-Field 2 DPU
- Intel Omni-Path (DAOS)
- InfiniBand
- Intel Optane Memories (DAOS)
- Intel Xeon Phi KNL, one Xeon-D node



PROJECTS AND ACTIVITIES

Overview

Projects and activities

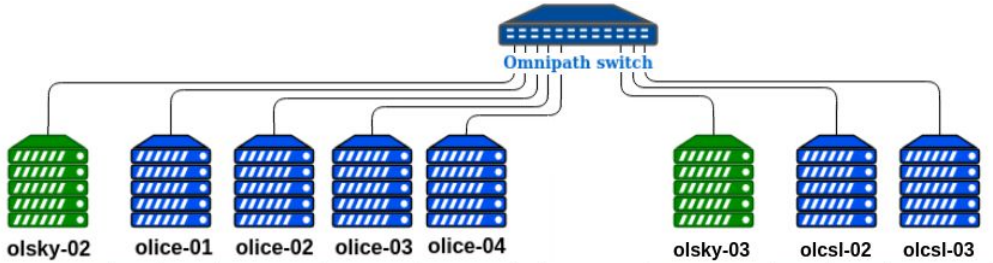
- E4/NVIDIA project
- Intel DAOS
- Benchmarking new platforms (with the HEPscore and HPC team)
- GPU support for ML
- Quantum Technology Initiative
- oneAPI and Intel tools

Main orchestration technologies

- **Puppet**, Ansible
- Docker, Singularity
- Production-like approach using most of the IT dept infrastructure tools
  - Foreman
  - Linux ai-admin tools
  - Gitlab
  - …and many more

DAOS (Distributed Asynchronous Object Storage)

- Open-source high performance storage
- Relying on distributed Intel Optane persistent memories and PCIe NVMe
- Current setup:
  - Ice Lake cluster: version 2.3 (development version)
  - Cascade Lake cluster: version 2.2 (official stable version)



DAOS Agent	Yes	No	No	No	No
DAOS Server	No	Yes	Yes	Yes	Yes
DAOS Head Node	No	No	No	No	Yes

DAOS Agent	Yes	No	No
DAOS Server	No	Yes	Yes
DAOS Head Node	No	No	Yes

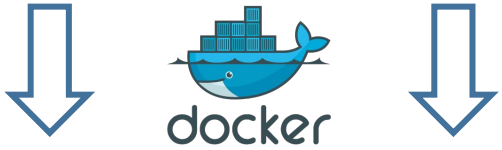
- Communication through TCP over Omni-Path fabric interface (ofi+tcp)
  - opx tests from v2.4

Main use-cases
ROOT
ATLAS

GPU - Machine Learning activities

Cope with high demand

Flexibility for the users



- Jupyter notebooks
- Support for CUDA versions, TensorFlow and other packages
- Plan for the future use of Kubernetes

Main use-cases
openlab ML
openlab QTI

Kubeflow (ml.cern.ch) with openlab GPUs

Thanks to Dejan Golubovic and Ricardo Rocha!

- Using Kubeflow interface created by the PW-PI team with openlab resources dedicated to openlab users (done through egroup)
- Plans to add more GPUs (testing process still on-going)

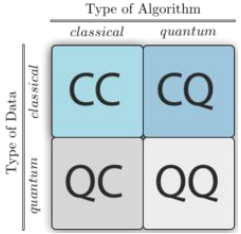
Quantum Technology Initiative

- Four project-dedicated GPU nodes with quantum environments: **Cirq**, **PennyLane, Qibo, Qiskit, cuQuantum SDK**
- See our documentation

- ATOS QLM Appliance
- Remote access to IBM Quantum resources

QML for High Energy Physics

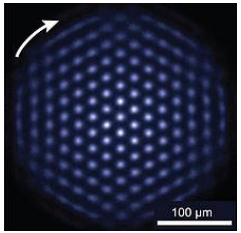
Quantum Machine Learning



Free energy based reinforcement learning FERL

Quantum Classifiers

Quantum Neural Networks



Quantum Key Distribution

Quantum Simulations

Supporting multi science projects

Some examples:

- Benchmarking new hardware (HS06, HEPscore23,…)
- Support to experiments building software on ARM and IBM Power architecture

Geant4

oneAPI development

ROOT

What happens in our nodes?

BioDynaMo

Deep learning, cryptography

3DGAN + Quantum GAN

oneAPI and Intel tools

- **oneAPI** → Latest available version: **2023.0.1**

- Available on CVMFS:

```
$ source
/cvmfs/projects.cern.ch/intelsw/oneAPI/linux/x86_64/2023/setvars.sh
```

More information: see our documentation.
Register to the **intel-tools-announcements** e-group to get the latest updates!

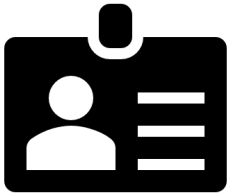
Heterogeneous Architectures Testbed - future plans

Aims to provide coordinated access to CERN developers and researchers to a broad set of heterogeneous architecture solutions

Apart from the architectures already shown (single entry point for the users):

- on-prem/remote via CERN openlab E4/NVIDIA project
  - NVIDIA CPUs and GPUS, including GRACE and HOPPER
  - ARM Marvell and Ampere CPUs
  - RiSC-V CPUs
  - EPI
  - FPGAs
  - AI-specialised architectures
- remote, via HPC supercomputers testbeds
  - Leveraging on the strong connection with PRACE and EuroHPC
- remote, via commercial cloud-hosted systems

How to contact us



- Open a SNOW ticket here
- Check our website: https://openlab-systems.web.cern.ch/
- ...or write directly to us!

  openlab-systems@cern.ch

Thank you!