Report on the meeting with the EuroHPC Federation Platform, and perspectives

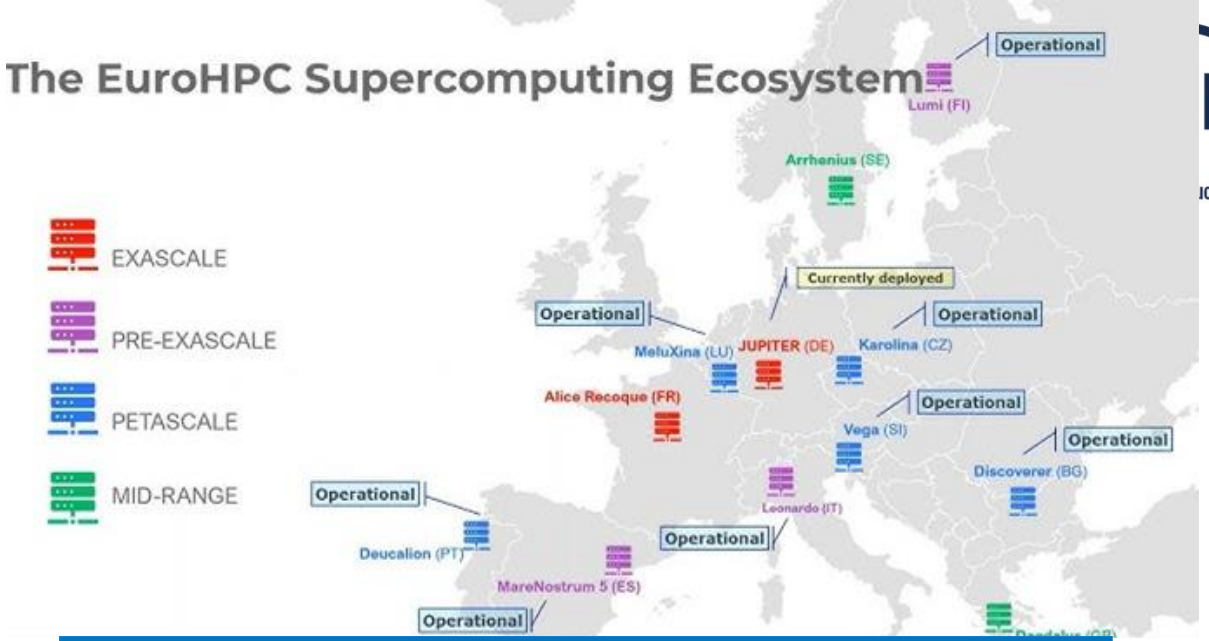
Tommaso Boccali (INFN Pisa)
with Openlab – Maria Girone

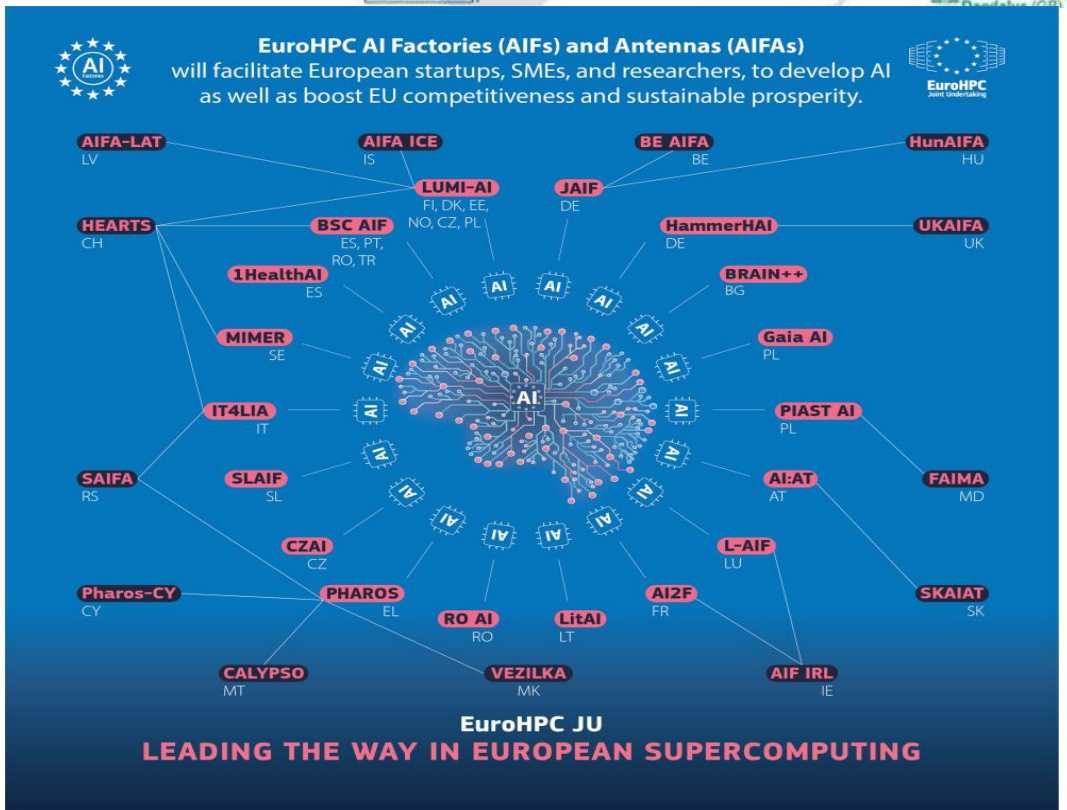
Outline and prior history

- What is the EuroHPC Federation Platform, and why it is!
- Previous contacts with CERN/WLCG
- The technical handshaking meeting, September 25th
- What now
  - EFP short and medium term plans
  - What/Where we can realistically do

The EuroHPC Federation Platform

- Europe HPC landscape circa 2025:
  - 2 Exascale systems (DE: operational, FR: 2026+) – around 1000 Petaflops
  - 3 Pre-exascale (IT, ES, FI) – means 200+ Petaflops
  - A number of Petascale systems (LU, CZ, BG, PT, SI) – means 10+ Petaflops
- Cofunded by the EuroHPC JU and local governments (~50%/50%)
- 2026+: the AI factories come online
  - 19 AI factories, 13 Antennas (“affiliates”)
  - All Exaflops level – in AI terms





How to access the systems

- Political:
  - For the quota owned by the EuroHC JU, access via projects
  - For the local coutry quota, “local contacts”
- Technical:
  - Very few dictates from the JU, every center has its own policies. Examples:
    - AAI solution
    - 2FA needed yes/no
    - Open/closed networking
    - CVMFS yes/no
    - Users allowed to instantiate “services” yes/no
    - Data mover solutions (if any)

- Benchmarking access
  - Devel access
  - Regular access
  - Exterme scale
  - Strategic
- Emergency Access
- Access for AI
- (see here)

A huge problem for (our) experiments: needs a different handshaking per center

Solutions?

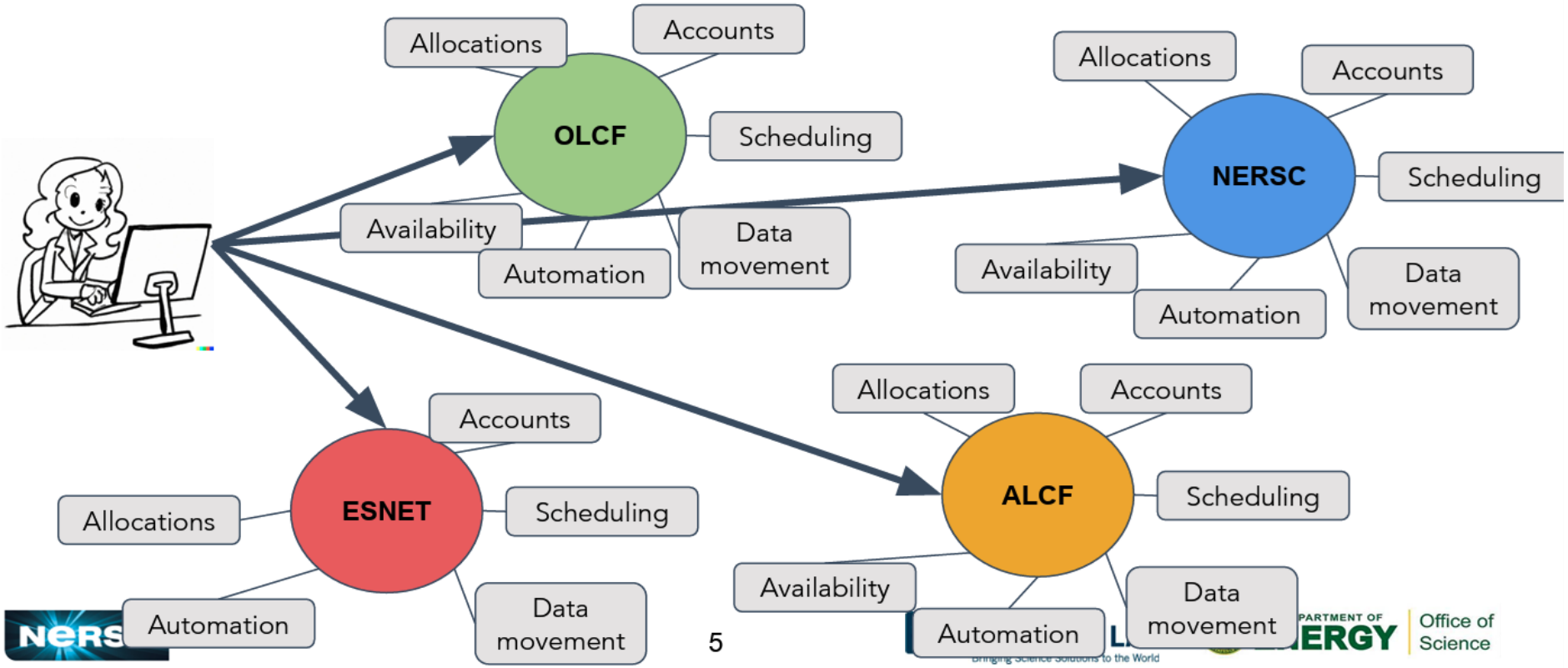
- The problem of fragmentation is perceived also at the funders level, but solutions are not obvious since the centers are mostly independent from the technical point of view
- The problem is indeed global; very similar situation in the US
- Attempts ongoing:
  - EU: The EuroHPC Federation Platform (20MEur grant with Finland/CSC and main beneficiary)
  - US: DOE Integrated Research Infrastructure (IRI)

Next MB?

- Quite different in scope, funding and reference communities, use cases

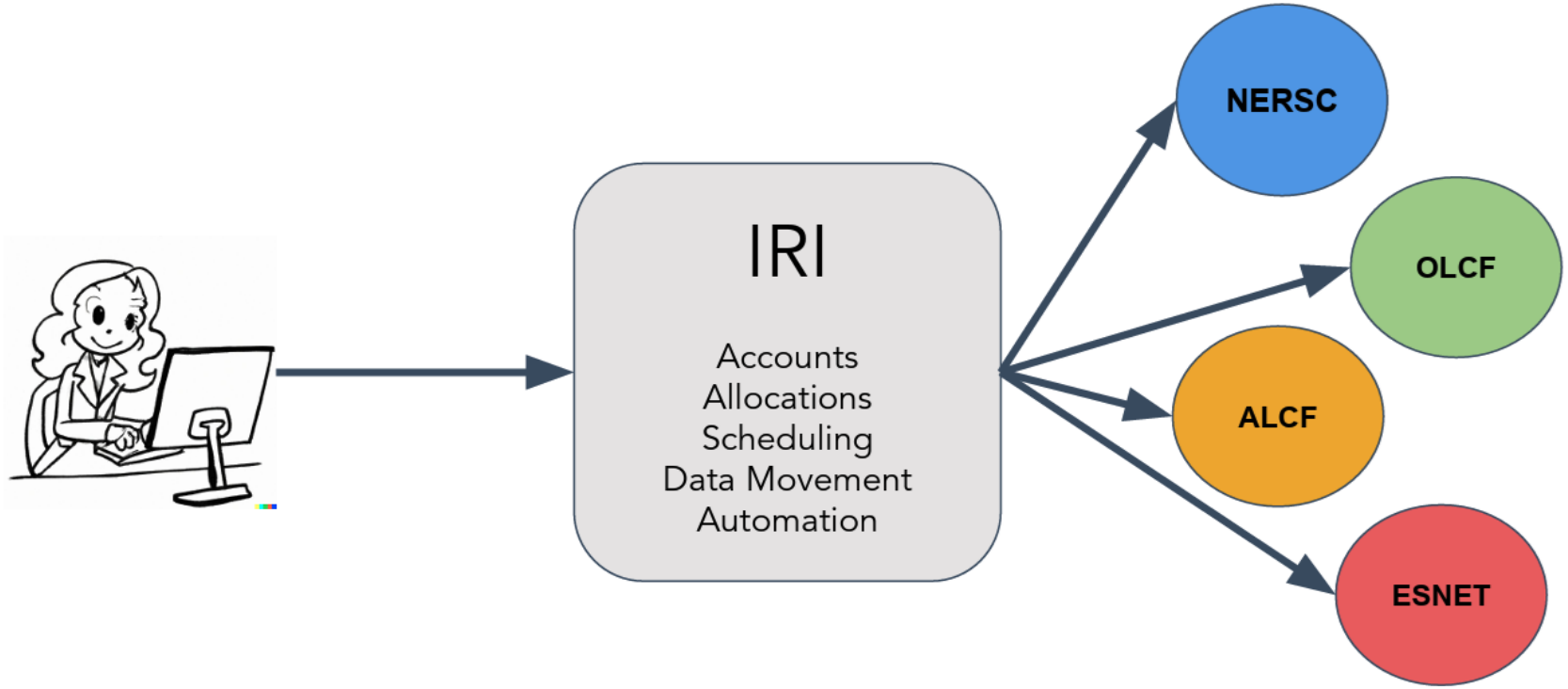
In a nutshell … from

In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible



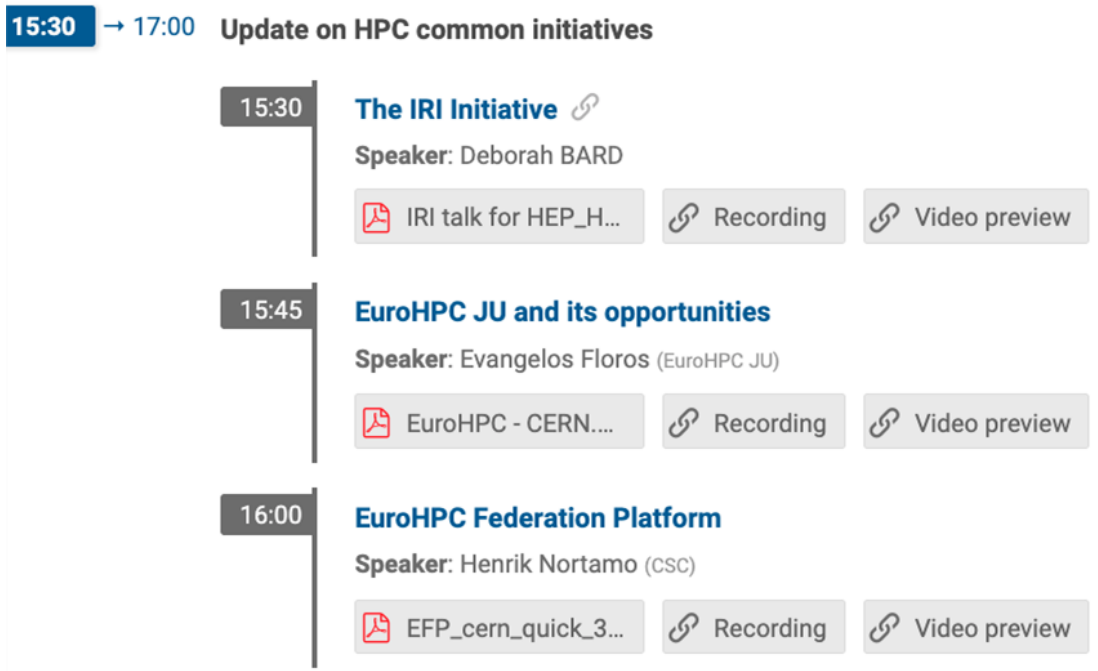
In a nutshell … to

In an Integrated Research Infrastructure, simple things should be simple, and complex things should be possible



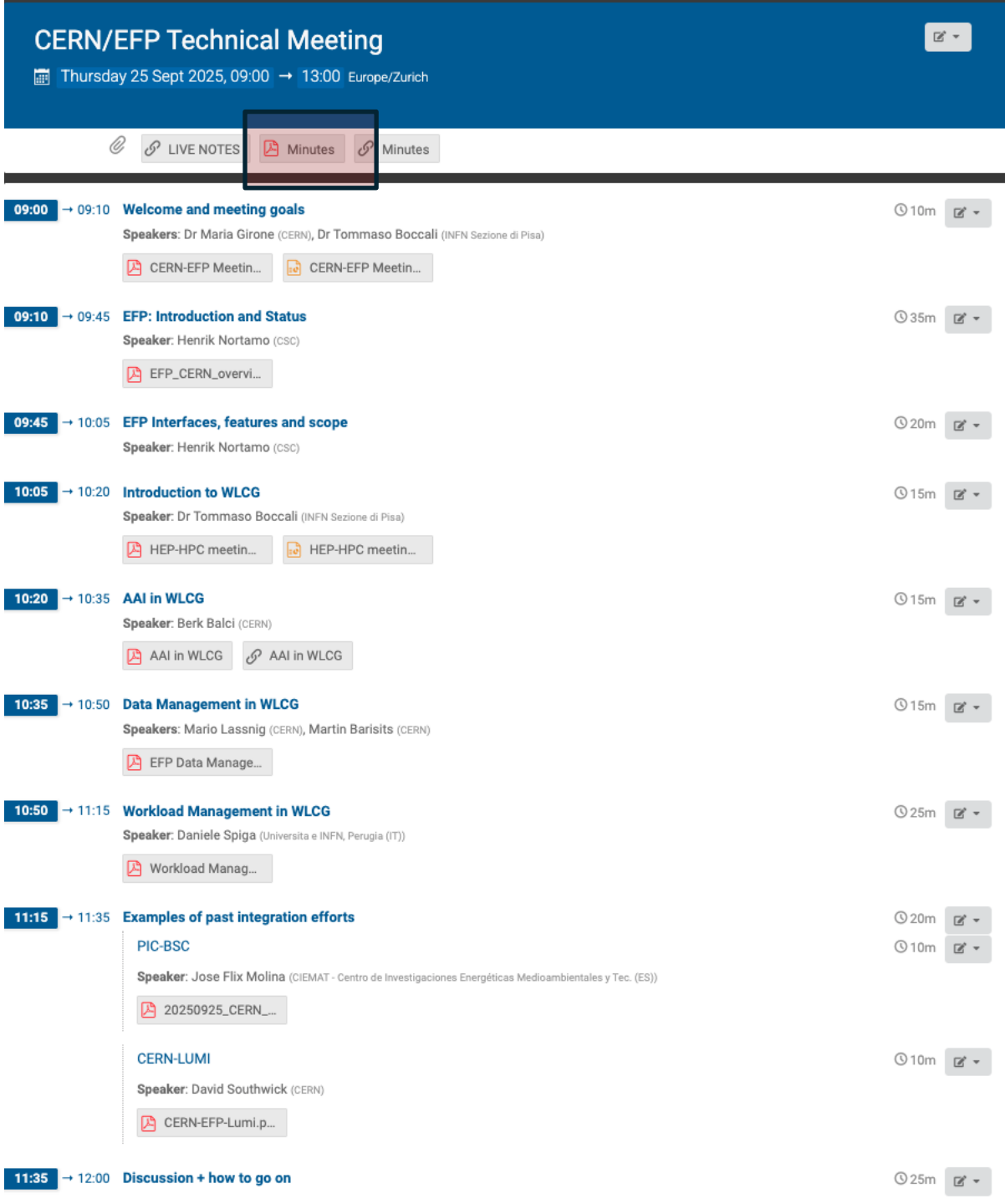
Contacts [all co-coordinated with Openlab]

- We had IRI and EFP talks at the HEP/HPC strategy meeting early 2025
- With EFP, we had some discussions + planned a technical meeting for handshaking
- We need to get back into contact with IRI
- September 25th → the CERN/EPF/WLCG technical meeting



Agenda and goals

- The main goals are getting to a better mutual understanding
- On our side:
  - Understand the solutions EFP has chosen
  - Understand the timeline (Phase1, Phase2, …)
  - Understand the capabilities for adding use cases, requests
  - Understand the opoprtunities for early tests etc
- On their side:
  - Undestand what we need but also what we can offer (as use cases, expertise, …)



I will not discuss our contributions here (many thanks to all the WLCG speakers!)

- EFP targets an “uniform” interface on all the systems funded by EuroHPC
  - Covering in principle the “EuroHPC quota” but by extension also the country quota
- The platform will provide
  - Unified solution for users to manage their access and resources across all EuroHPC systems.
    - E.g., managing all the grants, limited interest for us
- Unified software offering across the systems
  - This translates to some standard SW being distributed via EESSI/CVMFS
- Advanced graphical interface
  - No interest for us?
- Features for power users accustomed with direct access to systems and APIs
  - This is what we want .. See later

EFP timeline

- Timeline is divided into a first phase (2025-2026)
  - where a **Minimum Viable Platform** will be deployed, including federated AAI, resource allocations, and complex workflow manager, User interfaces and helpdesk
  - The **first release** is expected Q1/2026, supporting all currently online 9 EuroHPC systems, and will enable user testing;
  - A **second release** in Q4/2026 will include more systems and features.
  - The second phase (2027-2029) will include enhancements and support more systems.
- The feature set for the first phase is frozen, since long (the project submission?)
- For the first phase, if it happens a scheduled, realistically we can be beta testers, but we cannot ask for development

- As soon as first release is out, EFP would be happy to consider new devels / requirements

- **Important point: EFP has no strong mandate to be prescriptive towards the centers. This limits a lot the level of uniformity we can expect**
- *"The EFP does not impose technical changes / features on the HPC centers, which remain ultimately in control of the resources. Services like CVMFS, access to edge services, remote SLURM API interfaces can be available via the EFP, but only if a center decides to expose the service."*

The main building blocks

- Access:
  - Access via Portal to get an **ssh key**. Its lifetime to be negotiated with the centers
  - Third party integration (**for example with Indigo-IAM**) is scheduled for the seconf phase
    - Optimistic?
  - **Chain/robot accounts disfavoured** ("one account / one passport") – again a site decision
    - It was like this with X509 years ago …
  - **WLCG JWTs are in principle compatible with EFP solution**, modulo a needed (?) token translation service
- Sw distribution
  - **EESSI uses CVMFS**, but the enabling of other repos must be negotiated (especially if it has implications on much larger caches)
  - Some centers have still "worries" on CVMFS, but they will sync a list of repos "daily or hourly".
- DM:
  - Most (?) sites going to **expose an S3** on top on the Posix FS [separate areas or not? Sandboxing?]
    - Worries on our side on S3's performance
  - An alternative solution could be to bypass completely EFP DM, and **deploy edge services** on nodes which directly see the Posix. Again a site decision
- WMS:
  - Direct access (and then SLURM) via **ssh always possible** (a-la Bosco?) possible; also a **Vacuum-like** mechanis with self starting pilots
  - **API access** possible depends on the site
  - **Virtualization generally accepted** by sites –at least apptainer

The main components

- Edge Services
  - Most of the centers have **cloud systems "close" to the HPC** (in some cases, sharing Posix partitions). If available, they could be instantiated / steered via EFP
  - They would be a general solution to most of our "problems"
    - DM as a user task, bypassing the center DM (if not adequate)
    - Edge WMS nodes to get/distribute/steer processes
  - Some centers are **already allowing this**
- Extermal access
  - Today some HPCs do not allow for external IP addressing from the nodes (see PIC's experience)
    - Very close to a "no-go"
  - Optimism on the fact that **AI "is our friend here"**: sites are committed to host AI workflows, which depend on external training data.

Two “extreme” integrations experiences

- Pepe [PIC/BSC]
  - The major obstacle is the **absence of (outgoing) networking** from compute nodes, and no CVMFS
  - This has triggered **targeted devels** (condor, tunnels, fat containers, …)
  - Input data **must be preplaced** via an especially deployed transfer service via PIC
  - The result is an FTE need that **would not be sustainable** if scaled to all the systems
- David [CERN-IT/LUMI-G]
  - No CVMFS: various solutions attempted
    - Complicated by nodes w/o disk
  - Integrated CERN/SSO with eduGAIN
  - File transfer using ssh (nodes have connectivity)
  - Object storage accessed via s3cmd, rclone, …
- Other tests with LUMI reported by the experiments

How to go on?

- It is quite clear there is no space for new requirements for the first phase; discussions should continue to have eventually some of the features added for the second phase (discussions after Q1/2026)
- We expressed the intention to be beta testers already in phase 1
  - Clearly it depends on what the experiments / the MB think – is it worth?
  - A beta testing group for EFP has not been defined yet – there is space
- In view of more interesting possibilities for the second phase, we would like to prepare a mailing list with contacts to exchange info and prepare.
  - Can the experiments/centers tell us (WLCG-Office) about their interest in participating, and give names?
- Caveat: there is no user-ready documentation on EFP side, but it will be there by the time of the release.
  - HEP is asking for access to it as soon as available, together with the inclusion with the beta test users