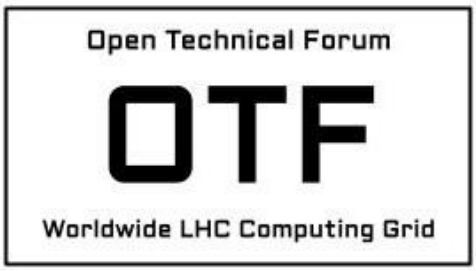
Live Notes for the Open Technical Forum #4 on Tuesday, April 1, 2025 at 13:37 (CET)

Indico: https://indico.cern.ch/event/1532650/

OTF#4 is co-hosted at the HEPiX Workshop in Lugano on the topic of mid-to-long-term evolution of facilities: https://indico.cern.ch/event/1477299/sessions/591035/#20250401

Welcome to the WLCG Open Technical Forum (OTF).



The WLCG Technical Coordination Board (TCB) is responsible for the technical evolution of WLCG services in line with the needs of the experiments and the capabilities of the infrastructure providers. The TCB defines a multi-year roadmap for such evolution and is responsible for its implementation. The TCB achieves these goals with a bottom-up approach through an Open Technical Forum (OTF) which welcomes the participation of all contributors to the technical evolution in the WLCG community.

Attending: ~90 in the room, 23 on Zoom

TABLE OF CONTENTS:

Introduction

Q&A

- [Questions inaudible from Zoom] Technical coordination in WLCG means listening to the experiments and their requirements, listening to the facilities, and writing down the agreed-upon plans, providing a forum where requirements can be discussed and agreed upon.
- [Encouragement from Jeff] : given limited (person) resources at sites, any coordination between experiments is highly appreciated and increases our effective capacity as a community. For example, not having experiment-specific Analysis Facility infrastructures. OR, having some basic functionality required for setting up AFs in userland, where sites support the common basic functionality and all experiment-specific stuff is not done by “site people” but by “experiment people”. This might even enable Alessandro’s “100 analysis facilities”.

  Also : the same thing could be useful from sites - not having experiments having to deal with 5 different types of AAI infrastructure.

Storage challenges (DOMA + capacity vs performance)

Speaker: Shawn McKee

Summary of the main points: There have been two WLCG data challenges so far, in 2021 and 2024. Significantly improved monitoring was a positive outcome of the first challenge. Issues to be followed up after DC24 with Rucio, FTS, monitoring, IAM, determining target rates (re-evaluate realistic computing models), and networking R&D (packet marking, SDNs). Token refresh rate (IAM) needs to be agreed upon between the experiments. Mini challenges help us to explore capability evolution, recently in the UK and the US and a CMS test for jumbo frames in transatlantic data transfers.

Q&A

- [Comment inaudible on Zoom] Some capability challenges require new monitoring, and will drive the creation of new monitoring.
- Challenges based on the new LHC schedule. There are varying opinions within DOMA about the target rates, whether to re-evaluate, etc.

HEPiX Technology Watch Working Group Report

Speaker: Andrea Sciabà (start at 14:12)

Summary of the main points: Technology for scientific computing, in HEP for online and offline use cases. Predictions of cost evolution and experiment requirements between now and Run 4. Recent evolution of CPU and GPU architectures and models. GPU is not evolving in a useful way for HEP. DRAM is not scaling any more. Concerns about HDD performance evolution. Networking standard evolution. Fabrication roadmap until 2036. Environmental sustainability.

Q&A

- [Question inaudible on Zoom] Someone is doing something at CERN but we don't have details.
- [Question inaudible on Zoom] Vendors understand our needs but they do not influence their decisions.
- [Comment inaudible on Zoom, sounded like Pepe] There are bigger entities that need tape other than us. Doesn't seem to be a risk for tape in the next few years.
- [Question inaudible on Zoom] Even optical storage is far from eternal.
- [Question inaudible on Zoom]
- [Question inaudible on Zoom] If we are more united, can we get better deals on hardware purchases? Improving the purchase process, but each site faces a different market.
- The main discussion was about "how can we compete with the "big 5"". We will need to ask AndreaS to write this down.

Italy Vision

Speaker: Daniele Spiga

Summary of the main points: INFN distributed infrastructure, for LHC (15% of the total globally) but 20% of the capacity is dedicated to non-LHC work (10 domains in total). Era changes of networking (Tbps, nationally), heterogeneity, and external resources (new HPC in Napoli). Creation of an Italian data lake. Motivations are meeting requirements, easing operations, and providing high-level solutions (R&D).Transparent site extension of CNAF to Leonardo HPC with 1.6 Tbps connection to CNAF storage, but no IPv6 support at this time. Direction of future EU "AI Factories" also in other countries e.g. BSC, LUMI. Directions of resource provisioning from

traditional Grid to cloud and k8s on bare metal. R&D focuses on connecting centers without outbound connectivity, cloud provisioning (?), **distributed storage (which will persist),** enabling smaller communities to use “Rucio as a service”, high level interfaces, and analysis facility demonstrations. In summary, abstracting the complexity of the underlying computing infrastructure to be logically simpler, integrate external resources (HPC), and be backbone for more than HEP.

Q&A

- What is the impact on the infrastructure direction on the Tier-2s? Italy is upgrading the Tier-2 infrastructure including networking.
- Do you foresee a unified Rucio federation? Will the Tier-2 storage be distributed or separately visible by the LHC experiments? There is an effort in INFN to coordinate a “data cloud” with a coordinated management and operation of distributed Tier-2 storage.
- [Question inaudible on Zoom] Resource location within the INFN context, there were several studies in WLCG, getting resources from the hyperscaler seems not a sustainable solution even at the funding level. Reusing and generalizing tools such as Rucio … Let’s take what we can get from cloud-based approaches.
- [Comment inaudible on Zoom, from Pepe] Comment perhaps about integrating external resources. Mitigation of risk factors regarding hardware and policies.
- Why not everything in the cloud? We need to retain expertise.

Coffee Break - 30m

Schedule is running 30m late.

IDAF @ DESY: Interdisciplinary Data and Analysis Facility: Status and Plans

Speaker: Yves Kemp (start at 15:50)

Summary of the main points: IDAF is a data-centric computing infrastructure. NAF focuses on interactive analysis with fast turnaround, user storage, and **user support**.

Q&A

- [Question inaudible]
- [Question inaudible] If jobs go by official tools, Rucio identifies the data set, but if users have their own framework, the users must know the datasets are in the storage.
- [Question inaudible]
- What is the level of effort needed for user support? Is it proportional to the amount of storage that each community has? Only logarithmically. Difficult to quantify the amount of effort dedicated to this. Looking for a technical solution to the user support problem will not work, in their opinion.

German University Tier-2s evolution

Speaker: Michael Boehler

Transition to ramp up Tier-2 CPU to 3 HPCs (Aachen, Gottingen, and Karlsruhe), and concentrate Tier-2 mass storage at 3 Helmholz centers (DESY, KIT, GSI). Helmholz centers will still provide significant CPU. Tier-2 groups will supervise the HPC resources. Performance studies of scalability of processing. Investigating how to improve access the more centralized storage at the Helmholz centers with different approaches including caches at the HPCs and Federated dCache.

Q&A

- Have you analyzed ARC caches? Not yet.
- How will the HPCs evolve over time, with relation to the projections? Note that the plots are based on the current HPC capacities. No procurement numbers at the moment.
- On slide 10, [Question inaudible]... Are there special privileges needed to run the xrootd caching proxy? No, but they were granted access.
- Federated dCache, how much is this based on pre-existing work?
- The Karlsruhe group has good relationships with the HPC.

Evolution of US ATLAS Sites

Speaker: Ofer Rind

Summary of the main points: The US ATLAS Cloud comprises a Tier-1 at BNL, 5 multi-site Tier-2 sites, and 2 HPCs (NERSC and TACC), and 3 analysis infrastructures (not a focus of this presentation), all with high-speed network connectivity and storage. NET2 has been using OpenShift/OKD for site management to try to reduce administration effort needs, after an initial investment to set up the site. Accounting needed some R&D, container-native k8s accounting called Kuantifier. The NE Tier-2 also has a 50 PB tape system which will be used to serve user analysis. Studying optimal data placement and workflow management in a heterogeneous environment (REDWOOD project). Facility R&D efforts have mostly focused on analysis use cases, including identity management with IAM, ServiceX (with IRIS-HEP), bursting from an AF into Tier-2 resources, and SDNs. Benchmarking analysis infrastructure regularly.

Q&A

- To what extent does the influence of the partnerships with external groups like IRIS-HEP affect goals? Likely affected US CMS more than US ATLAS, a lot of synergy with groups like OSG and IRIS-HEP, on Coffae-Casa for example. There is a lot of alignment between the analyzers and the external groups more so than with the experiments. There are regular meetings on the AGC. There is convergence of technologies and analysis interests now, greatly improved from a few years ago.

- Any differences seen running with or without HTCondor on a k8s substrate? NET2 is in production and runs well, no major issues with this.
- [Room lost audio connection for ~30s]

"Round table" - Facilities in WLCG Technical Roadmap

(Started at 17:07)

- Going back to the scope of the session, fostering a strong connection between WLCG and the facilities. A lot of R&D projects to understand analysis. Analysis use case is not well-framed and user support is a key pillar.
- We heard from Italy and Germany about distributing or concentrating compute and storage, and integrating HPC resources.
    - What about Spain?
        - BSC and their FA agreed that they are a strategic project and access a portion of the HPC, with proposals for allocations made every 4 months, for 50% of the LHC computing needs, and run simulation there. The 50% is based on the needs of the experiments to run simulation. Motivation was to save money for the HEP program. Challenge was to integrate MareNostrum HPC with no external network connectivity.
        - There are multiple risks, large other projects competing for resources, and the risk that our frameworks need to evolve in order to run on new infrastructures.
        - User support at MareNostrum – there is some support but no flexibility. They are getting more flexible as they see examples from other countries. Different countries have different access policies even though these are all EuroHPC resources, and we need to lobby EuroHPC to move to have more uniform access policies.
    - What about the UK?
        - There is almost no HPC capacity in the UK. Future UK investment will be in AI and GPU by the FAs. The cheapest way to do WLCG computing is to buy hardware that matches the job requirements and put it on the Grid. You don't need Infiniband, large shared memory, etc., though there may be good political reasons to use HPCs.
        - No plan yet, pending understanding the funding situation. They need written plans from WLCG. Quantify requirements, including for GPUs.
    - The Netherlands has a flat-funding model. Improvements have to come from performance improvements. How much can we expect over the next 5 years?
- GPUs: Someone comments that WLCG probably won't need GPUs on the Grid.
- Security (J.C. Luna, WLCG Security Officer):
    - There are two main aspects that are worrying, or implications of changes: tokens and AFs, that are not very well-defined. We need to understand more details of both so the security implications can be understood.

- Education and workshops.

Meeting time: 4:02

Executive Summary

WLCG held OTF#4 hosted at the HEPiX Workshop in Lugano on Tuesday, April 1st. The scope of the session was to deepen the strong connection between WLCG and the facilities and explore the main issues about facilities that should be addressed in the WLCG Technical Roadmap. We learned that many of the R&D efforts ongoing are focused on understanding analysis use cases including user support needs as a key pillar. We heard from Italy, Germany, Spain, and US ATLAS about plans for distributing or concentrating compute and storage and integrating HPC resources. The UK does not have HPC capacity at this time but would find concrete requirements from the experiments useful to help argue for funding. US ATLAS has moved in the direction of using OpenShift/OKD for more automated site management, enabling job submission through K8S API, as well as deploying tape to a Tier-2, and considering the possibility to serve also user analysis. There are many efforts to optimize data placement and data and workflow management on heterogeneous and distributed architectures. Proliferation of differing interfaces and access policies at HPC facilities is seen as a challenge, the nature of the HPCs (constantly pushing the technologies limits) implies the need to foresee expertise in our community to integrate them. Security implications of new infrastructures and authentication methods were discussed.