Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
INFN

Servizi e risorse di HPC-BD-AI

Giacinto Donvito

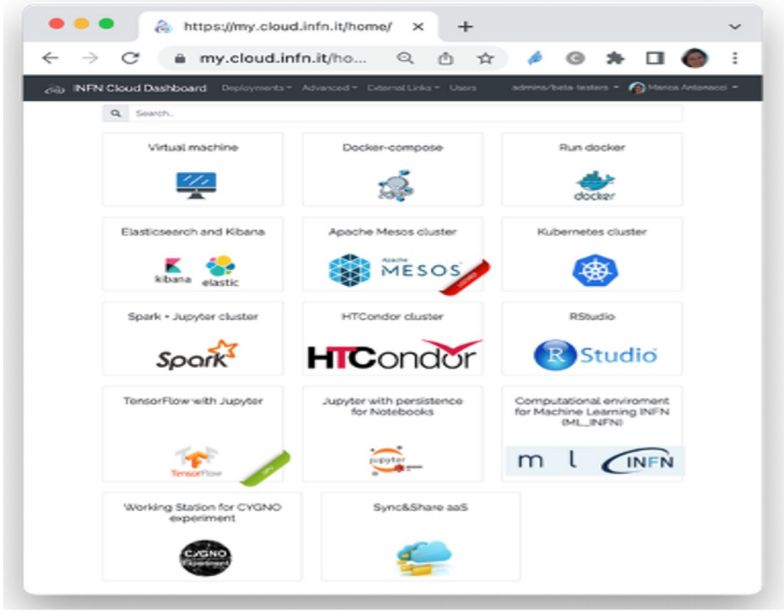
Workshop Terabit 2024

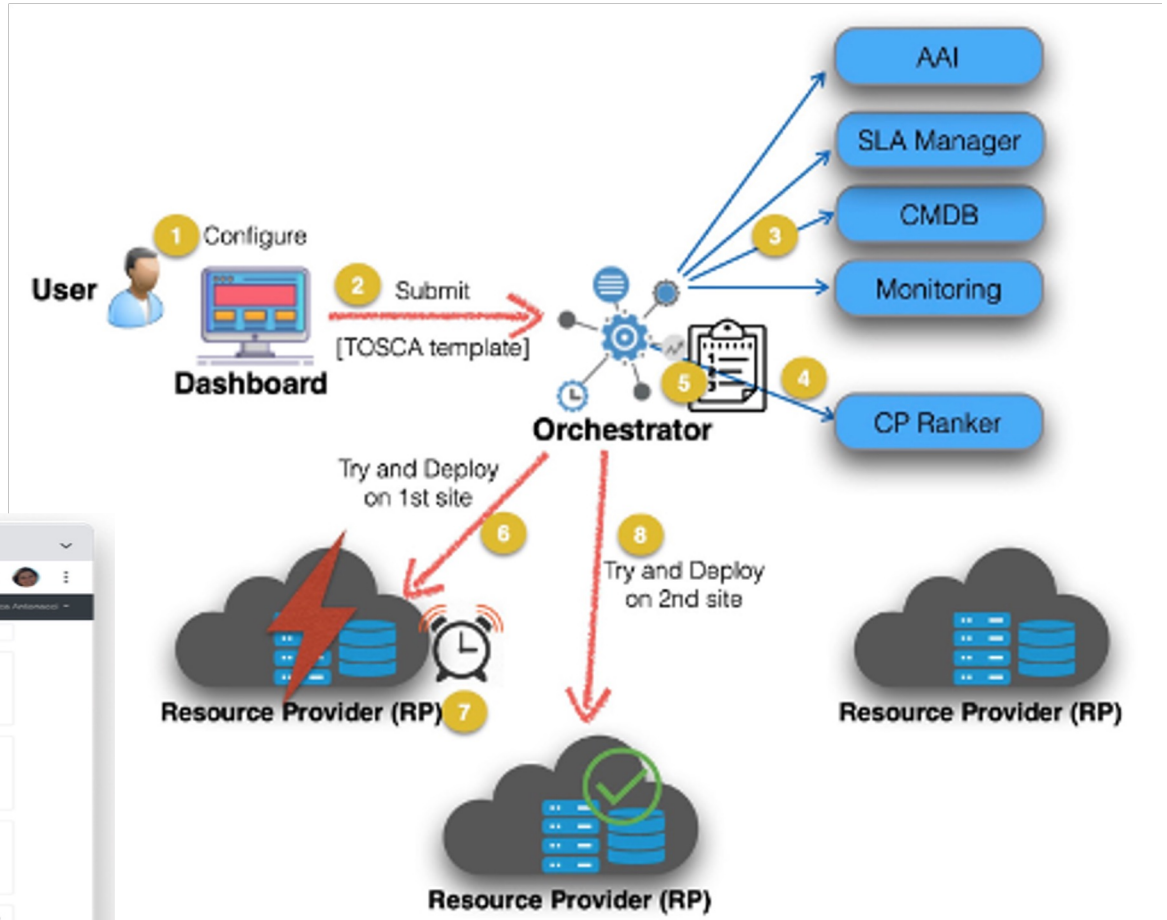
Claudio Grandi

Facilità d'uso

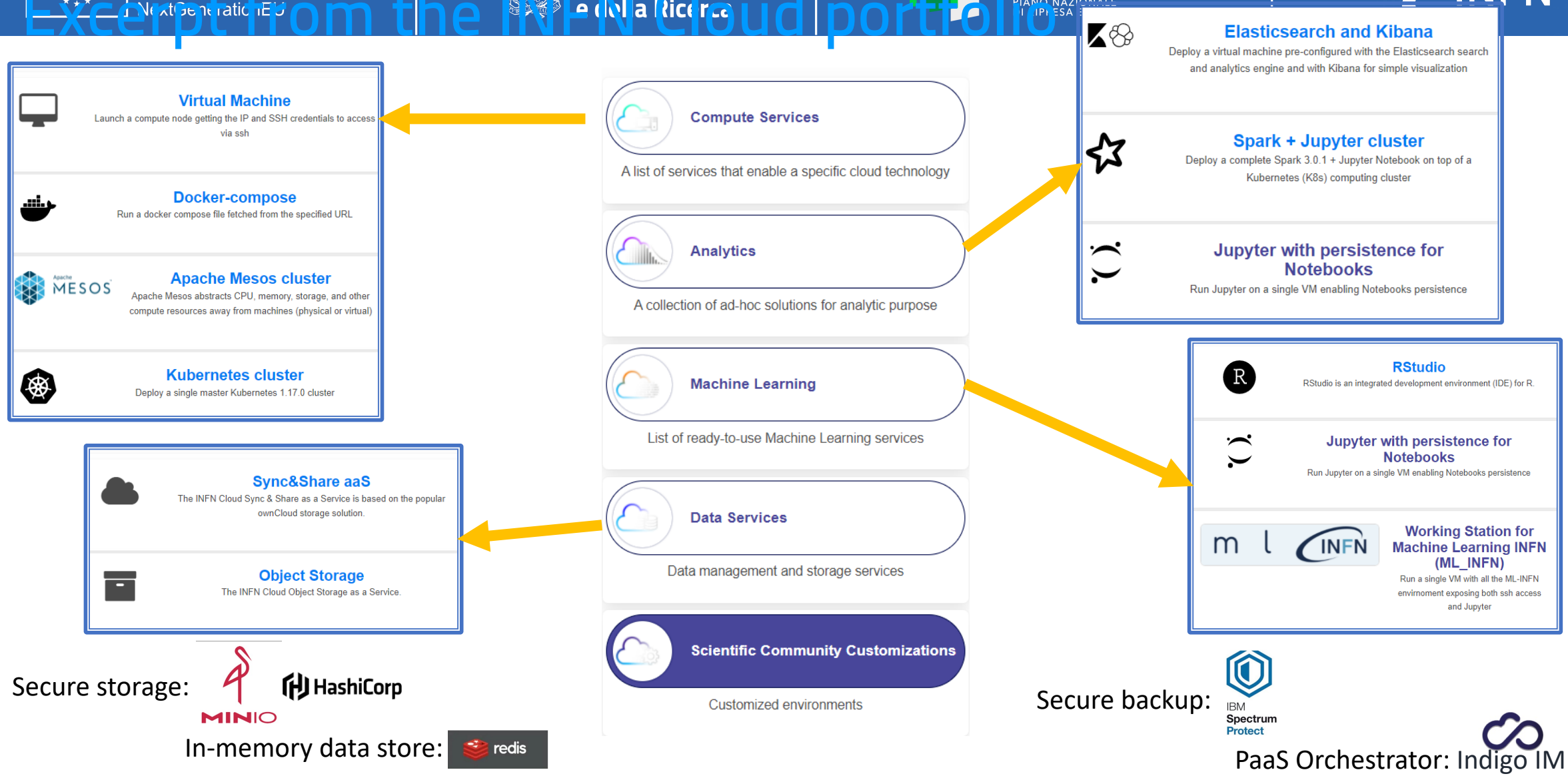
La federazione servirà utenti con diverse competenze sul calcolo

- Nascondere all'utente finale la complessità dell'infrastruttura sottostante
- Esperti dei diversi campi sviluppano piattaforme che consentono l'uso efficace delle infrastrutture attraverso la composizione di servizi e risorse





Excerpt from the INFN Cloud portfolio



Secure storage: MINIO HashiCorp

In-memory data store: redis

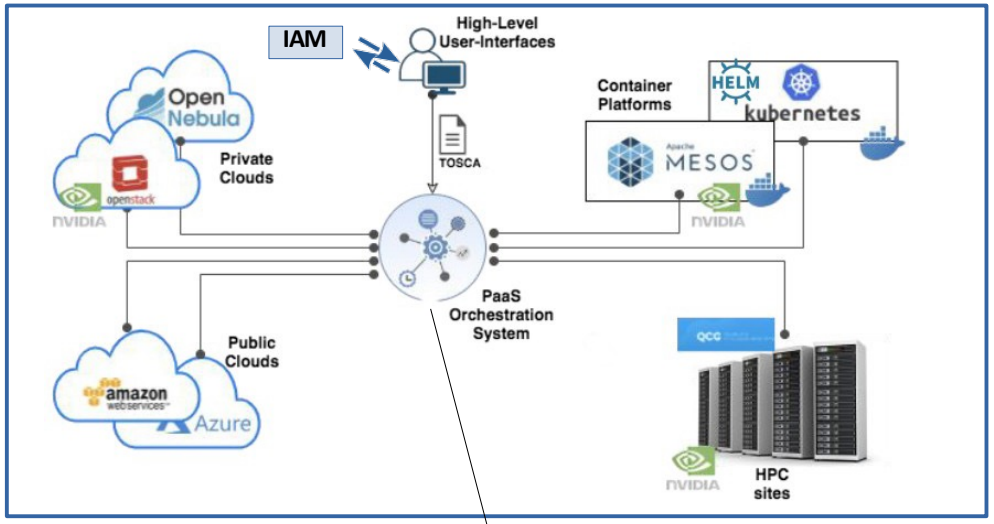
Secure backup: IBM Spectrum Protect

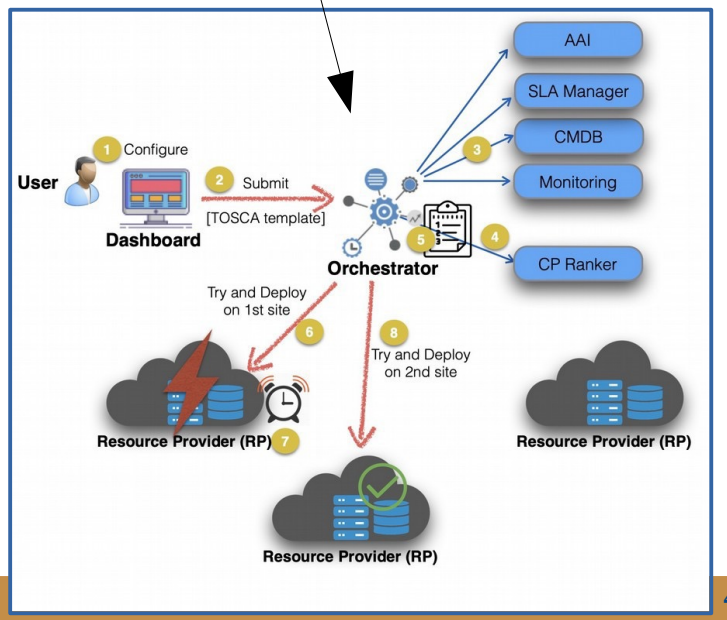
PaaS Orchestrator: Indigo IM

Selectable storage QoS levels: fast (SSD), normal (HDD), archive (tape-backed), remote replicas

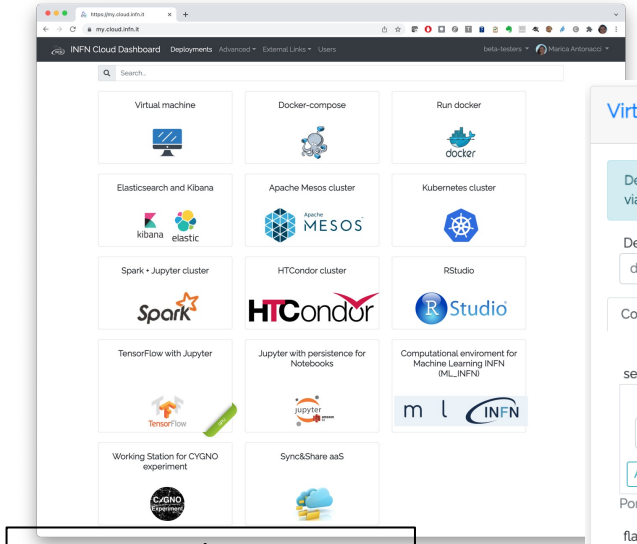
Architectural key points

- Open source, vendor neutral architecture.
- Dynamic orchestration of federated resources
  - via the INDIGO PaaS Orchestrator across all participating Cloud infrastructures, according to agreed SLAs, hardware available, data location and Rules of Participation.
- Consistent authentication and authorization technologies and policies at all Cloud levels
  - via OAuth and OpenID-Connect, supporting also legacy AAI solutions, via INDIGO-IAM (Identity and Access Management).



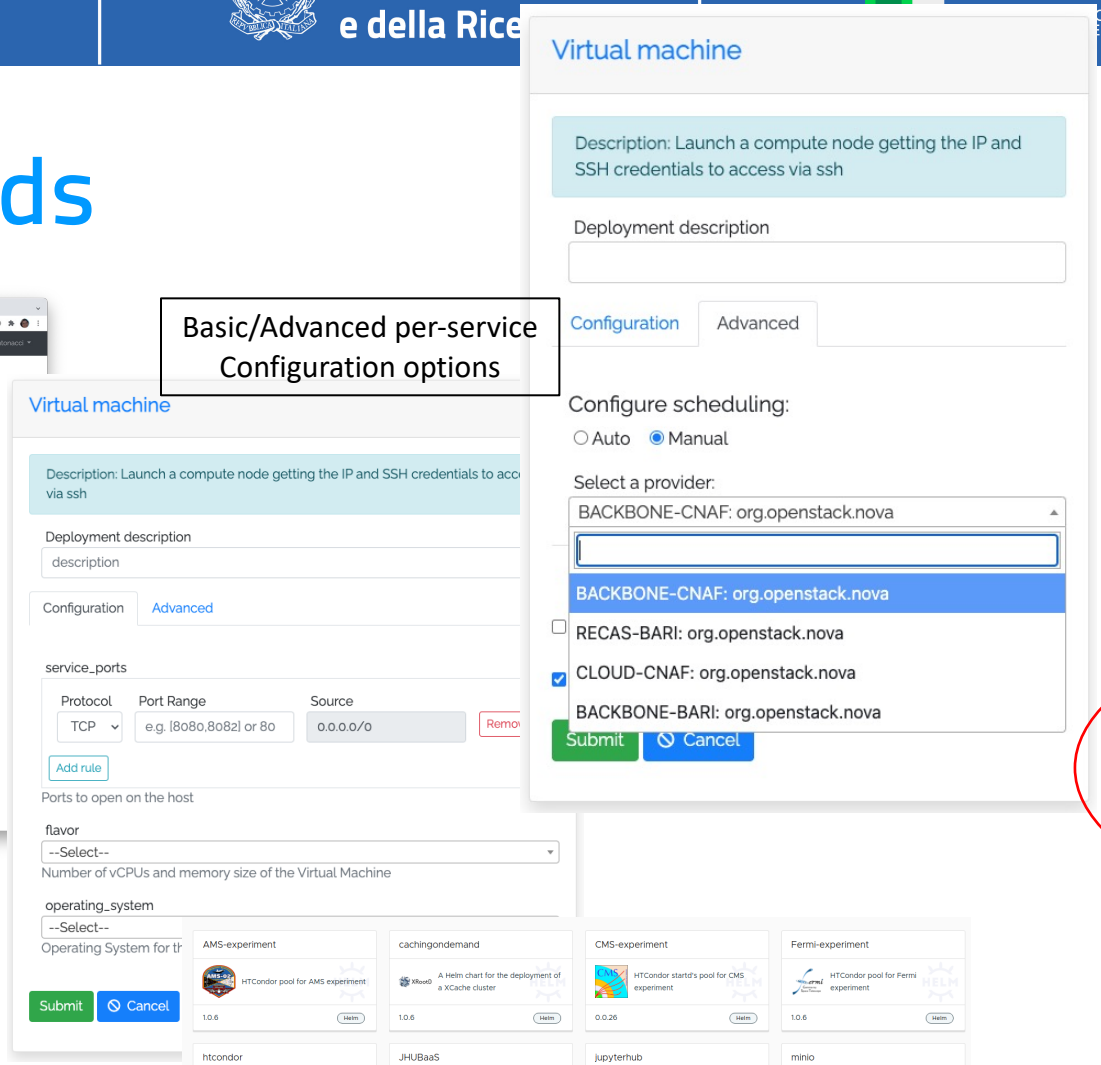


Status, Dashboards

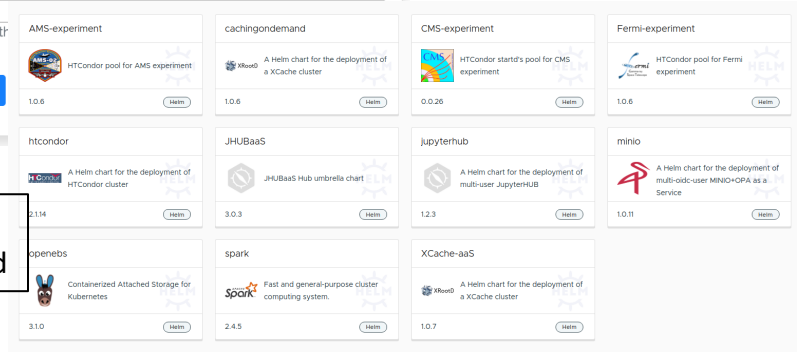


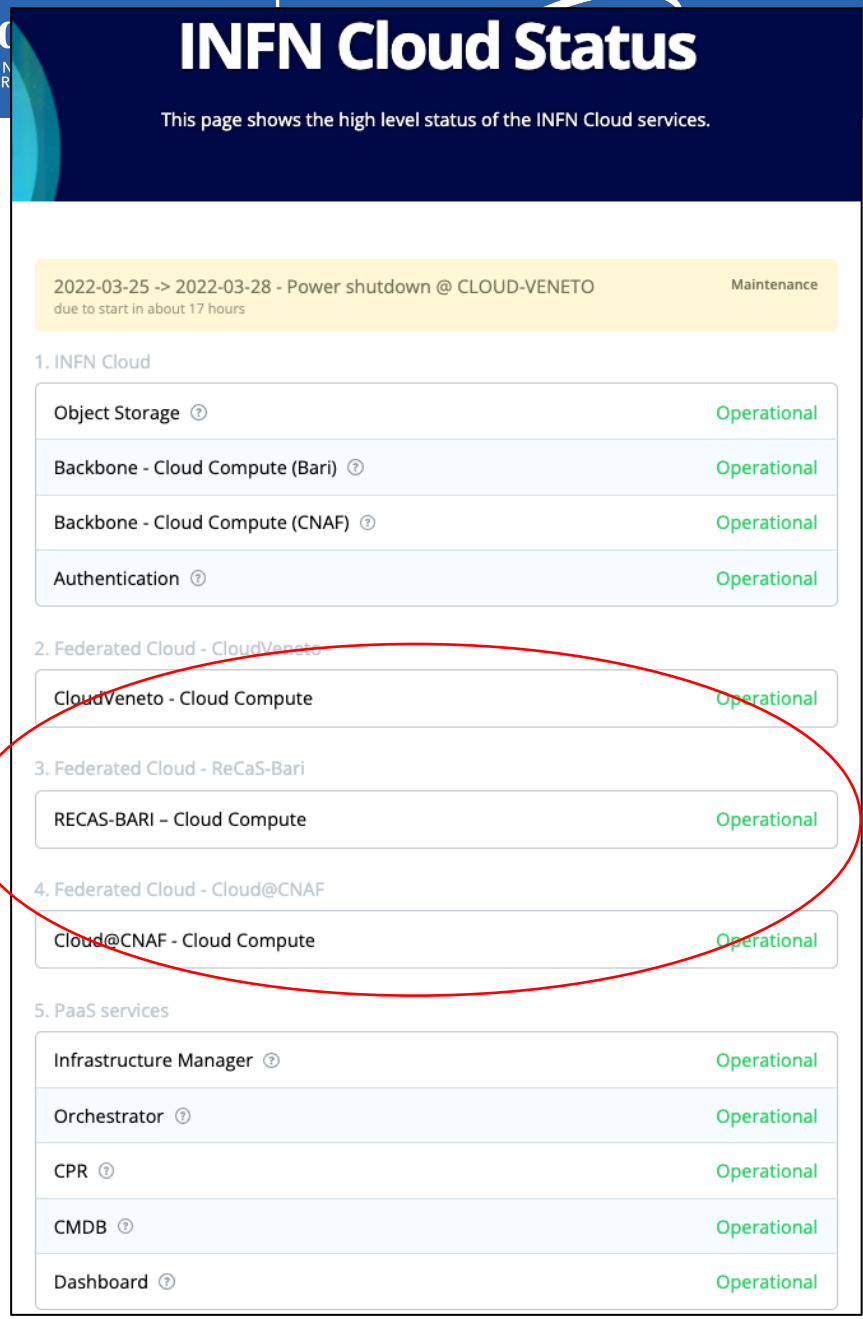
Per-user/per-group
View of the main dashboard

Basic/Advanced per-service Configuration options



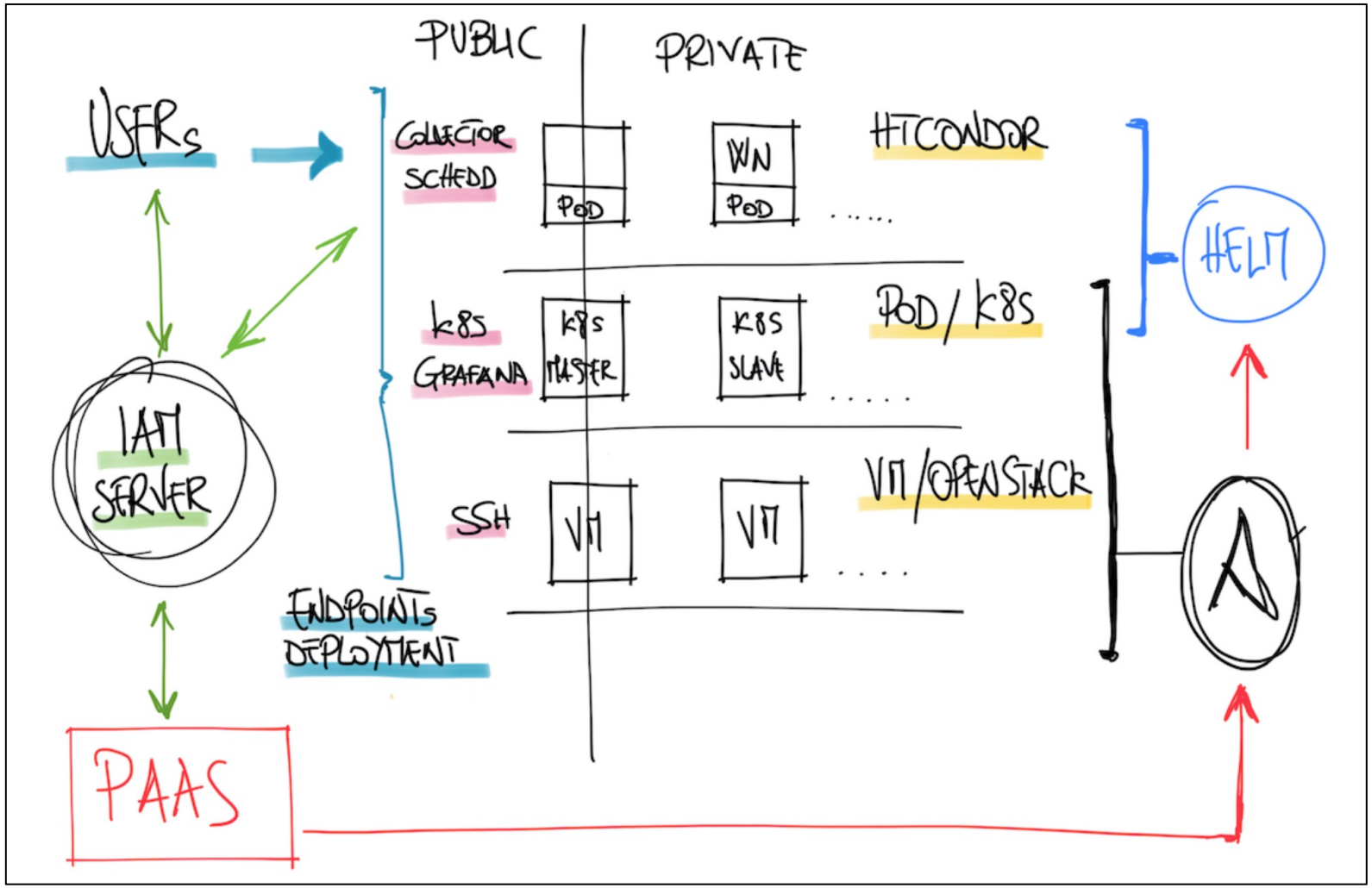
Kubeapp-level
Service dashboard





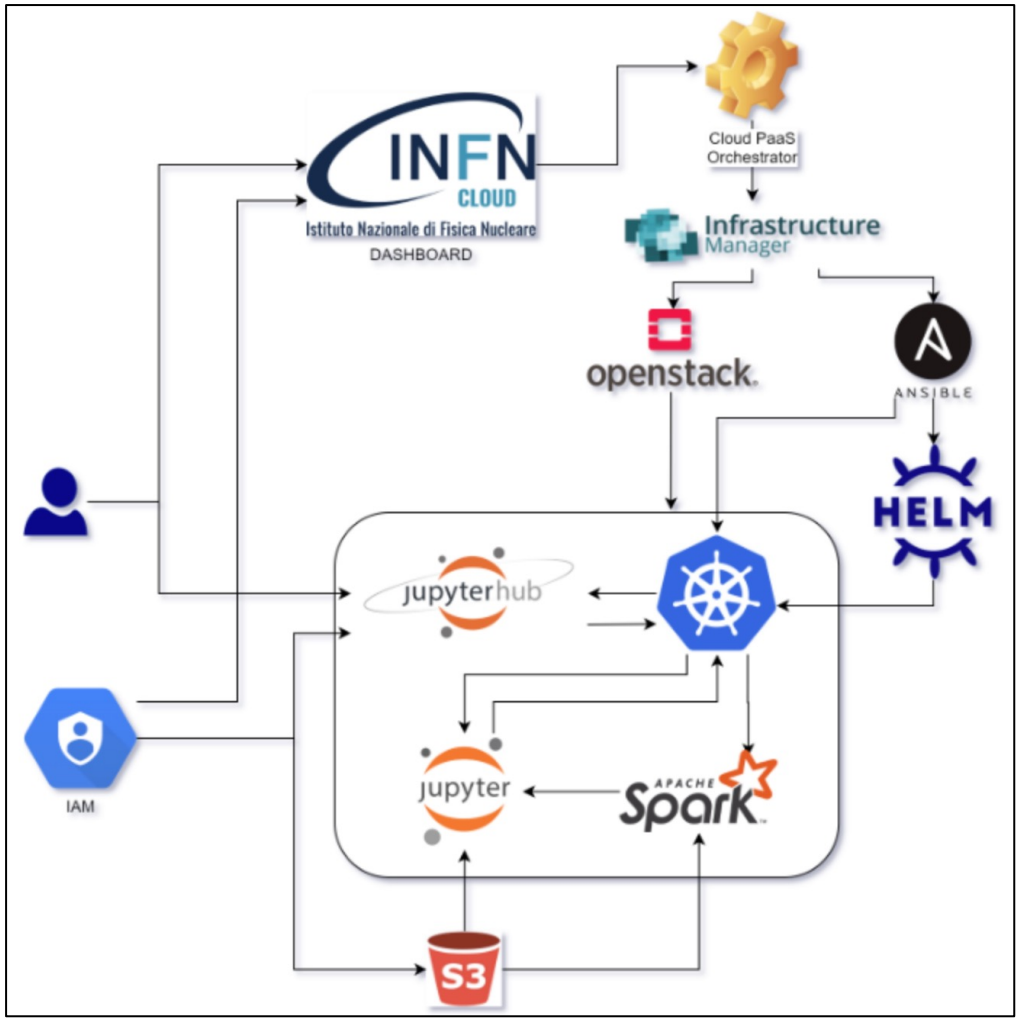
Examples of INFN Cloud services: on-demand HTCondor

- This service instantiates a k8s cluster which is then used to automatically deploy an HTCondor cluster.
- HTCondor services are deployed using dedicated PODs.



Examples of INFN Cloud services: on-demand Spark + Jupyter

- This service creates a Spark cluster on-demand, based on Kubernetes as resource manager and on JupyterHub for the user interface.
- The JupyterHub application, the Jupyter notebooks launched from it, the Spark driver and executors are all Docker containers orchestrated by a Kubernetes cluster deployed on a federated cloud (the box in the picture on the right). The components outside the box implement the workflow to instantiate the service.



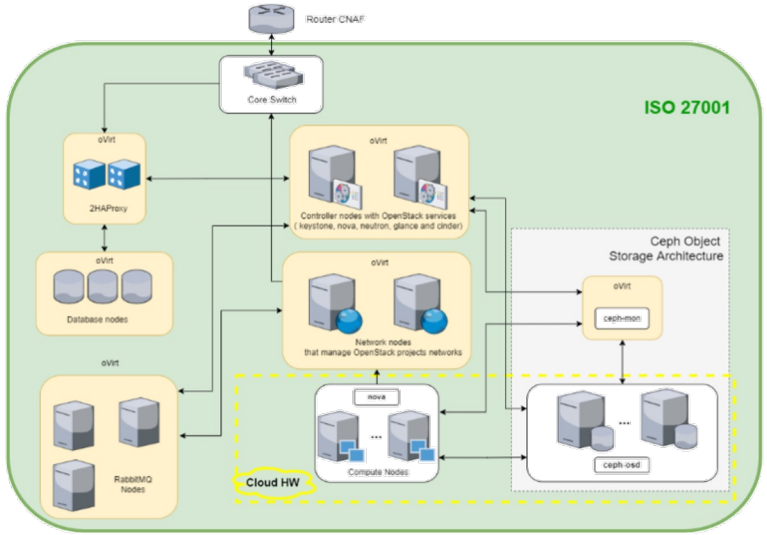
Flessibilità

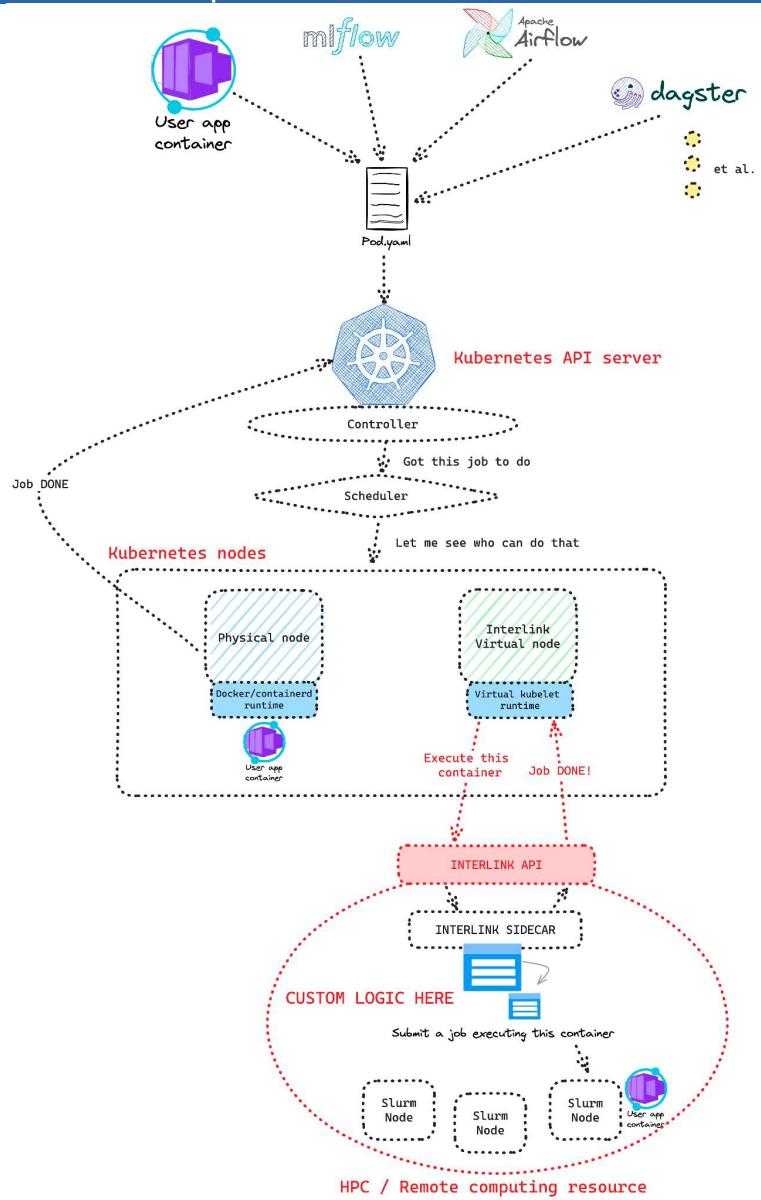
Supporto a diversi metodi di accesso alle risorse, con attenzione a:

a. Trasparenza e facilità d'uso
b. Efficienza ed efficacia

Supporto a requisiti specifici delle applicazioni

Ad es. Piattaforme con requisiti particolari sulla privacy





Accesso alle risorse di «compute»

Attualmente sono disponibili risorse di calcolo con accesso secondo diversi paradigmi:

- OpenStack
- K8s (still not production-ready)

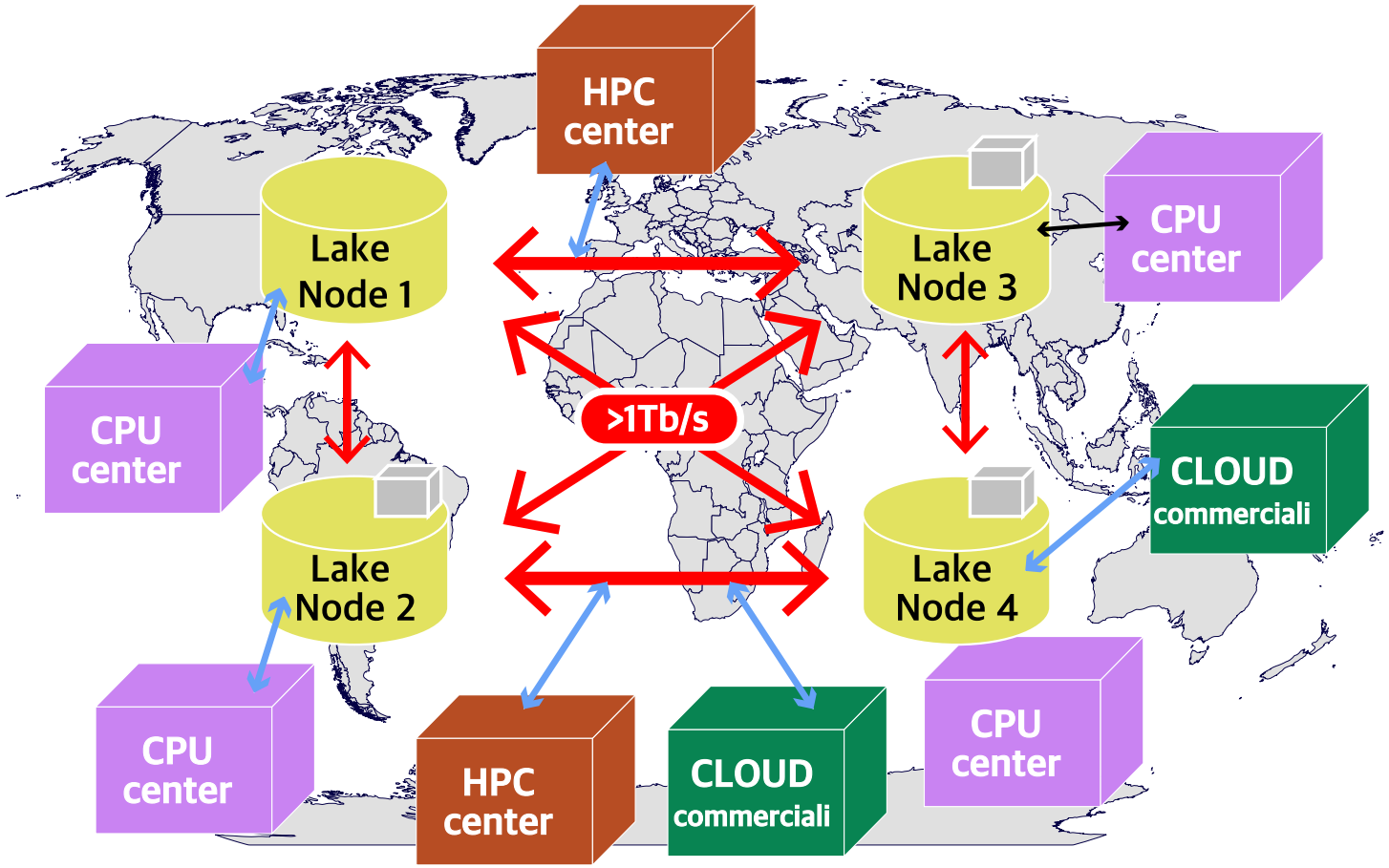
} PaaS Federation Layer

- HTCondor
- SLURM
- Grid-approach

} Dedicated access or overlay services

Modello data-centrico

- Disaccoppiamento di storage e CPU
- Nodi storage interconnessi tramite una rete a banda larga
- Nodi eterogenei possono accedere ai dati in modo indipendente dalla locazione



Accesso allo «storage»

Attualmente sono disponibili risorse di storage:

- S3
- OpenStack Swift

} Object Storage

- Webdav
- Xrootd
- gridftp

} From posix-to-access or transfer

Gestibili via RUCIO-based data lake

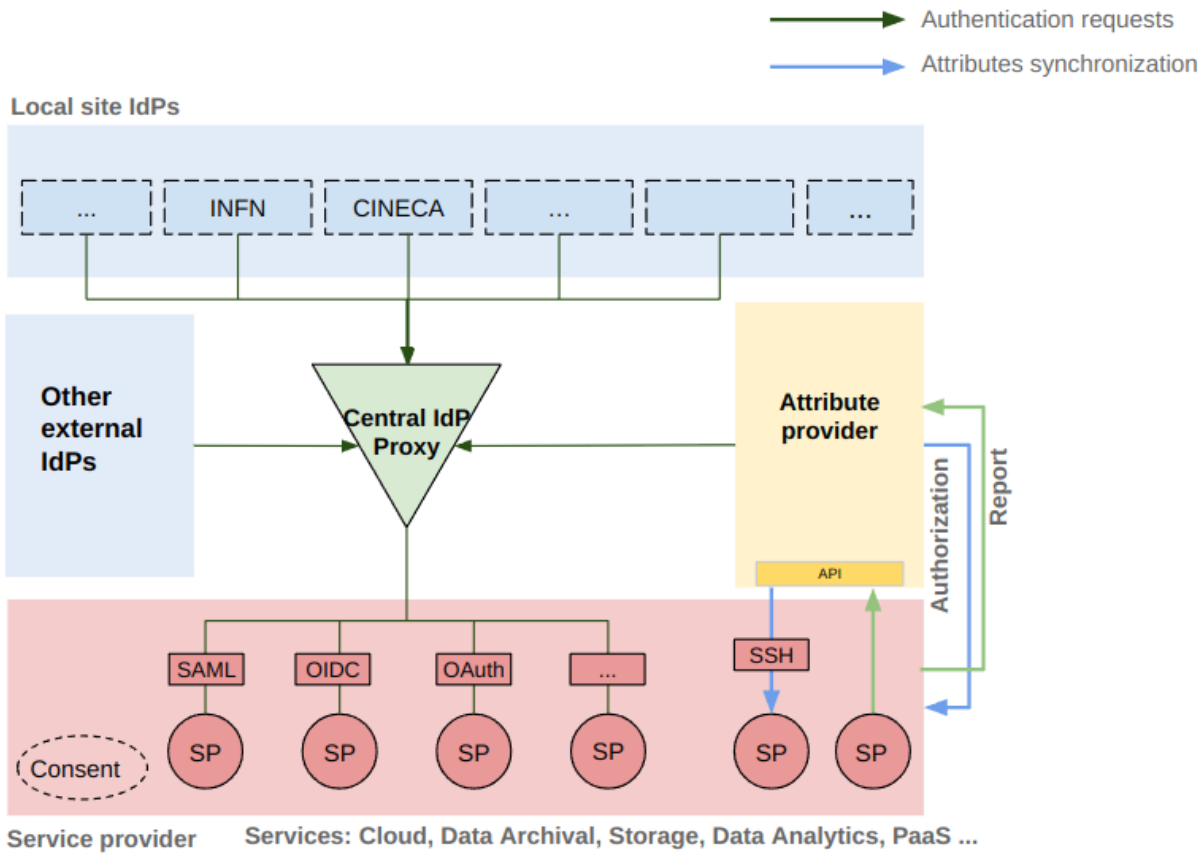
Inclusività

La federazione includerà data centres che sono già in produzione, e parte di comunità internazionali

- Le procedure per entrare nella federazione devono essere non intrusive
- Bisogna usare standard quando possible, e sviluppati quando non ci sono

La federazione servirà utenti di diverse organizzazioni in diversi campi

- Le procedure di ingress devono essere il più semplice possible
  - Ad esempio tramite l'uso di federazioni di identità



HPC Bubbles

Nodo CPU	192 core fisici 1.5TB RAM DDR5 IB NDR 400G 20TBL (SSD) + dischi di sistema
Nodo GPU	Come CPU + 4x NVIDIA H100 SXM5 con minimo 80GB e memoria HBM2e
Nodo FPGA	32core RAM 768GB DDR5 IB NDR 440G 4 x XILINX U55C o 4 x TerasicP0701
Nodo Storage (CEPH Bricks)	64 core fisici 1TB RAM DDR5 384 TBL HDD + 25.6 TBL NVMe
Accessori	Switch IB, Switch ETH Cavi IB, Cavi ETH Transceiver vari Assistenza 3+2

Risorse HPC bubbles

Accordo Quadro Nazionale

- Listino prezzi per nodi + accessori
- 2 anni di validità
- Lotto1
  - CPU, GPU, FPGA
- Lotto2
  - Storage (1 nodo 380 TB raw; ≥ 250 TBN)
- Contratti stipulati per entrambi i lotti
- Ordini effettuati

Sito	Nodi CPU	Nodi GPU	Nodi FPGA	Nodi Storage
CNAF	26	30	4	52
BA	24	6	0	32
MI-BI	0	0	4	0
PI	8	0	0	0
TO	6	6	0	0
LNGS	0	6	0	12
NA	18	1	2	8
RM1	12	0	0	0
PD/LNL	10	6	0	0
LNF	20	6	0	6
CT	12	0	0	8
MI	4	0	0	0
TOTALE	**160**	**61**	**10**	**118**