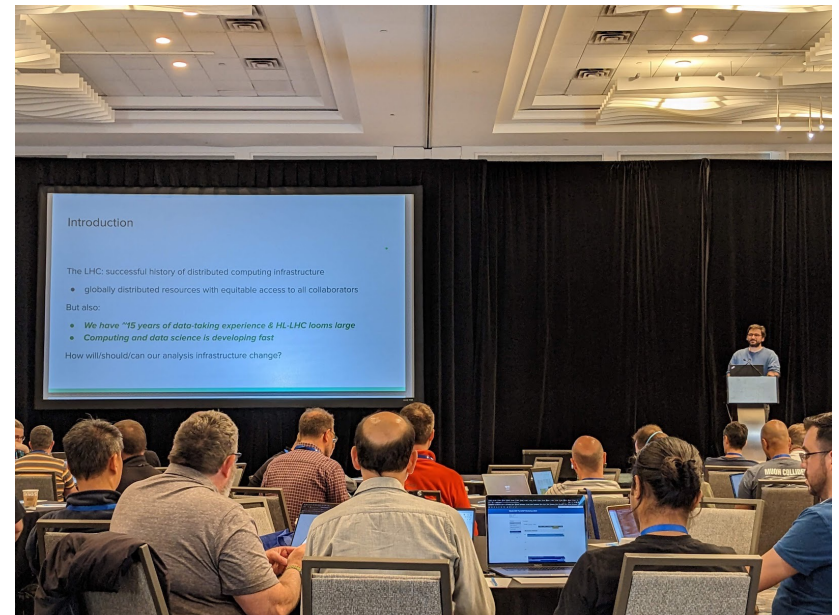# WLCG Workshop
Summary and Next Steps

S. Campana (CERN)

WLCG/HSF workshop

Two sessions (half a day each) in the weekend before CHEP2023 (Norfolk)

Attended by ~150 people in person and few remote. Which is really a lot for a Saturday/Sunday event.

- Analysis Facilities: how will the end user analysis look like at HL-LHC and will we need specialised facilities with dedicated services and hardware? Impact on WLCG?

- Heterogeneous Architectures: what is the progress made by the WLCG community to be able to use non-X86 architectures and what should be the focus of WLCG in the coming years?
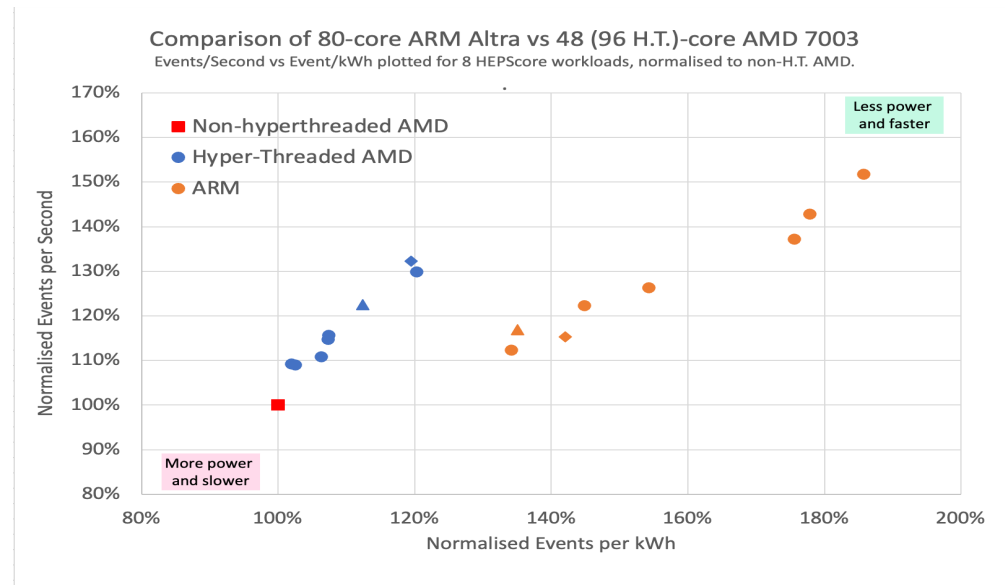
WLCG/HSF workshop: [heterogeneous architectures](#)

ARM is the most promising non-X86 CPU architecture. High HS23/KW ratio. Many HEP workloads ported, some sites planning to provide resources. HS23/EUR varies in different countries. Too early to pledge, but the trend is clear

Power8/9: available at some HPC but no future prospects. Current CERN offering (machines for porting and building) is adequate

GPUs: ALICE running offline reconstruction on CPU/GPUs hybrid (O2). CMS plans to be ready by end of 2023.

WLCG sites should work with experiments to gain experience running offline applications on CPUs+GPUs.

Comparison of 80-core ARM Altra vs 48 (96 H.T.)-core AMD 7003
Events/Second vs Event/kWh plotted for 8 HEPScore workloads, normalised to non-H.T. AMD.

Heterogeneous Architectures – Next Steps

- The HSCORE23 Task Force is closed but the HEPIX Benchmarking Working Group should work on the evolution of the benchmark for different architectures

- We need to revive the effort in Technology Watch. There is a HEPIX WG also on this, a bit dormant. Need to revive or start an initiative in WLCG

- Start a WLCG-wide discussion (sites-experiments) about the timeline for pledging ARM resources in WLCG. We need enough supported workflows

- Gain experience is scheduling jobs in a mixed CPU-GPU environment in WLCG, with the available applications

- Start discussing a model for pledging GPUs in WLCG. Review the current opportunistic GPU offering in WLCG and the experiments' needs

WLCG/HSF workshop: Analysis Facilities

- Work organised within the **HSF Analysis Facilities Forum**, created early 2022 (sites, users, experiments)
- An update was presented to the June WLCG Overview Board and to the LHCC

- A loose definition
  - *Infrastructure and services that provide integrated data, software and computational resources to execute one or more elements of an analysis workflow. These resources are shared among members of a virtual organization and supported by that organization.*
- Do we have analysis facilities at the moment?
  - CERN (all WLCG users), NAF (only German users), FNAL (CMS), BNL (ATLAS),…
    - But also T3s, grid, commercial clouds….
    - GSI and Wigner for Alice are highly specialised sites for analysis trains without interactivity

---

Why are we looking at this?

- Evolution of technologies and analysis techniques
- HL-LHC increased data and interactive computing requirements
  - Much higher trigger rates - many more events for analysis - may need to prototype on larger datasets
- Improve the users' experience
- Improve site maintainability   Technology Evolution   Techniques Evolution
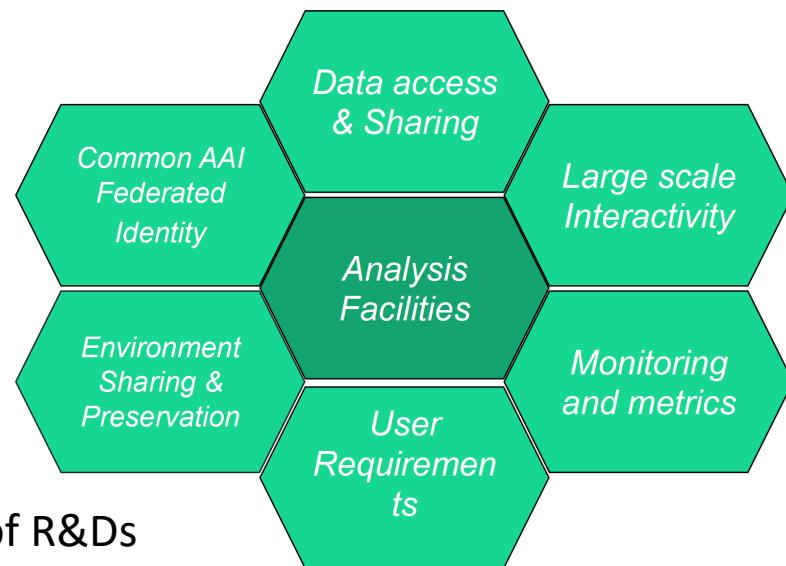
User Requirements and R&D

Starting points: analysis user experience

- Analysers often leave the global grid early in the analysis cycle to work on local resources
- Different access systems between grid and local resources (ssh, batch, interactive)
- Sharing outputs is a concern, usually via local file system
- Interactive analysis mostly defined by what fits onto a single machine

Lead to a set of R&Ds to improve:

- Scalability
- Ease of use
- Data access
- Interoperability
- Collaboration

Data access & Sharing

Common AAI Federated Identity

Large scale Interactivity

Analysis Facilities

Environment Sharing & Preservation

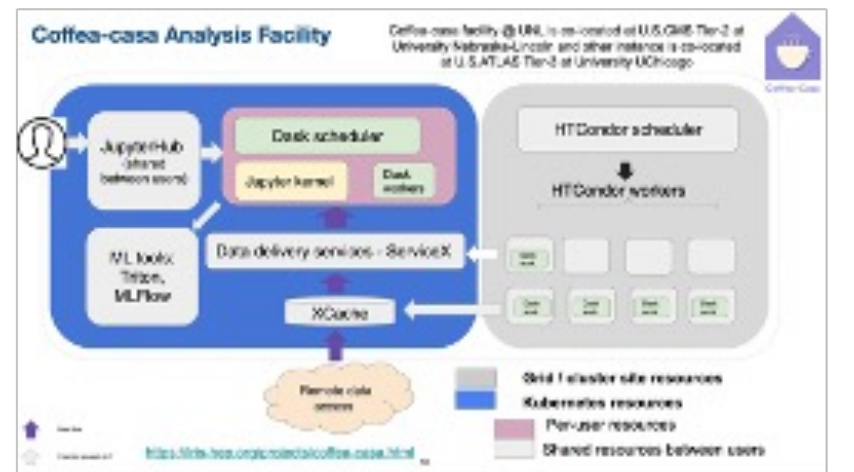User Requirements

Monitoring and metrics

Areas of R&Ds

Analysis Facilities implementations

- Analysis Facilities are supported and are being developed at different sites
- The R&Ds are guided by facilities and users
- Most of the building blocks for AF R&D are available
- Still many open questions (e.g. need of ssh, Jupyther Hub, "other" workflows, POSIX ?)

Analysis Facilities next steps

The discussion focused a lot so far on technologies. Many building blocks were identified

At the level of use cases there is a lot of uncertainty: discussions about scaling of interactive analyses and what "interactive" really means. No current real definition of the ML needs

The **LHCC recommends** that experiments engage in the process of developing and defining the structure of the future Analysis Facilities and **requests** they produce a document which defines the use cases in order to establish realistic benchmarks. This process should be coordinated with the HL-LHC Computing and SW review panel. The document is expected to be regularly updated in the process towards HL-LHC.

The GDB could organise focused discussions around selected R&D topics: AAI, data sharing, data access, environment and software sharing, monitoring