



Under the cover An HPC-as-a-Service Architecture

Philippe Bricard

Budapest – 20/06/2022



“Mayor”



Client

Vendor

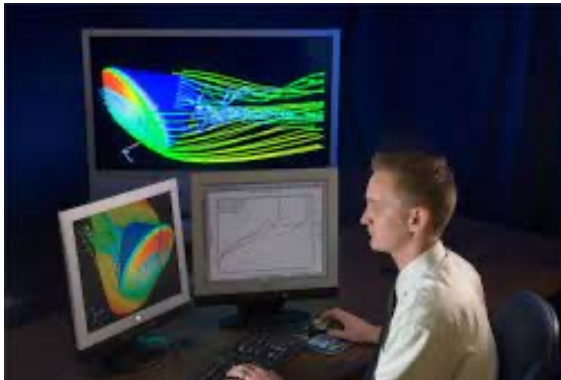
Choose products, negotiate prices
and leave the market with the
products

Provide their Products or
Resources



The HEROES Platform “Administrator”

“Mayor”



Users can bring their own workflow and if they agree on the terms, they launch the related jobs/steps



Client

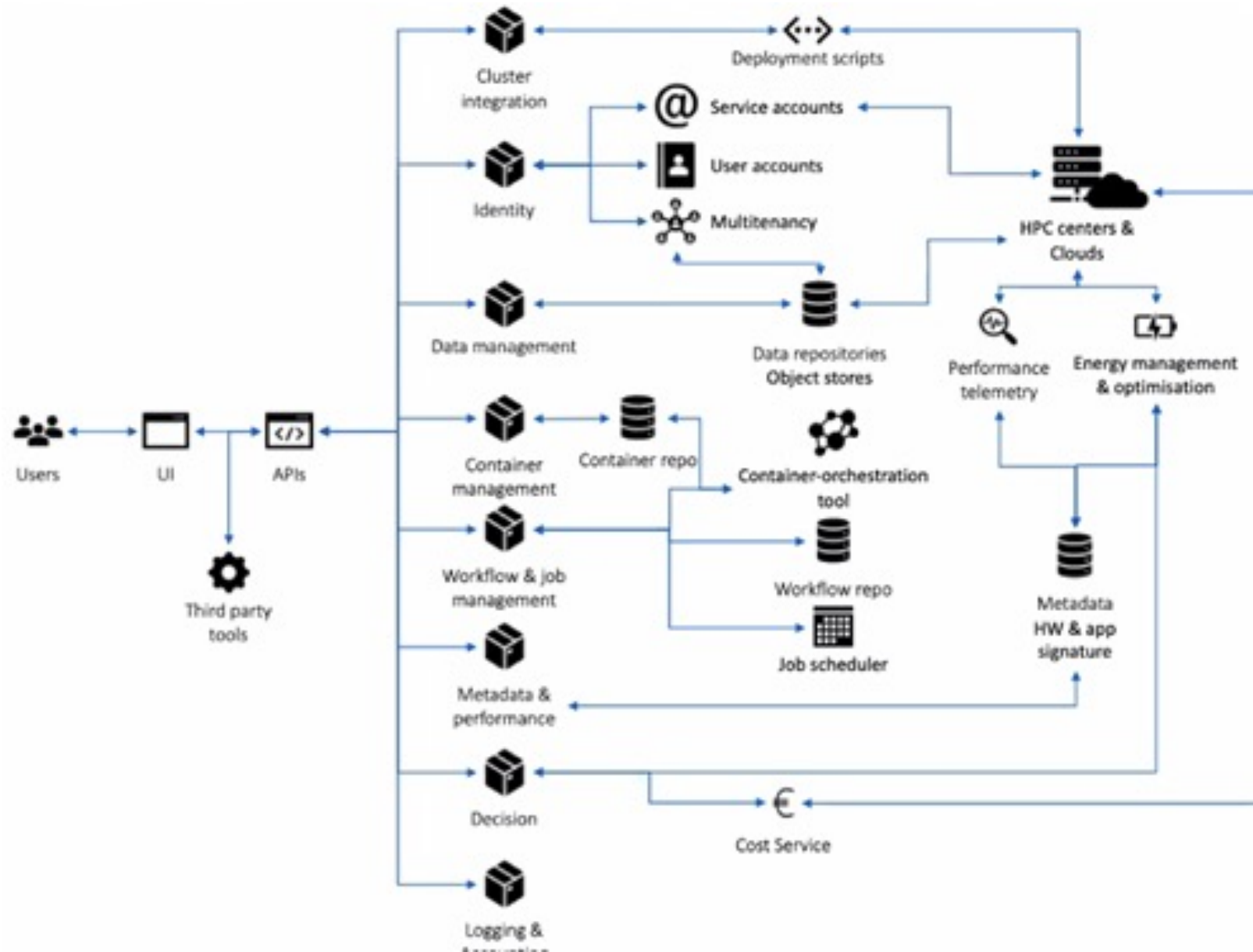
Vendor

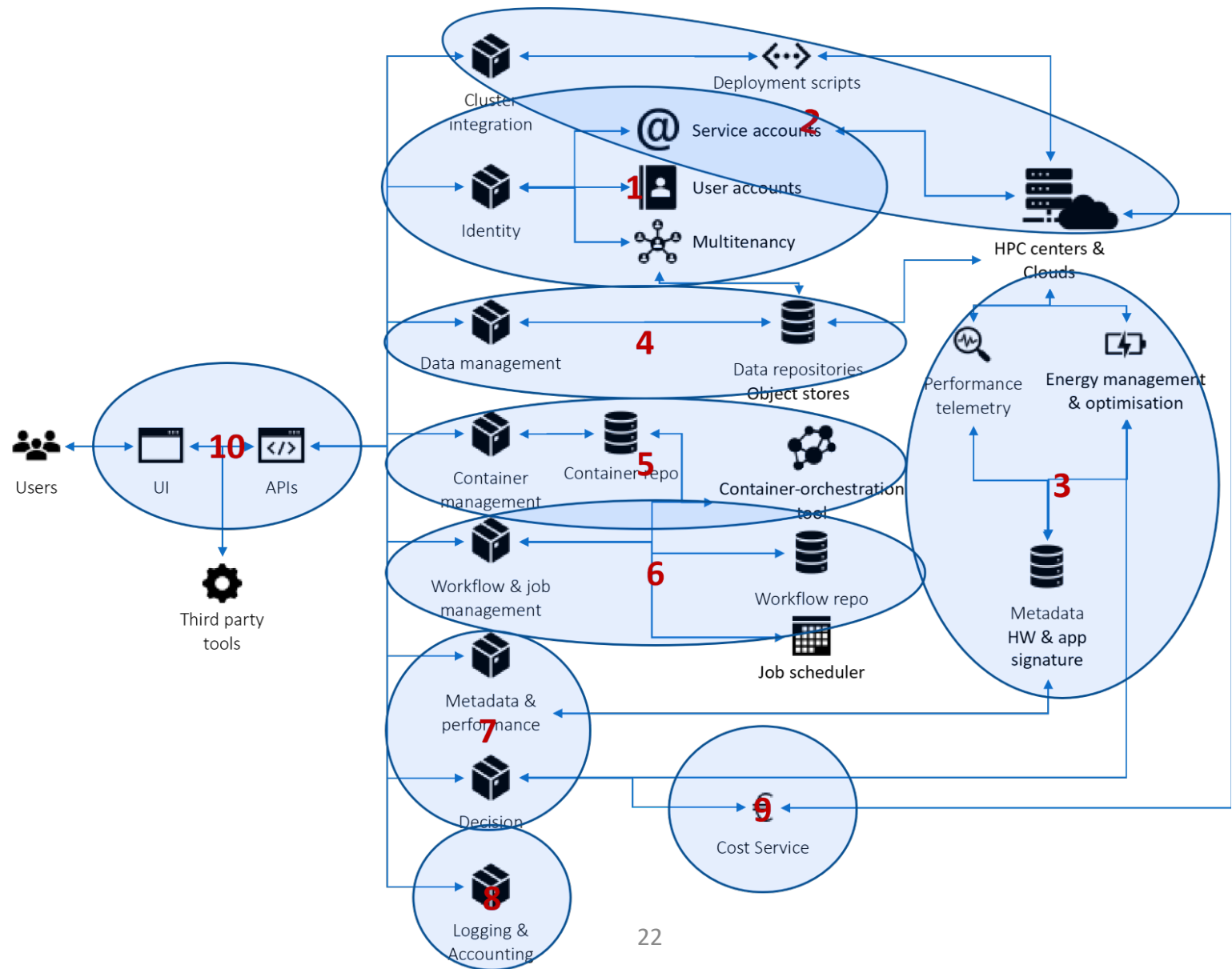


Provides HPC Compute & Storage at their own Terms and Conditions which can vary over time

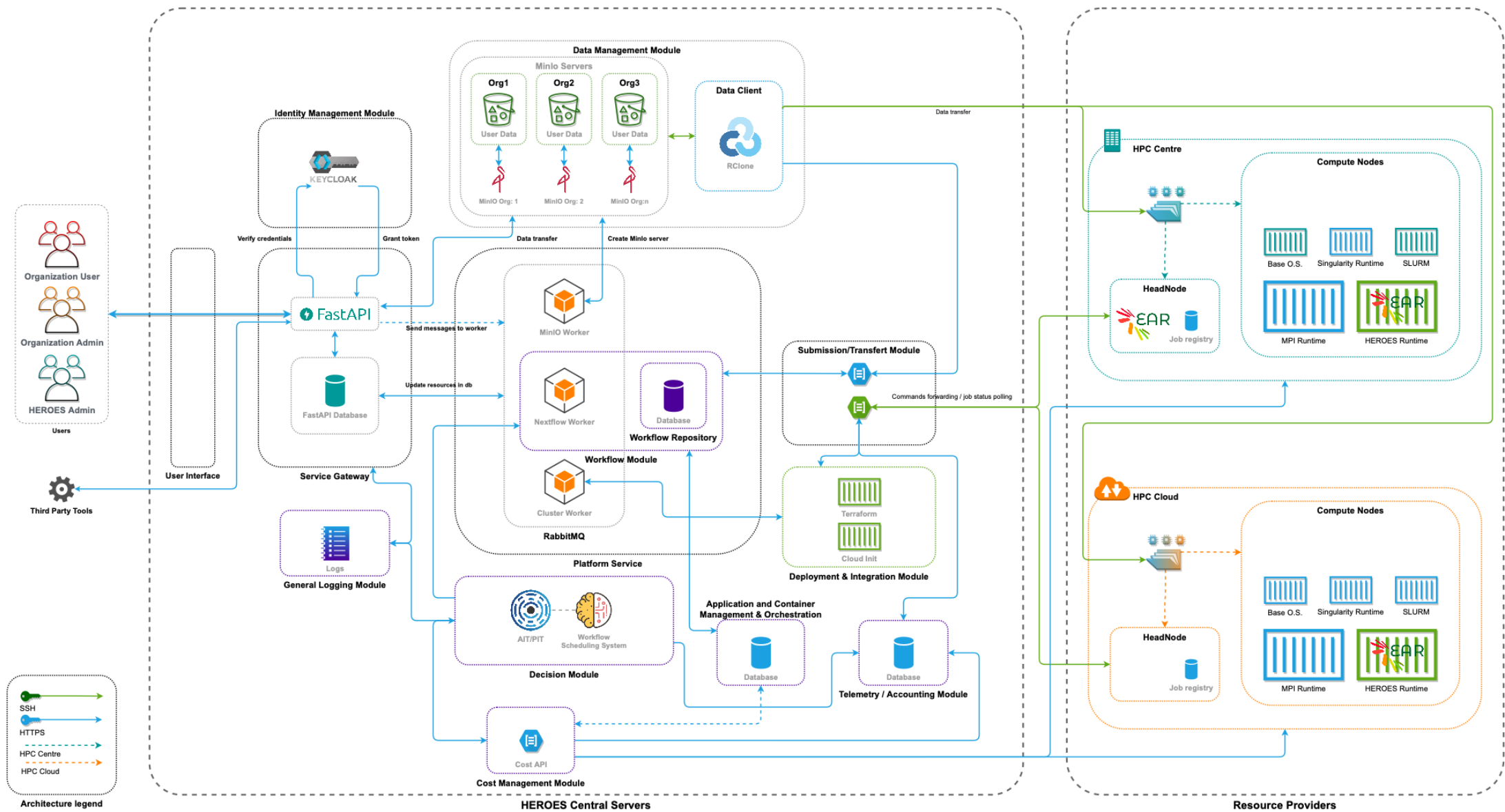
Some Vendors will provide HPC & AI Workflows

Summary of the work done 1/6

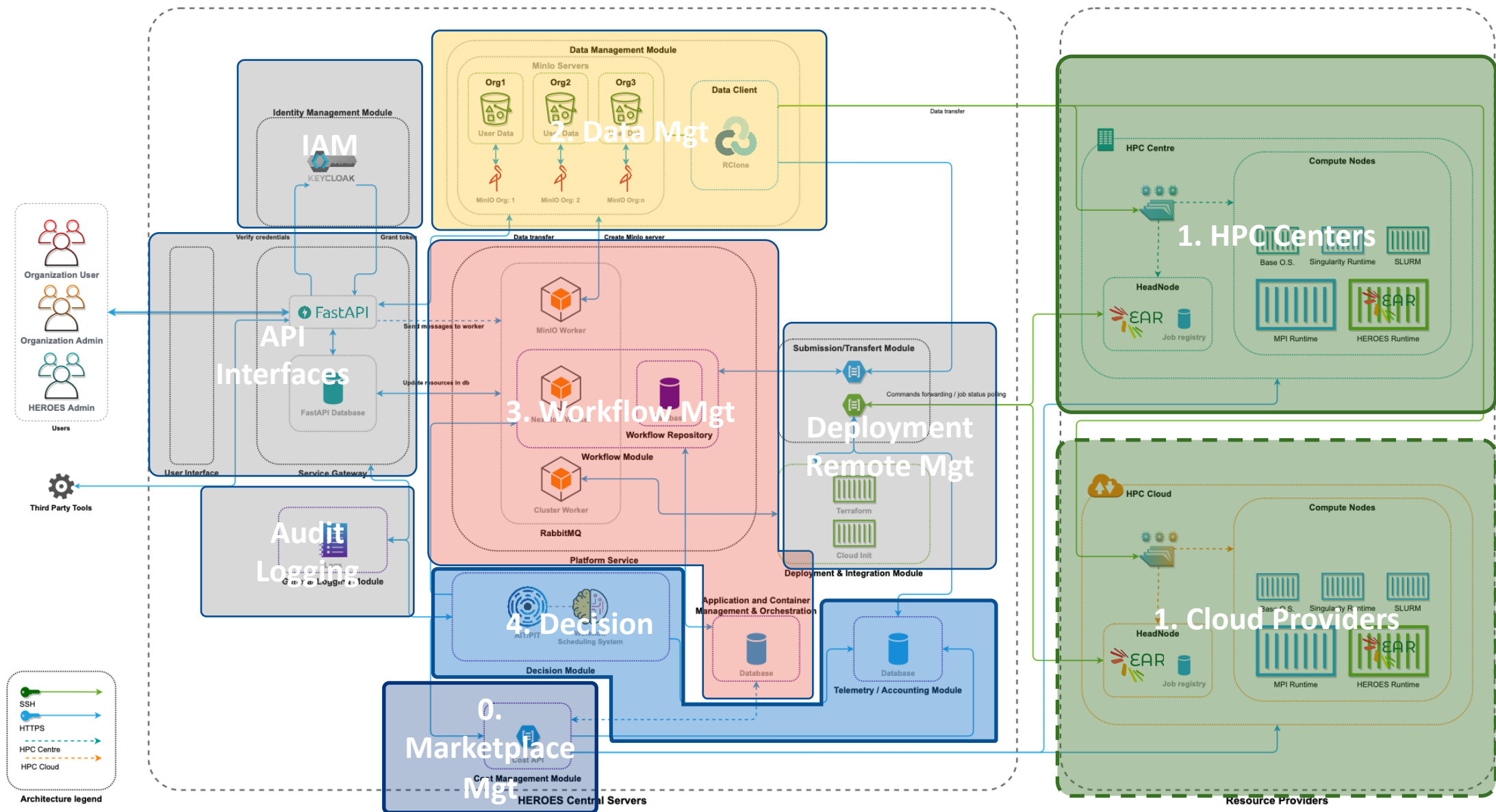




Architecture



Architecture



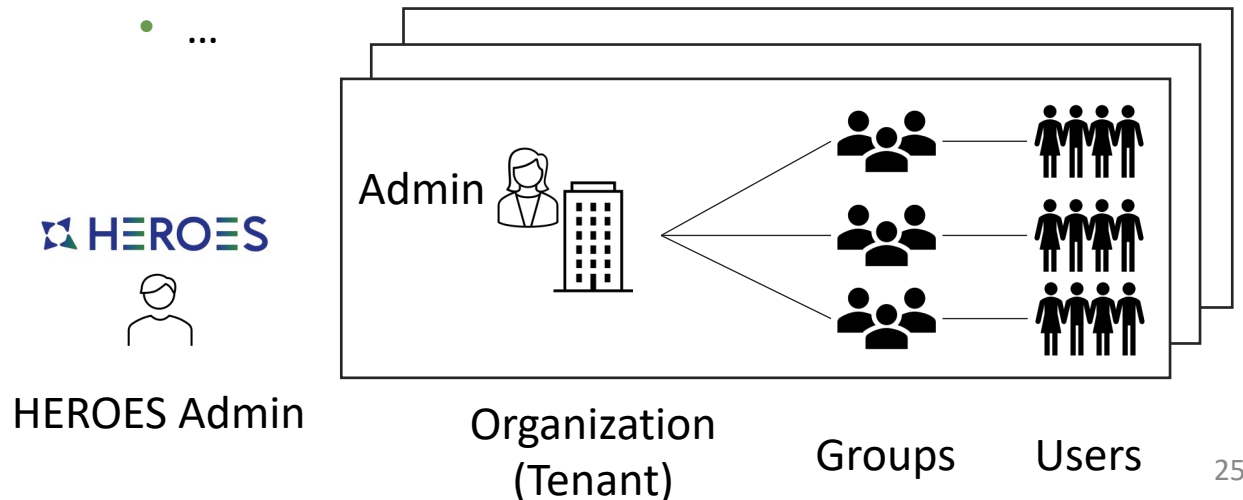
1. Organizations

- Organization User actions

- Login into the platform
- Access their own existing data, if any
- Select a workflow / job template
- Input job parameters
- Receive available alternatives, based on energy, performance, cost parameters
- Select the platform for the job
- Run the job
- Recover job results
- ...

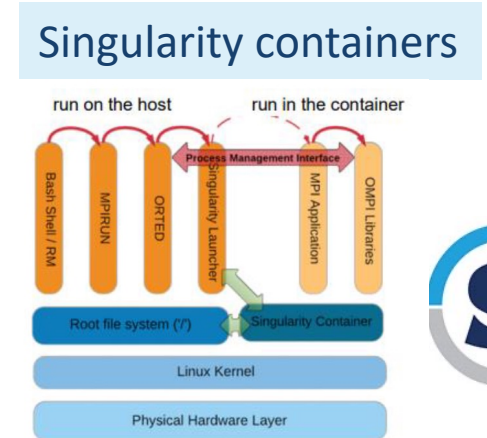
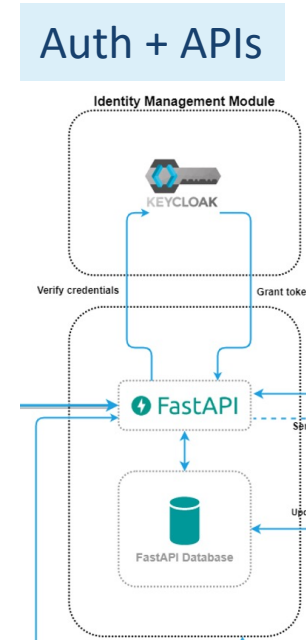
- Organization Admin actions

- Login into the platform
- Access the management interface for the Organization
- Create/Change/Delete Users
- Manage data storage for the Organization and the Users
- Select, review and deploy cloud infrastructural objects (Storage, Cloud HPC)
- Select which on-premises HPC infrastructures should be available for usage within the Organization
- Display Organization resource consumption and generate reports
- ...



1. Access to HPC Centre & Security

- Effective segregation of user data and runtime environment on platforms that allow for a single service user through HEROES-FS & container isolation techniques
- Platform isolation, with a single entry point for end-user interaction
- Centralized user management through KeyCloak
- Containerized applications through Singularity that allows for reproducibility, portability and data isolation
- “HEROES FS”: hierarchy of directories per organization/project/user, mounted in container and prevent directory traversal
- Single SSH account,



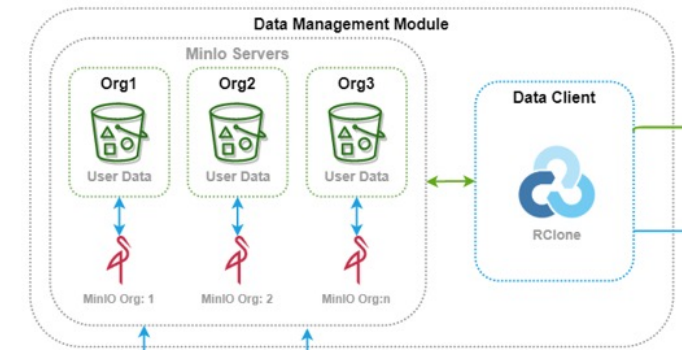
HEROES-FS

```
/efs/  
├── huser  
│   └── heroes-fs  
│       ├── orga8cae-8413-4856-909d-1fb28a290fe1  
│       │   ├── alicexxx-appl-ewor-thxx-1fb28a290fe1  
│       │   └── bobxxxx-bana-nara-maxx-1fb28a290fe1  
│       ├── orgb40aa-c236-4bfd-a810-60cc89d8fc82  
│       │   ├── anaxxxx-park-enxx-xxxx-60cc89d8fc82  
│       │   ├── johnxxx-doex-xxxx-xxxx-60cc89d8fc82  
│       │   └── tomxxxx-bosl-eyxx-xxxx-60cc89d8fc82  
│       ├── orgc7736-1c92-4957-860f-f653591d96a3  
│       │   ├── johnxxx-brya-nxxx-xxxx-f653591d96a3  
│       │   └── timxxxx-call-agan-xxxx-f653591d96a3  
│       └── heroes-rt  
│           └── ear  
│               ├── ear-1.21pre-16.el5_7.7.x86_64  
│               │   ├── ear.1  
│               │   ├── ear.2  
│               │   └── ear.3  
│               └── ear-1.21pre-16.el5_7.7.x86_64.tar.gz
```

2. Data Management

- Centralized data management between HPC centers/CSPs
- Object storage: MinIO, Multi-tenancy structure
 - 1 server per organization
 - 1 bucket per user
- Integrations with workflows between platform endpoints
- Keycloak managed tokens for authenticated & authorized users
- RClone over SSH as actual data mover

Data Management Module



Data Management API

| | | | |
|--------|---|--|---|
| GET | /organization/data/list | List all buckets available for the authenticated user | ▼ |
| POST | /organization/data/bucket | Create a new bucket in organization minio server | ▼ |
| DELETE | /organization/data/bucket | Delete bucket from organization minio server | ▼ |
| GET | /organization/data/bucket/{bucket}/list | List all objects presents in the target bucket of the authenticated user | ▼ |
| GET | /organization/data/download | Download file from bucket | ▼ |
| POST | /organization/data/upload | Upload file to bucket | ▼ |

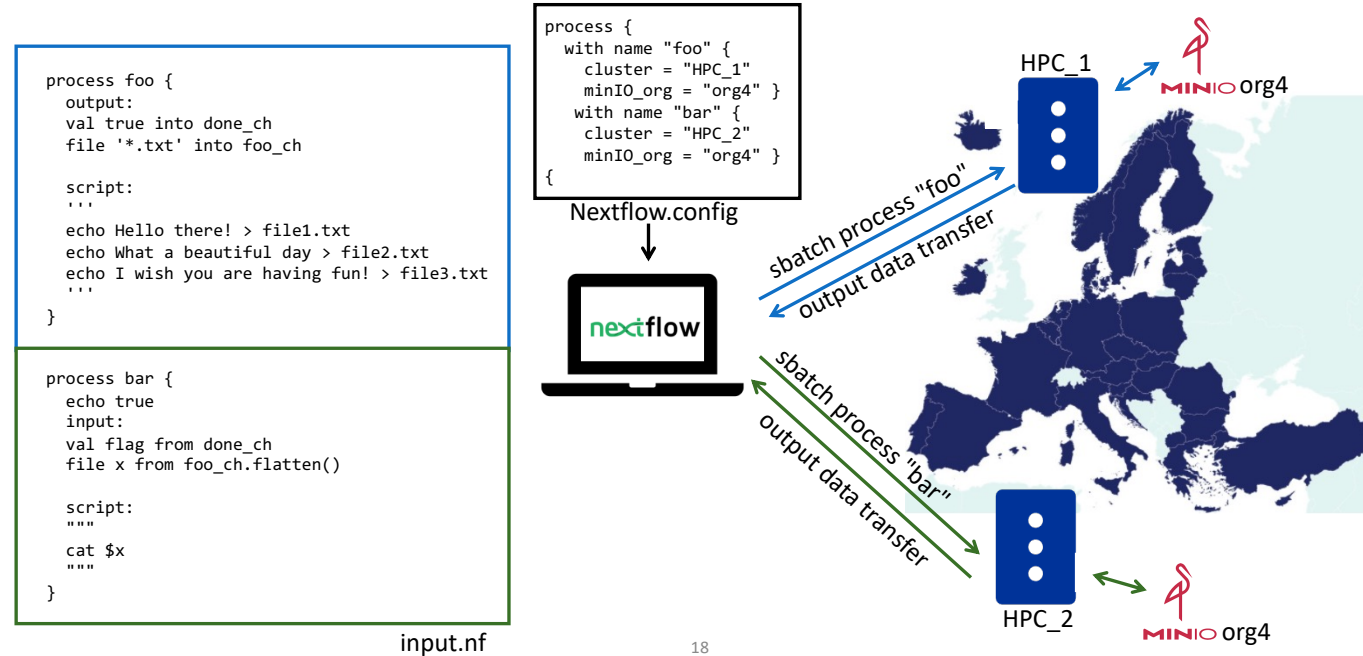
3. Workflow manager

- **Integration of NextFlow**

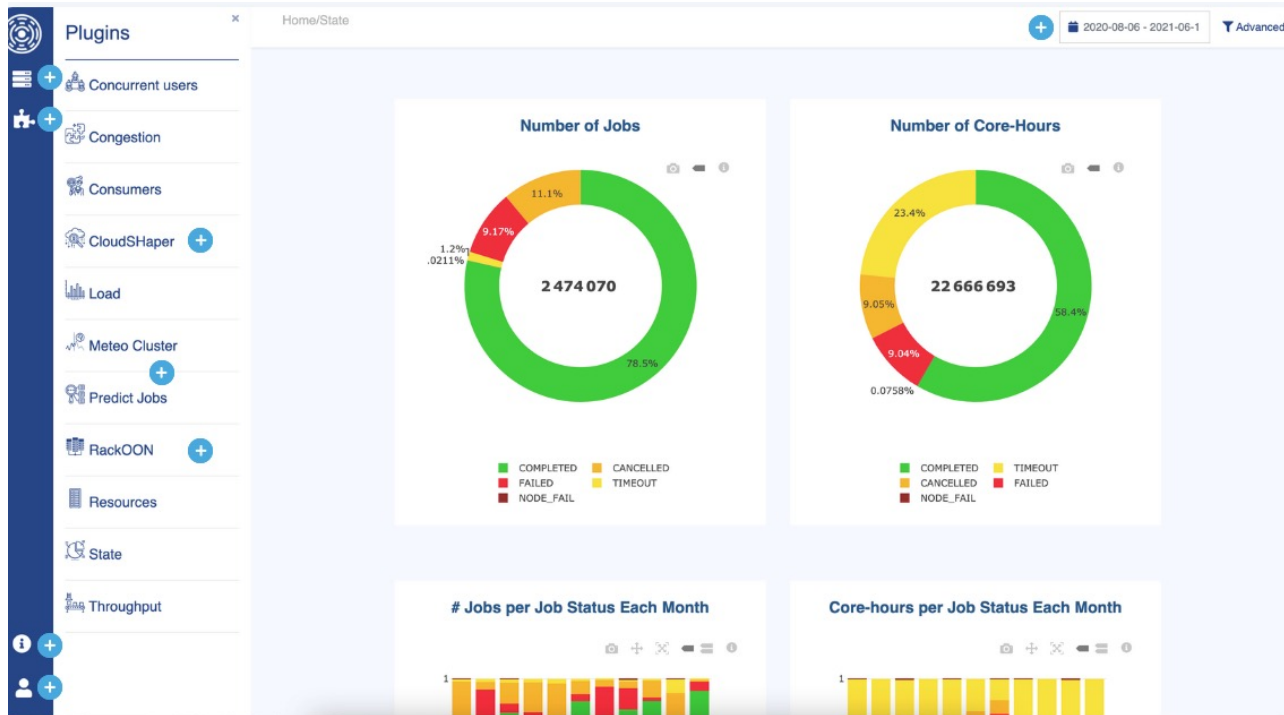
- Native integrations for some technologies used on the HEROES platform and other features
- Support for different jobs schedulers
- Singularity support

- Nextflow extensions for HEROES

- **Configuration parameters:** cluster credentials and security key, name of the servers in RClone configuration.
- **Call to external processes** to perform **data transfer** to remote clusters and file systems.
- **Execution** of jobs on **remote** clusters
- Management of the exit status and **data transfer** for each workflow process.



4. Decision module - OKA



OKA is the Data Science platform for HPC Environments
It provides a global view of HPC clusters and their usage
OKA is extensible, adaptable and dynamic

In HEROES, OKA is the core of the Decision module:

- Centralization of metrics from multiple sources
 - Jobs/Workflows accounting logs/metrics
 - Infrastructure metrics (node/core availability...)
 - Cost & Energy consumption tracking
- Provides both analytics and predictive capabilities



<https://oka.how>

4. Decision module – EAR

System software for energy management

EAR main goal: From powerful Data Centers to Powerful and Energy-Efficient Data Centers



Be cost-effective

Consume what you need
and no more
Optimize your electricity
bill



Be eco-responsible

Resources are limited
Carbon footprint
minimization



Be energy-efficient

Understand/Optimize your
system energy
consumption
Know why!

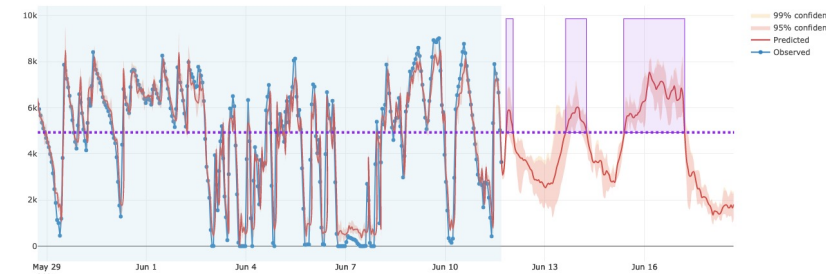


<https://www.bsc.es/research-and-development/software-and-apps/software-list/ear-energy-management-framework-hpc>

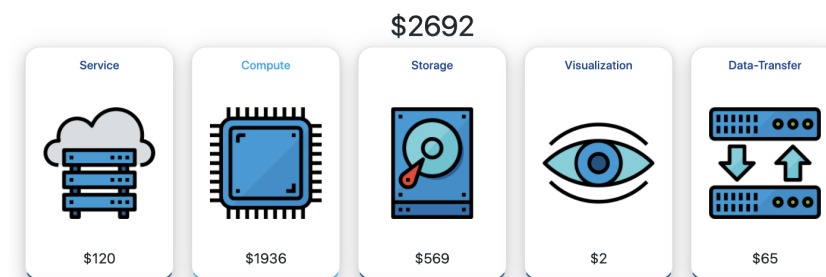
4. Decision module

HEROES Plugin in OKA

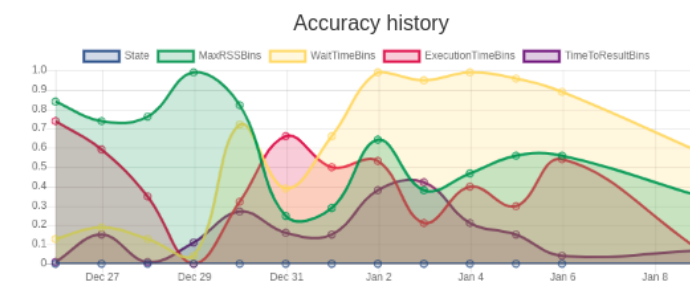
- API to get the placement (cluster selection + submission parameters) of a set of jobs based on:
 - Details about the jobs (user, job name, application, number of cores required, timelimits...)
 - An optimization choice: $f(\text{cost, performance, energy})$
- Placement algorithm will rely on
 - Constraints matching (job technical requirements, e.g., arch, min #cores/nodes...)
 - AI/ML models to predict
 - The global behaviour of HPC/Cloud platforms thanks to MeteoCluster
 - The cost associated with a workload thanks to CloudShaper
 - The needs and behaviour of individual jobs thanks to Predict-IT
- Integration options from Recommendations up to applying automated actions.



MeteoCluster



CloudShaper



Predict-IT



www.heroes-project.eu

heroes@ucit.fr



This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 956874. The JU receives support from the European Union's Horizon 2020 research and innovation programme and France, Spain, Italy.