Analysis Facilities

A. Forti

GridPP52

28 August 2024

Areas of work

- Manchester WP-D
  - DOMA and GPUs for Analysis Facilities 0.55
- Planning
  - GPUs
  - Container registries
  - Data access
- Done
  - White paper publication
  - WLCG WS organisation
  - LHCC charge

Driving paradigms

AFs do not exist in a vacuum and should not be considered isolated

AF are not grid sites due to a strong interactive non centralised component but should be integrated to make it easy to exploit the grid resources as needed.

GPUs

- Manchester has 36 T4 GPUs and 3 v100
  - Distributed between the T2 and the T3
  - All maintained with the same puppet instance and reinstalled if they need to change destination
    - Not k8 automatic but it's sufficient
- Currently there are several users asking for the interactive GPU access
  - Interactive users usually tend to install also their own software
    - There is no solid software distribution for GPU users and what they are doing is likely not portable
  - When their needs for data grows they move to CERN
    - Storage is what makes the difference

GPUs (2)

- On the grid the GPUs are basically empty
  - Grid access is dedicated to ATLAS, DUNE and manchester users
- ATLAS runs only tests
  - HammerCloud Basic Functional Tests I wrote
  - ART reconstruction test which are WIP
- Looking at the submission the CEs are very limited
  - ARC-CE doesn't pass any GPU requirement (not even in ARC-7)
  - HTCondor-CE can request a GPU and pass the number of GPUs but not much else
- For now it is ok but it is not clear devs will put any work if there isn't an increase in usage
- Quantities the user may really want to request the the GPU memory, the number of GPUs and the libraries version
  - Currently there is also disagreement on who should maintain the libraries
  - Experiments are putting them in cvmfs but not all libs work on all models

GPU(3)

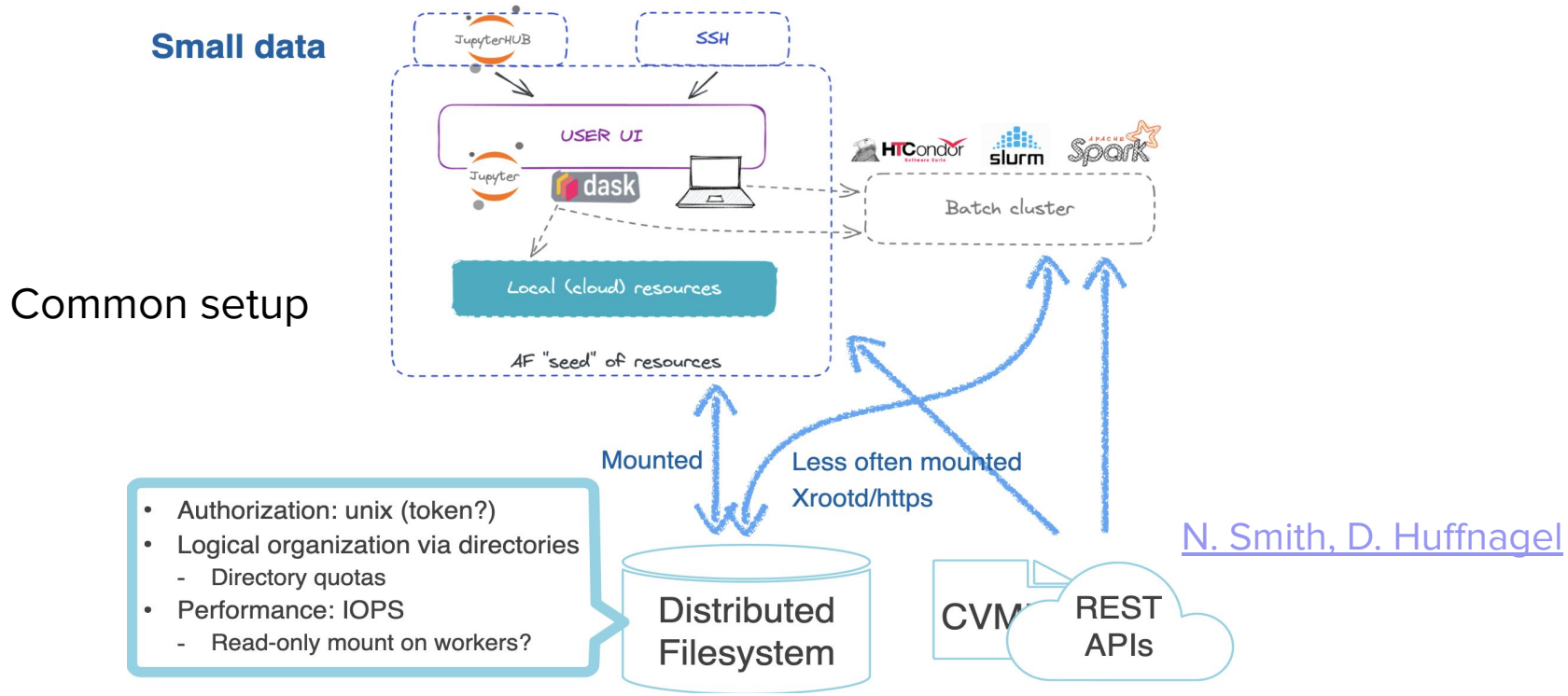
- Experiment monitoring can effectively monitor only if a job landed on a queue with GPUs
- Any other quantity is for CPUs walltime, HEPSPECs, memory are all tailored for CPUs
  - Glasgow is also interested in the benchmark for GPUs
  - Without even going in mixed jobs monitoring not ready
- prmon was modified at some point to include some GPU information but it only works if there is 1 GPU
  - the nvidia tool that was incorporated refuses to return information
  - Could be be something to work on
- CERN prototype AF adopted prmon too to instrument the jobs as recommended in the WP
- Even considering just the infrastructure there is plenty of work to develop or to push for development

GPU (4)

- Manchester ATLAS also put a successful bid to work on improving software efficiency and produce reccomdanations for users to use GPUs more effectively for the next two years
  - It involves reviewing and establishing methods and metrics for evaluating the cost of computing hardware, the effective policies in place for re-use, as well as improving the energy efficiency of scientific software on different computing architectures by analysing selected frameworks in pilot studies. The project also seeks to bridge the gap between individual scientists developing and running software on their own machines and larger-scale software development needed.
- Cooperation with the Tier2 for this is already foreseen

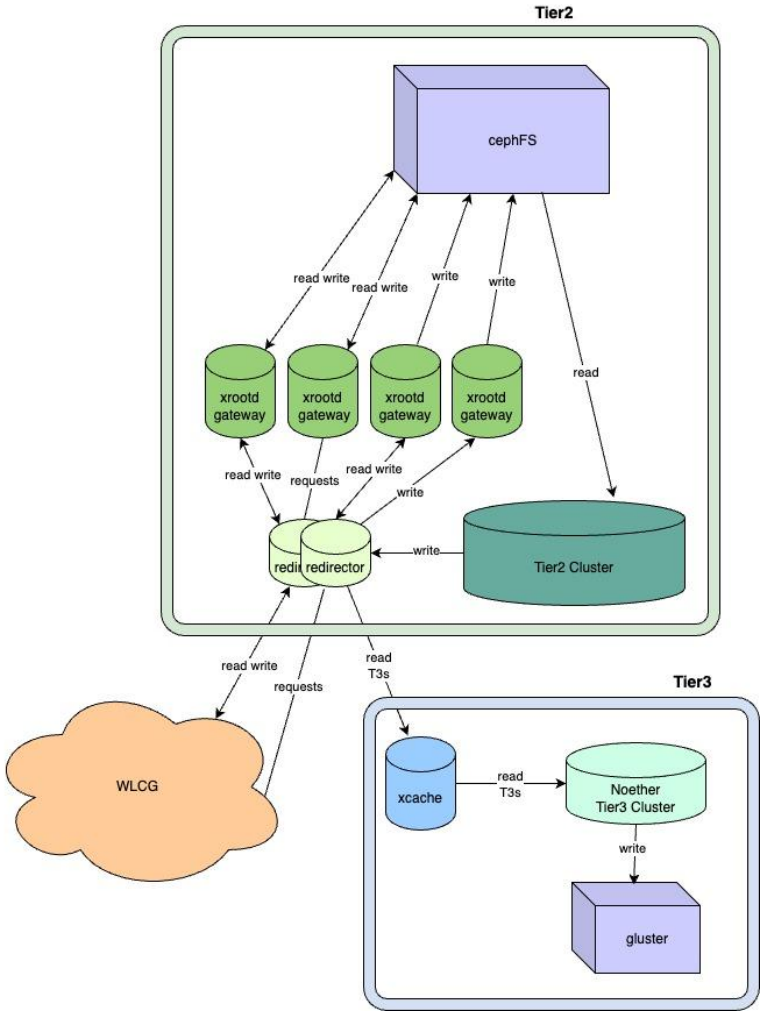
Data Access

- For an AF data access is extremely important
  - Even with reduced formats like PHYSLITE and nanoAOD.
- WLCG/HSF WS talks below show this quite effectively
  - 200 Gb/s analysis challenge
  - Data Access discussion (POSIX vs Object Stores 🥷)



Data Access (2)

- Without starting from the extremes discussed in those talks (yet) working with xcache is the first step
  - Also strongly related to Virtual Placement solution in ATLAS which still require work.
  - Now that we moved off DPM to xrootd+cephfs we can plug-in in the current developments more easily
- ServiceX, the data transformation service used in the US AFs and in the 200 Gb/s challange is something we should like to look at.
  - T1 has also some interest here



Container registries

- Containers have been considered a solution to portability
  - But… users have always been in the wild as far as a proper registry infrastructure was concerned.
  - We allowed to use any registry and there was no place where to point the users and no real attempt at curating the images or offer base images at scale.
- New WG to work on the registry infrastructure has been proposed by CERN IT (time scale for the initial testing is 6 months)
  - Plan and test the scalability of container registry usage within WLCG:
  - The registry service should be able to handle the required scale in number and size of images, as well as throughput for pushing and pulling image artefacts
    - For an estimate of scaling targets, ATLAS suggests 20TB and order of 10k images
  - Develop a pilot for a container registry serving a distributed computing infrastructure with cache replicas at multiple locations.
- This is an important part of the infrastructure for AF (and grid) jobs.

White paper

- Analysis Facilities white paper is progressing its publication iter
  - On ArXiv as a pre-print since April
    - https://arxiv.org/abs/2404.02100
- Submitted to Springer Computing and Software for Big Science
  - Reviews are quite positive 👍
  - Should be published in the next few weeks

HSF-TN-2024-01
April 2024

Analysis Facilities White Paper

D. Ciangottini[1,b], A. Forti[2,b], L. Heinrich[3,b], N. Skidmore[4,b], C. Alpigiani[5], M. Aly,[2], D. Benjamin[6], B. Bockelman[7], L. Bryant[8], J. Catmore[9], M. D'Alfonso[35], A. Delgado Peris[17], C. Doglioni[2], G. Duckeck[10], P. Elmer[11], J. Eschle[12], M. Feickert[13], J. Frost[14], R. Gardner[8], V. Garonne[6], M. Giffels[36], J. Gooding[15], E. Gramstad[9], L. Gray[16], B. Hegner[31], A. Held[13], J. Hernández[17], B. Holzman[16], F. Hu[8], B. K. Jashal[18,19], D. Kondratyev[20], E. Kourlitis[3], L. Kreczko[21], I. Krommydas[22], T. Kuhr[10], E. Lancon[6], C. Lange[23], D. Lange[11], J. Lange[34], P. Lenzi[1], T. Linden[24], V. Martinez Outschoorn[27], S. McKee[25], J. F. Molina[17], M. Neubauer[26], A. Novak[35], I. Osborne[11], F. Ould-Saada[9], A. P. Pages[28], K. Pedro[16], A. Perez-Calero Yzquierdo[17], S. Piperov[20], J. Pivarski[11], E. Rodrigues[29], N. Sahoo[30], A. Sciaba[31], M. Schulz[31], L. Sexton-Kennedy[16], O. Shadura[32], T. Šimko[31], N. Smith[16], D. Spiga[1], G. Stark[33], G. Stewart[31], I. Vukotic[8], G. Watts[5],

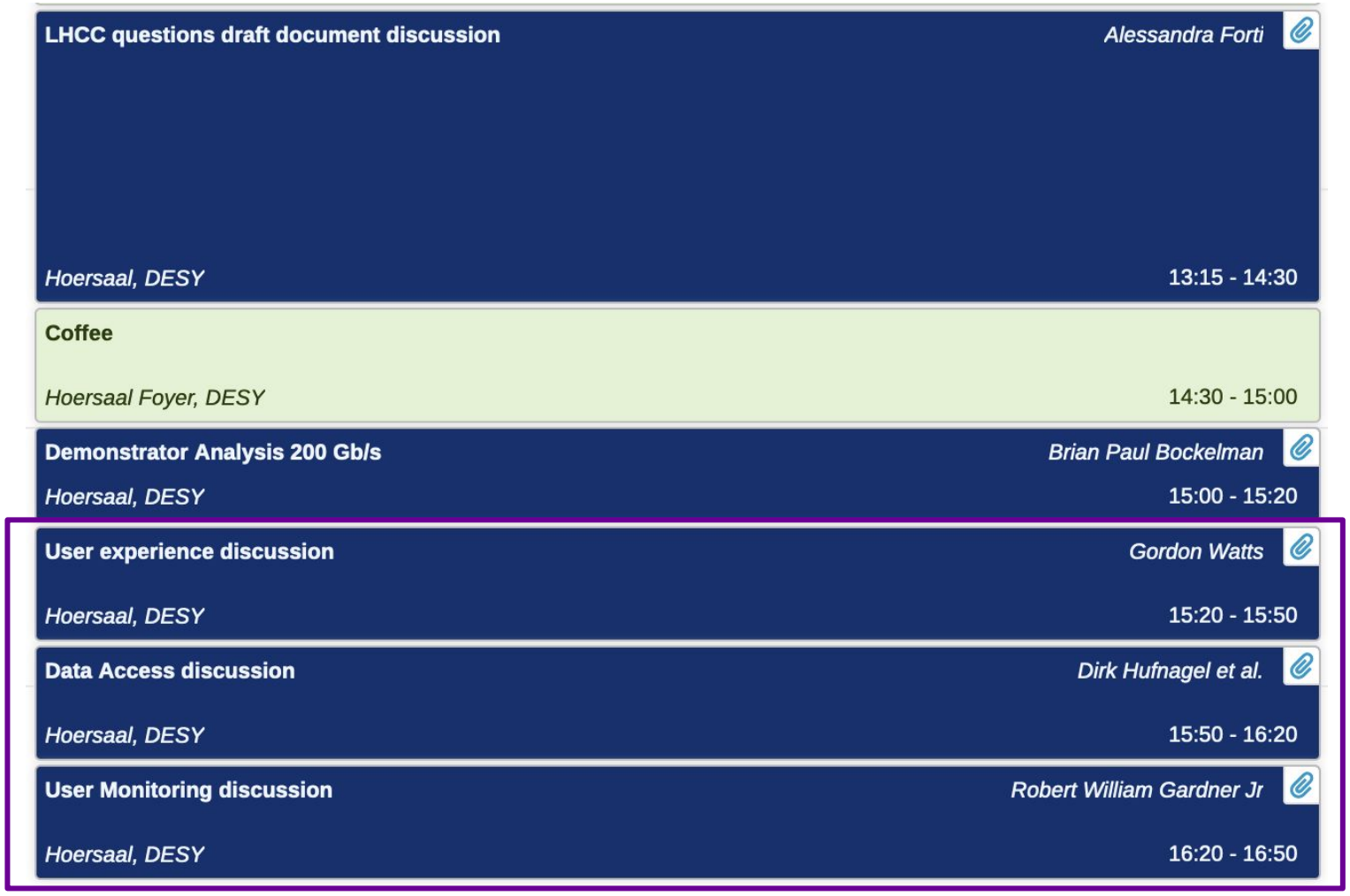
[b] *Editor,* [1] *INFN,* [2] *University of Manchester,* [3] *Technische Universität München,* [4] *University of Warwick* [5] *University of Washington* [6] *Brookhaven National Laboratory* [7] *Morgridge Institute for Research* [8] *University of Chicago* [9] *University of Oslo* [10] *Ludwigs-Maximilians Universität München* [11] *Princeton University* [12] *Syracuse University* [13] *University of Wisconsin-Madison* [14] *University of Oxford* [15] *Technische Universität Dortmund* [16] *FNAL* [17] *CIEMAT* [18] *IFIC* [19] *TIFR* [20] *Purdue University* [21] *University of Bristol* [22] *Rice University* [23] *Paul Scherrer Institute* [24] *Helsinki Institute of Physics* [25] *University of Michigan* [26] *University of Illinois at Urbana-Champaign* [27] *University of Massachusetts Amherst* [28] *IFAE* [29] *University of Liverpool* [30] *University of Birmingham* [31] *CERN* [32] *University of Nebraska-Lincoln* [33] *SCIPP UC Santa Cruz* [34] *University of Hamburg* [35] *Massachussets Institute of Technology* [36] *Karlsruhe Institute of Technology*

Abstract

This white paper presents the current status of the R&D for Analysis Facilities (AFs) and attempts to summarize the views on the future direction of these facilities. These views have been collected through the High Energy Physics (HEP) Software Foundation's (HSF) Analysis Facilities forum [1], established in March 2022, the Analysis Ecosystems II workshop [2], that took place in May 2022, and the WLCG/HSF pre-CHEP workshop [3], that took place in May 2023. The paper attempts to cover all the aspects of an analysis facility.

WLCG/HSF WS

- WLCG/HSF workshop at DESY in May
- Analysis Facilities common WLCG/HSF plenary

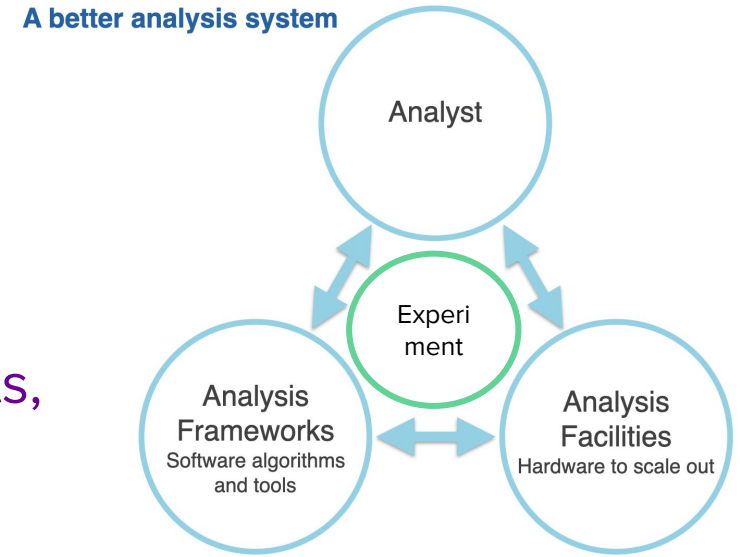


White Paper driven discussions

AF LHCC charge

Analysis Facilities is a broad topic. The use cases and the expectation of the experiments, and the analysers may evolve over time. The scope and the expected content of the document to be provided by the experiments must be clearly defined. To this end, a list of questions must be defined first that seek to define the expectations from experiments for Analysis Facilities. The questions, to be answered by experiments, must be picked such that the answers are useful for sites and are representative of a broad spectrum of analyses and analysers. The list of questions might evolve in the future.

- The questions have been chosen to
  - Represent a broad spectrum of analysis
  - Be useful to sites
  - Be answered by the experiments
  - May evolve
- A single list of questions
  - Unified separate lists from the experiments, reviewed with the community at the WLCG/HSF WS, finalised with comments
  - Experiments will answer separately



Questions structure

Run3
analysis model

Run4 (Run5)
analysis model

AF organisation,
support and resources

AF hardware
and technology

Analysis use cases and
benchmarks

AF current
and planned R&D

- 6 question with a number of sub-questions each divided in 3 categories
  - Analysis model(s) to understand the evolution
    - Run3 and then Run4 (Run5) to highlight the differences
  - What would these AFs look like?
    - AF organisation resources, support, technology and hardware
  - How is the experiment going to get there and measure the benefits
    - Analysis, use cases, benchmarks and R&D
- Ball is in the LHCC court now