

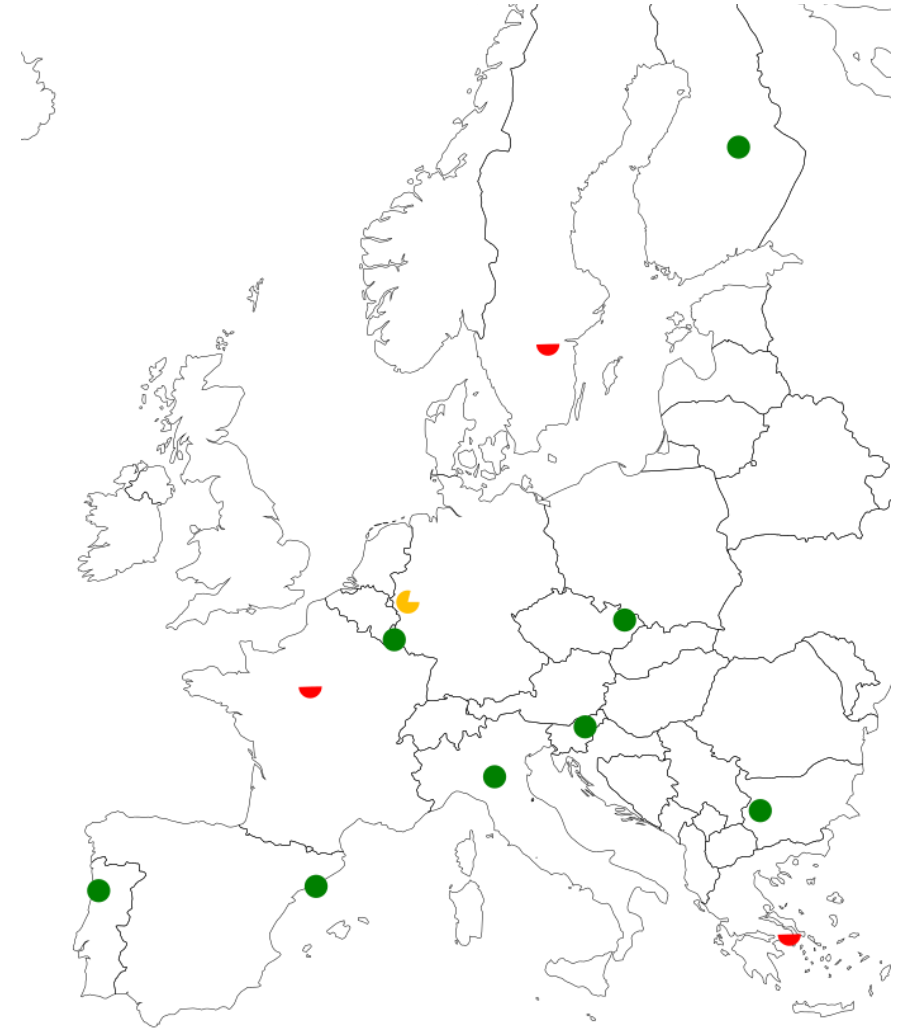
EuroHPC Federation Platform

HEP/HPC STRATEGY MEETING - ALL REGIONS, 30.01.2025

HENRIK NORTAMO, CSC - IT CENTER FOR SCIENCE

Federation targets for the platform

- Classic supercomputers, shown on map
- AI Factories , these are supercomputing infrastructures with AI-optimized computing capabilities.
- Quantum computers
- Connection to various dataspace, e.g. EOSC and SIMPL



Current Issues

- End-users acquire fully separate accounts and projects/allocations through completely different processes for each system.
 - Steps like initial user identification must also be re-done for each system
- "New" tools and workflows which require a lot of manual work to get working on the HPC systems.
- Growing heterogeneity of both compute hardware and environments
- Compute is much easier to move than data

The federation platform

The consortium delivering the service

A platform federating the access to all EuroHPC systems, with the main features being:

- **Federated identity and Single-Sign-On (SSO).** Users utilize the same login and identity (e.g. granted via their home institution) to authenticate to all services and access all supercomputers.
- **Resource allocation, management and monitoring across systems.** Users can see what allocations they have on each system in a single place.
- **Direct access utilizing SSH certificates.** Short lived certificates which are obtained via a login flow with optional MFA.
- **Interactive web-based usage** with e.g., remote desktop, shell sessions and Jupiter notebooks. Ability to launch batch jobs and browse files on the supercomputers.
- **Federated software catalogue** providing a pre-installed pseudo-uniform software stack on all systems
- **Advanced workflows and data transfer.** Workflow execution and data transfers across systems with smart scheduling capabilities

NORDUnet
Nordic Gateway for Research & Education

GÉANT
Networks • Services • People

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

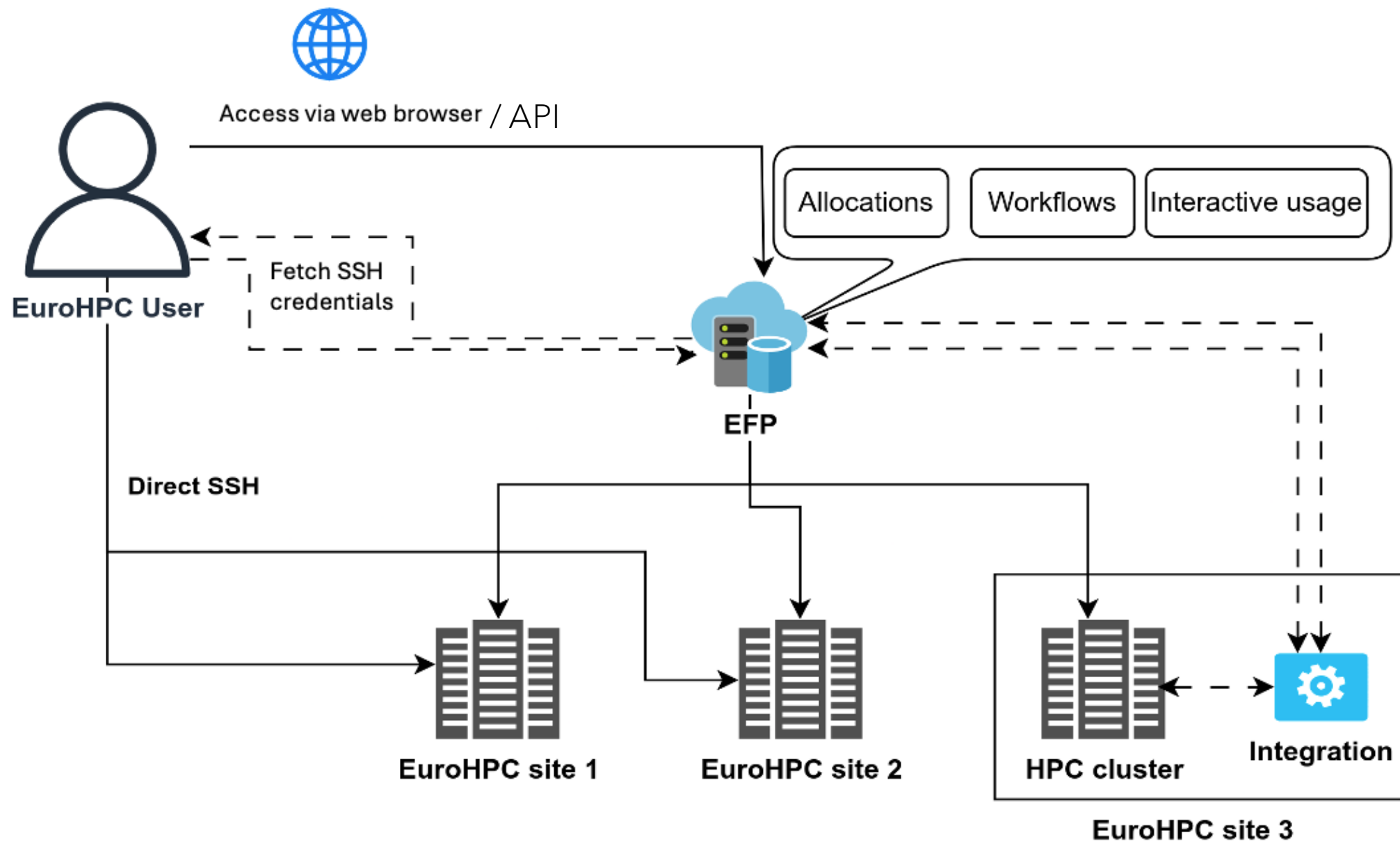


GHENT UNIVERSITY

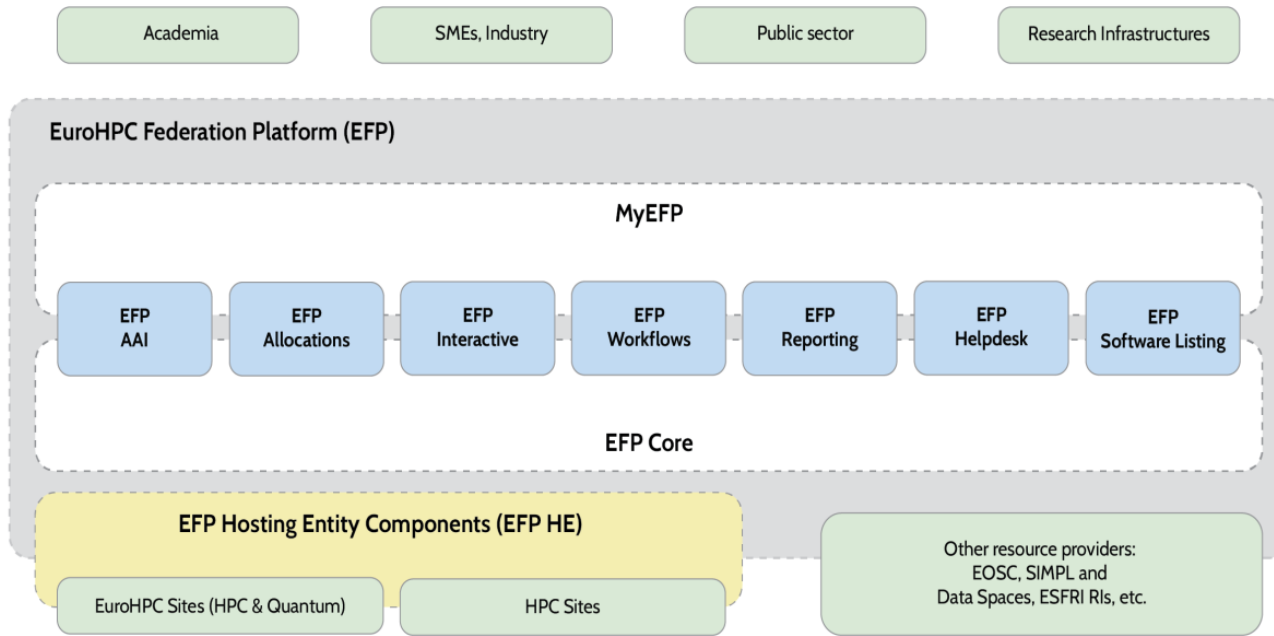


ICT Solutions for Brilliant Minds

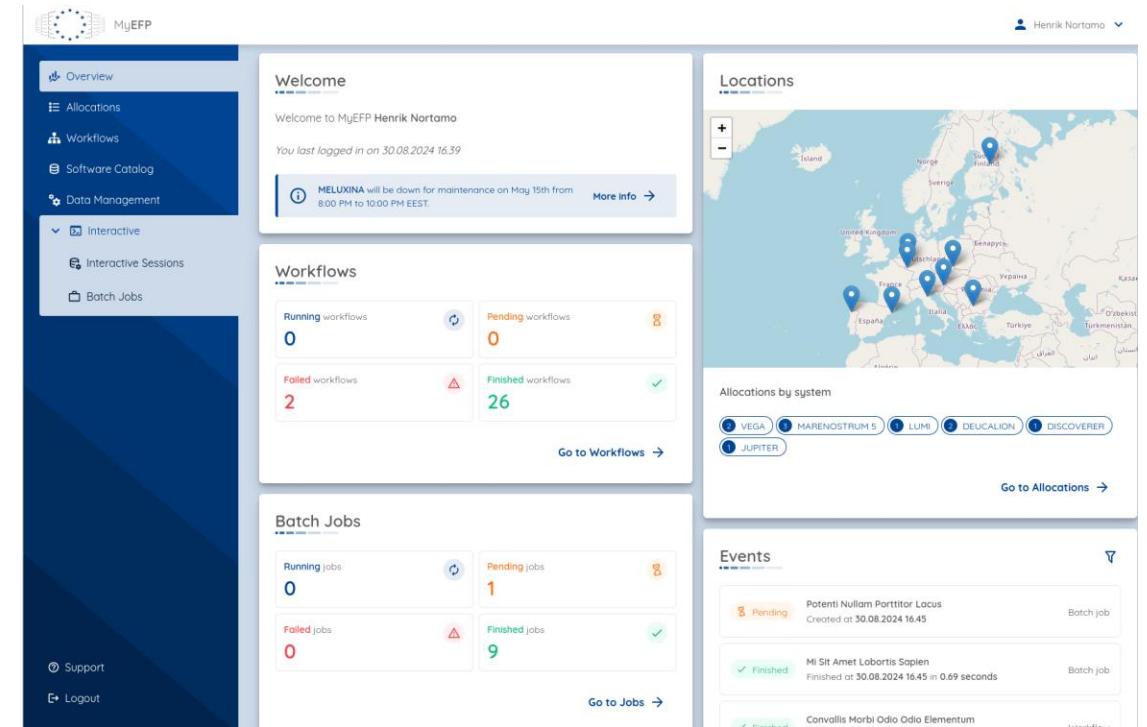
Architecture



Architecture



Modularity, Flexibility and API centric are among the core design principles, along with being as non-intrusive as possible.



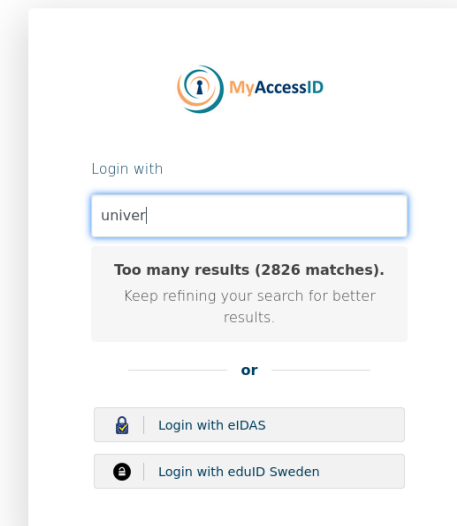
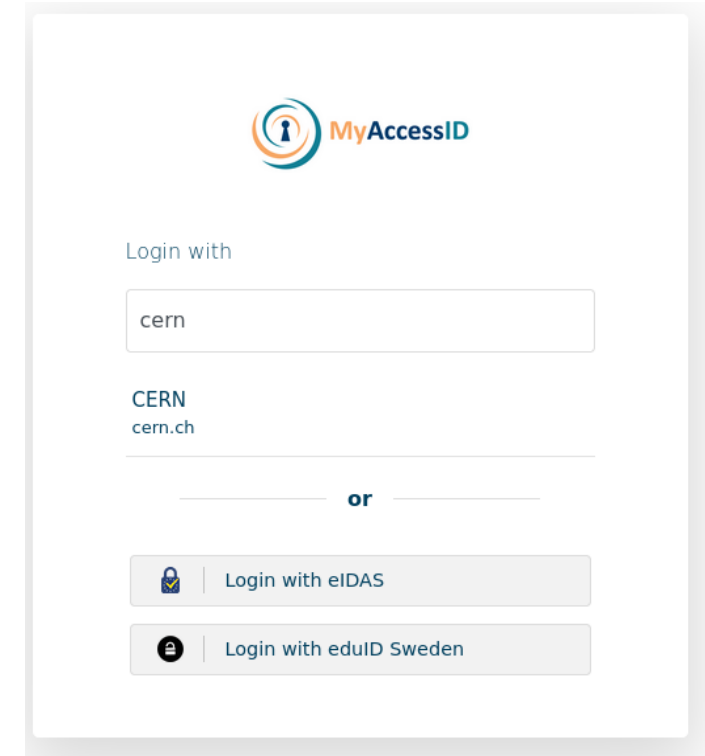
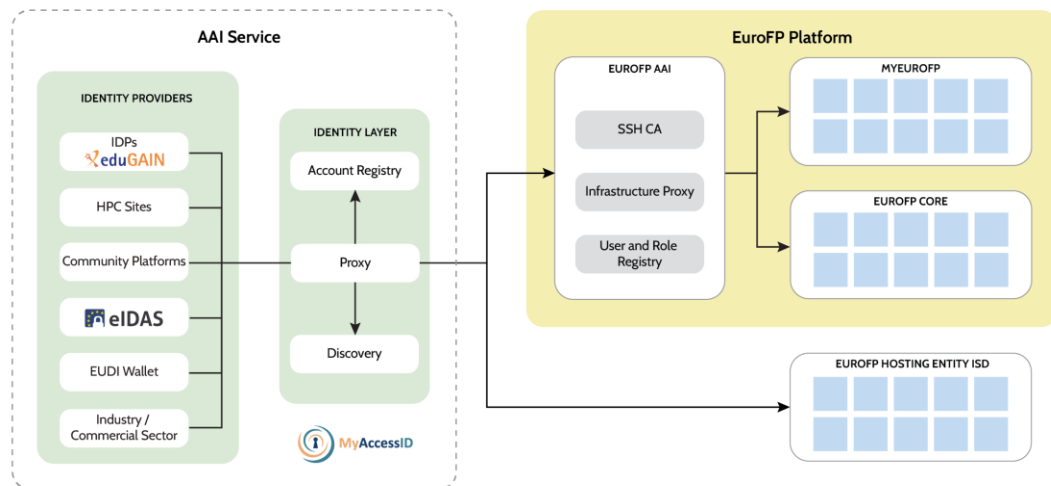
Overview of a users activities, coupled with UI:s for specific functionalities, all tightly integrated into a cohesive service.

Main components

- EuroFP AAI -> Leveraging [MyAccessID](#), SSH CA, Infrastructure Proxy
- EuroFP Allocations -> [Waldur](#) (same as Puhuri used on LUMI)
- EuroFP Interactive -> [Open OnDemand](#)
- EuroFP Workflows -> [LEXIS Platform](#), [HEAppE Middleware](#)
- EuroFP Reporting (also includes monitoring) -> [Grafana](#), [Icinga](#), [OpenSearch](#)
 - Waldur also shows some information on resource consumption
- EuroFP Helpdesk -> [Zammad](#)
- EuroFP Software Catalogue -> [EESSI](#), [EasyBuild](#) and [CernVM-FS](#)

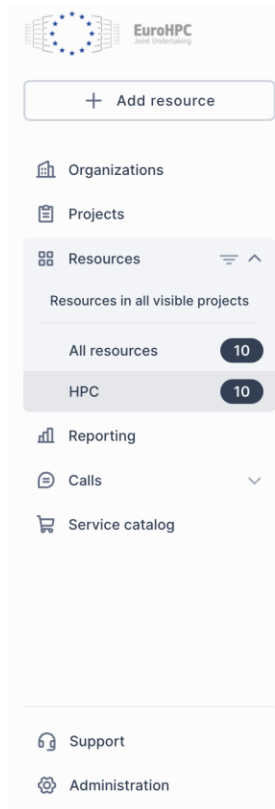
EFP AAI

- EFP Utilizes MyAccessID for identity federation, but does not provide it
- MyAccessID can be integrated as an OpenID Connect Provider or SAML Identity Provider
- Direct access to systems will be based on SSH Certificates
 - If you have used e.g. Leonardo, the process will be very similar. I.e. initiate a connection, authenticate via MyAccessID + IDP + possible MFA step, receive credentials
 - Once obtained, functionally equivalent to public - private key pair
 - Limited lifetime. Specific duration are a per-site policy
 - Allows for more restrictions, increasing security. Again per-site policy



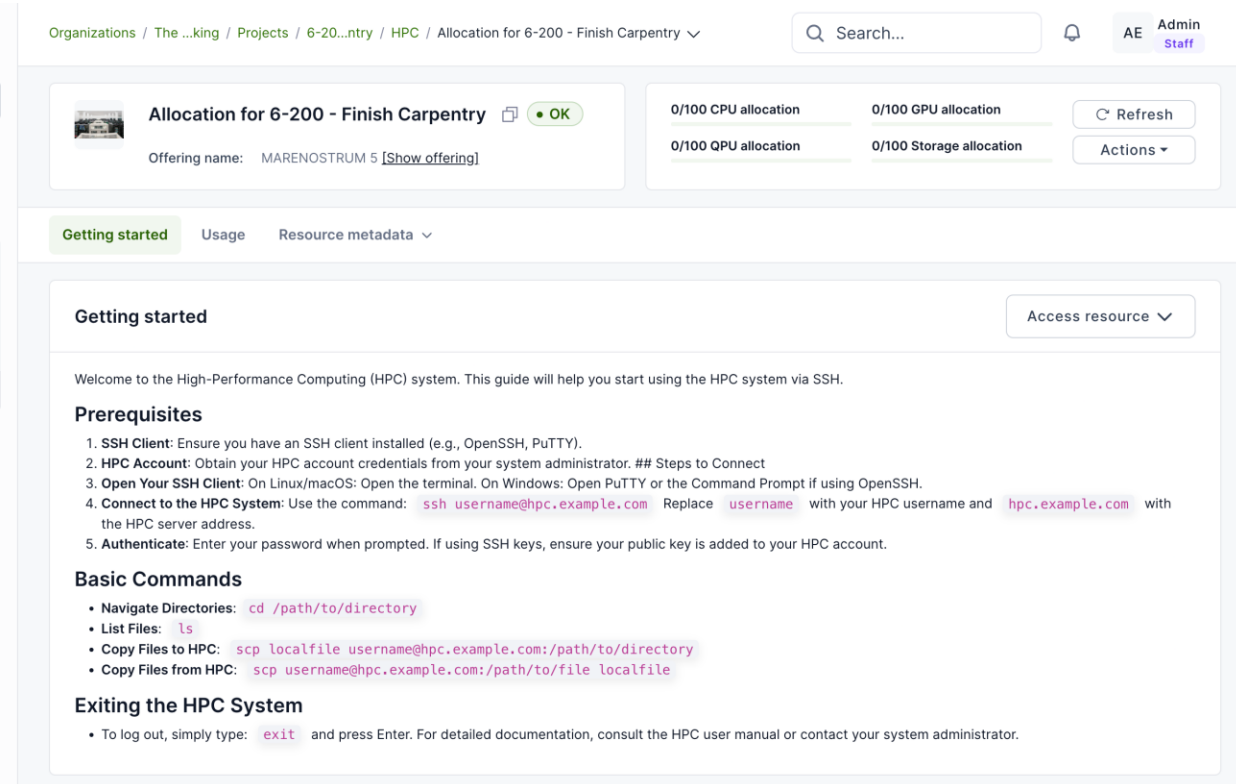
EFP Allocations

- Provides a unified allocation and project membership management capability
- Users and resource management portal provides project-centric dashboards.
- EuroFP Allocations allows project members to see and manage their allocations, request new ones either directly or via published calls for access.
- Support for projects having allocations on multiple systems



The sidebar navigation menu for EuroHPC includes the following items:

- + Add resource
- Organizations
- Projects
- Resources ⌵
 - Resources in all visible projects
 - All resources 10
 - HPC 10
- Reporting
- Calls ⌵
- Service catalog
- Support
- Administration



The screenshot shows the details for an allocation titled "Allocation for 6-200 - Finish Carpentry". The breadcrumb path is "Organizations / The ...king / Projects / 6-20...ntry / HPC / Allocation for 6-200 - Finish Carpentry". The offering name is "MARENOSTRUM 5". Resource usage is shown as 0/100 for CPU, GPU, QPU, and Storage. A "Getting started" section provides instructions for connecting to the HPC system via SSH, including prerequisites and basic commands.

Getting started

Welcome to the High-Performance Computing (HPC) system. This guide will help you start using the HPC system via SSH.

Prerequisites

1. **SSH Client:** Ensure you have an SSH client installed (e.g., OpenSSH, PuTTY).
2. **HPC Account:** Obtain your HPC account credentials from your system administrator. ## Steps to Connect
3. **Open Your SSH Client:** On Linux/macOS: Open the terminal. On Windows: Open PuTTY or the Command Prompt if using OpenSSH.
4. **Connect to the HPC System:** Use the command: `ssh username@hpc.example.com` Replace `username` with your HPC username and `hpc.example.com` with the HPC server address.
5. **Authenticate:** Enter your password when prompted. If using SSH keys, ensure your public key is added to your HPC account.

Basic Commands

- Navigate Directories: `cd /path/to/directory`
- List Files: `ls`
- Copy Files to HPC: `scp localfile username@hpc.example.com:/path/to/directory`
- Copy Files from HPC: `scp username@hpc.example.com:/path/to/file localfile`

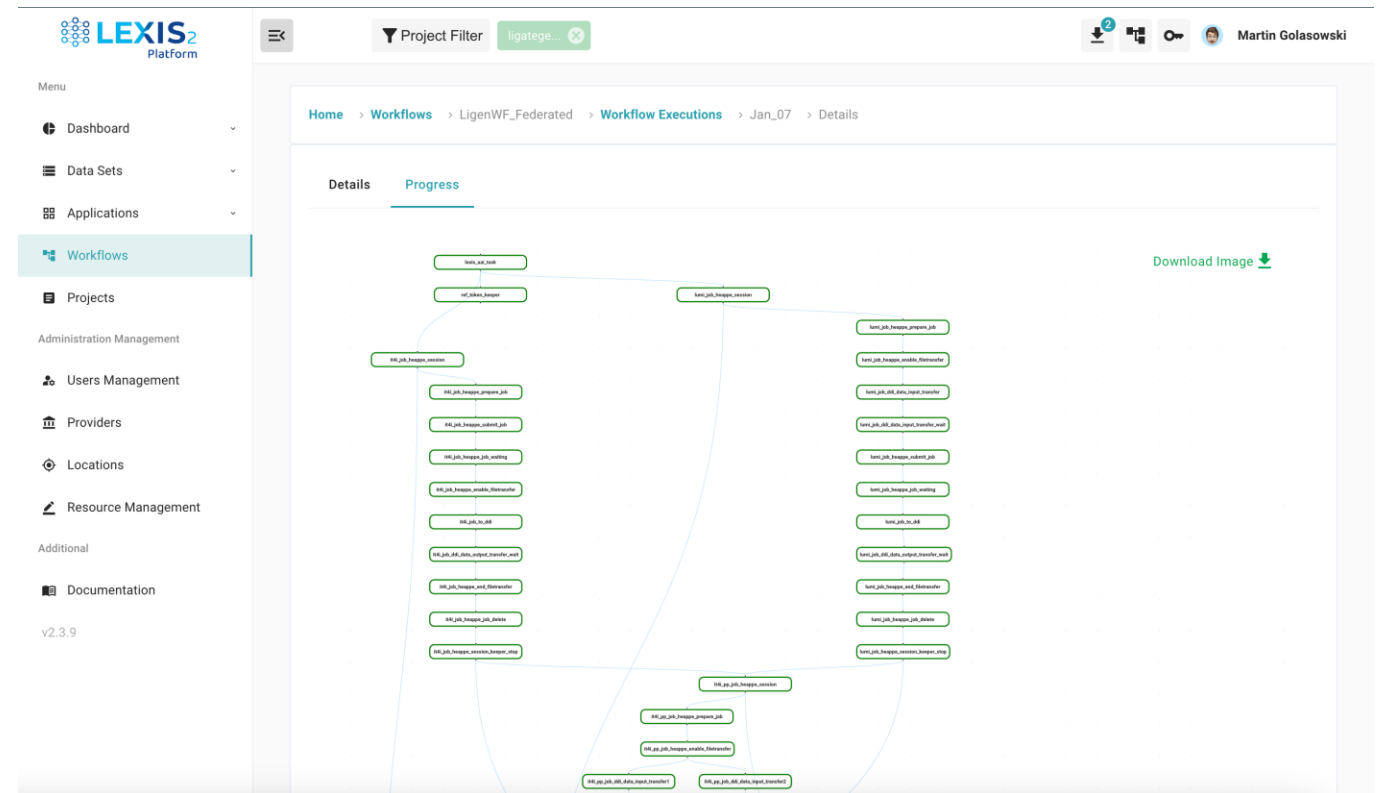
Exiting the HPC System

- To log out, simply type: `exit` and press Enter. For detailed documentation, consult the HPC user manual or contact your system administrator.

EFP Workflows



- Easy access to visually managed workflows and distributed data management across the federated resources.
 - Also allows API usage from various languages like Python or R.
- Smart scheduling policies across federated resources in workflow executions based on metrics from HPC sites.
 - Also available as a service
- Multi system workflows
 - Ability to combine allocations
- Able to schedule on both batch systems and e.g. Kubernetes
- Data staging from external sources

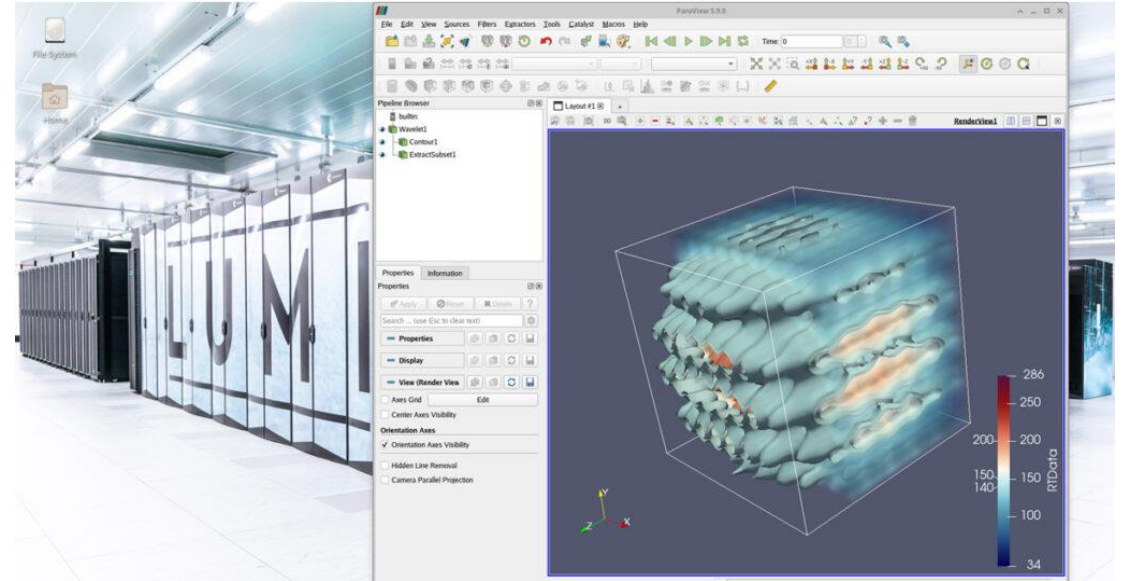


- HEAppE Middleware: REST API for restricted access to HPC infrastructure

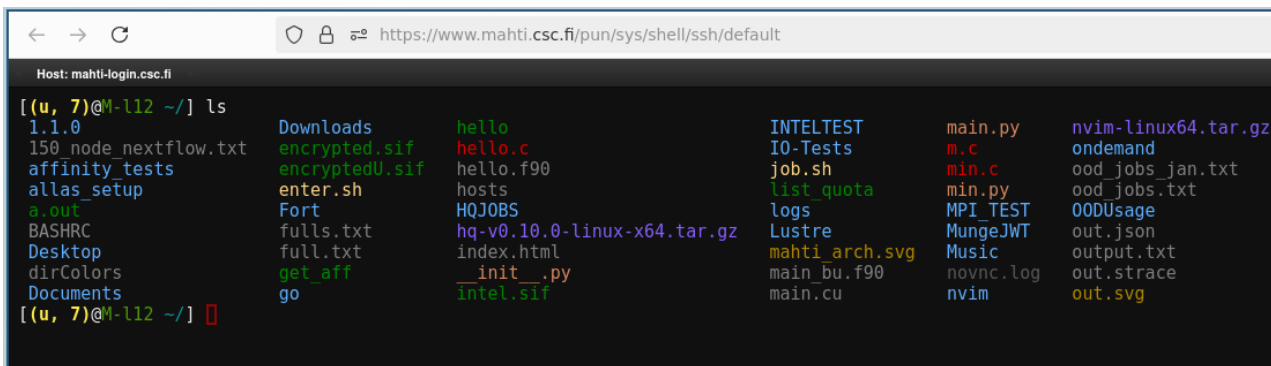
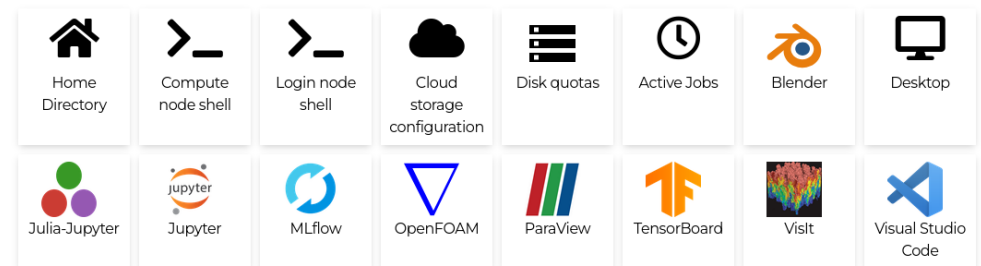
EFP Interactive



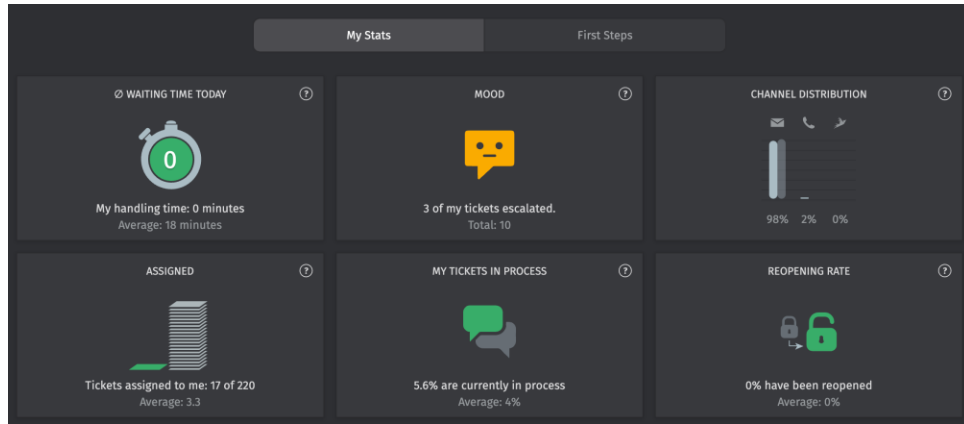
- Easy to use web interface
- Applications like Jupyter notebooks, remote desktops and shell
 - Native VNC support for high performance
- Fine-grained job level management and file management across the federated resources.
 - Accessing files on the HPC storage + smaller data uploads/downloads
 - Manage and view running batch jobs



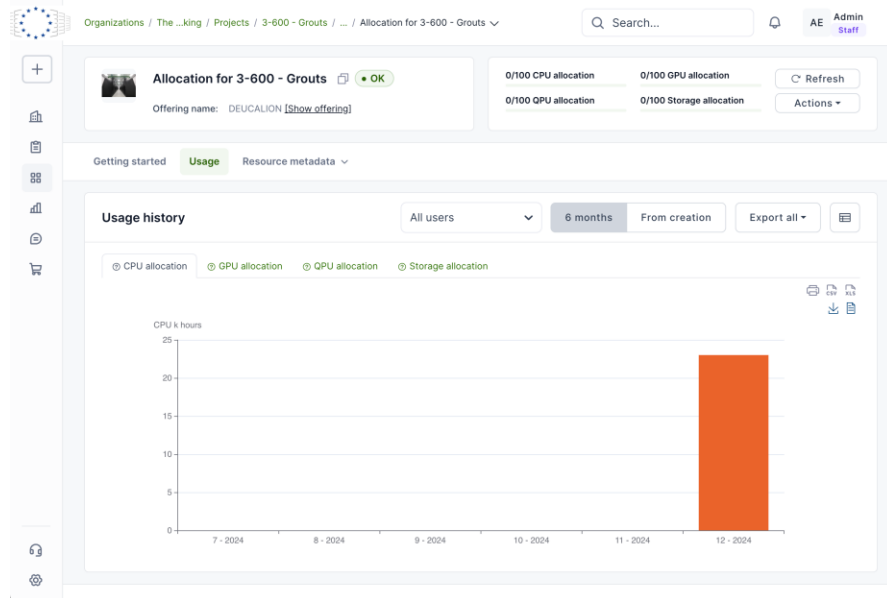
Pinned Apps



EFP Helpdesk

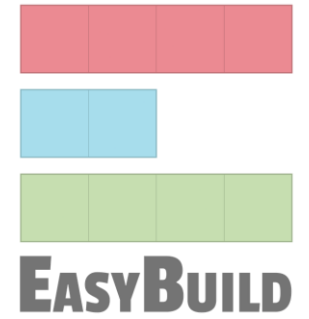


EFP Reporting



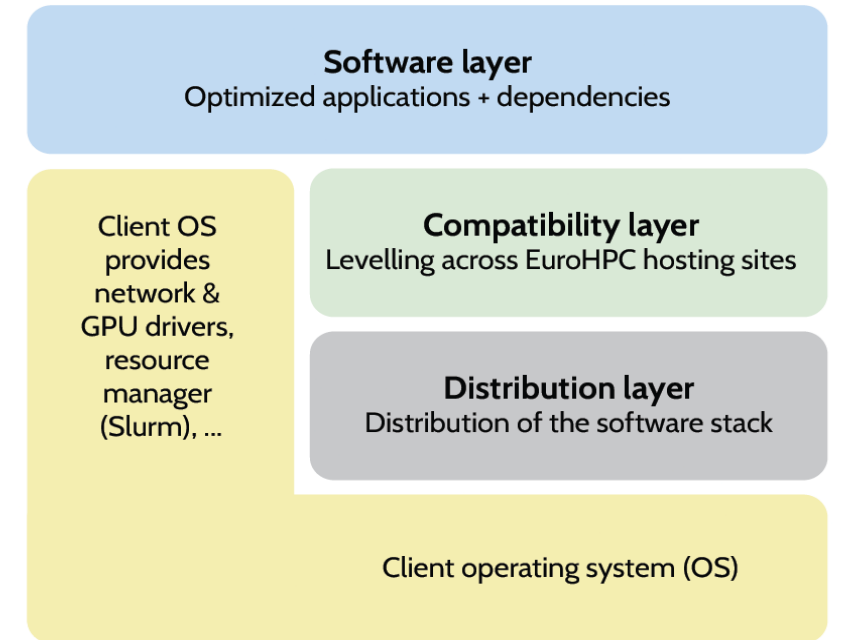
- Centralized reporting solution
- Resource consumption monitoring across different centers

EFP Software Catalogue



- Federated Software Catalogue based on EESSI
 - EESSI -> precompiled architecture optimized binary distributions of common scientific software
- EFP Software Listing provides list of the available software
- Ideally accessible on nodes via native installation of CernVM-FS, but alternative approaches are possible
- The EuroFP Federated Software Catalogue component will manage the heterogeneity of the different system architectures
 - So ideally, basic users would not need to care about the underlying GPU or CPU architecture.

Note: Different scale when compared to Cern's repositories
Stratum1: 120TB vs 300 GB, and much smaller pace of growth.



Timeline

- Work started January 2025
- Phase1 2025-2026
 - MVP released and available for users Q1/2026
 - MVP Includes the presented components excluding the federated software catalogue
- Phase2 2027-2029
 - More features for the components
 - More systems
 - License management
 - pre and post processing capabilities