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

Claudio Grandi

Le HPC Bubble

D.Cesini (INFN-CANF), G. Donvito (INFN-BARI)

C3SN
LNF 4 Settembre 2025

Claudio Grandi

HPC Bubbles

 Nodo CPU Lenovo Lenovo ThinkSystem SR665 V3	192 core fisici - Dual AMD AMD EPYC 9654 96C 360W 2.4GHz 1.5TB RAM DDR5 IB NDR 400G - NVIDIA ConnectX-7 NDR OSFP400 1-Port PCIe Gen5 x16 InfiniBand Adapter 20TBL (SSD) + dischi di sistema
 Nodo GPU Lenovo ThinkSystem SR675 V3	Come CPU + 4x NVIDIA H100 SXM5 con 80GB HBM3 (non HBM2e come da offerta)
 Nodo FPGA Lenovo ThinkSystem SR675 V3	32core - AMD EPYC 9124 16C 200W 3.0GHz Processor RAM 768GB DDR5 IB NDR 440G 4 x XILINX U55C o 4 x TerasicP0701
 Nodo Storage (CEPH Bricks) DELL PowerEdge R760xd2	64 core fisici – Dual Intel Xeon Gold di quarta generazione (Sapphire Rapids), modello 6428N, frequenza 1.8GHz, 32Core/64Thread, 60M Cache, DDR5-4800, 185W TDP 1TB RAM DDR5 IB Mellanox 400G 384 TBL HDD + 25.6 TBL NVMe
 Accessori	Switch IB, Switch ETH – NVIDIA Modello SN3420 - 12x QSFP28 100GbE + 48x SFP28 25GbE Cavi IB, Cavi ETH Transceiver vari Assistenza 3+2

Claudio Grandi

Terabit "HPC Bubbles"

Quantità nodi con solo fondi Terabit

	Nodo CPU	Nodo GPU	Nodo FPGA Xilinx	Nodo FPGA Terasic	Nodo storage
BA	0	4	0	0	18
CNAF	16	21	2	2	36
MIB	0	0	2	2	0
NA	10	0	2	0	0
PD	6	6	0	0	0
PI	8	0	0	0	0
RM1	8	0	0	0	0
TO	6	6	0	0	0
TOTALE	**54**	**37**	**6**	**4**	**54**

Core: 10.4 kcore fisici
Circa 34 HS/core

GPU: 148 NVIDIA H100
5.0PFLOPS FP64

40 FPGA

21 PB RAW

InfiniBAnd 400Gbs

Claudio Grandi

Quantità nodi HPC BUBBLES con fondi DARE – Terabit per Spoke8 in zona Certifcata ISO27001

	Nodo CPU	Nodo GPU	Nodo FPGA Xilinx	Nodo FPGA Terasic	Nodo storage
BA_DARE	12	6	0	0	6
BA_TerabitS8	0	0	0	0	0
CNAF_DARE	10	9	0	0	16
CNAF_TerabitS8	0-8?	0-8?	0	0	0-6?

Modalità di accesso

- Batch system (SLURM o HTCondor) – CNAF, TO, BA, LNF, LNGS, PD, NA, PI, CT
  - sottomissione da macchina di frontend
    - User interface (Grid o local batch) o jupyter notebook
    - AUTH su user interface via token IAM
    - Possibilità di job interattivi
- Integrazione in Cloud via OpenStack – CNAF, NA, CT, PD
  - AUTH via token IAM
  - Accesso interattivo e interattivo/grafico
- Cluster K8S – BA, MI, RM1
  - AUTH via token IAM
  - Per workflow già containerizzati
- Storage: CEPH/CEPHFS

Allocazioni sulle Bubble (non-ISO)

Allocazioni da:
- ICSC
- DUNE 2024
- JLAB12 2026

Per pledge 2026 su bubble non dovrebbero esserci problemi
- CSN5: 105kGPUhour ➔ 12GPU/yr

	Nodo CPU	Nodo GPU	Core fisici	GPU number	Allocazioni	CPU Allocated	GPU Allocated
BA	0	4	784	16	OH-HPC:50c,4GPU SARAW:192c,4GPU	242	8
CNAF	16	21	7104	84	DUNE:192c,1GPU JLAB12:192c,1GPU DATOME-IA:64c,4GPU Designer: 32,1GPU EARN-DATA: 192c,1GPU ML_G4: 90c,2GPU MoveToDisc:64c,2GPU Stat4Value:192,4GPU	1018	16
MIB	0	0		0			
NA	10	0	1920	0			
PD	6	6	2304	24	ENNIPIML: 192c,4GPU ENIPRED: 288c,6GPU	480	10
PI	8	0	1536	0			
RM1	8	0	1536	0			
TO	6	6	2304	24			
TOTALE	**54**	**37**	**17488**	**148**		**1740**	**34**

Federazione delle Bubble

INFN-Cloud services

Centralized services (SaaS):
•INFN Cloud Registry service (Harbor)
•INFN Cloud object storage service (based on rados-Gateway)

PaaS services:
•Virtual machine
•Docker run
•Docker compose
•Kubernetes cluster
•HTCondor mini
•HTCondor cluster
•Jupyter with persistence for Notebooks
•Jupyter + Matlab (with persistence for Notebooks)
•Spark + Jupyter cluster
•Working Station for CYGNO experiment
•Computational enviroment for AI_INFN
•Elasticsearch and Kibana
•INDIGO IAM as a Service
•Sync&Share aaS

IaaS services:
•Start and Stop
•Hostname choice
•Manage VM ports

INFN-Cloud Status

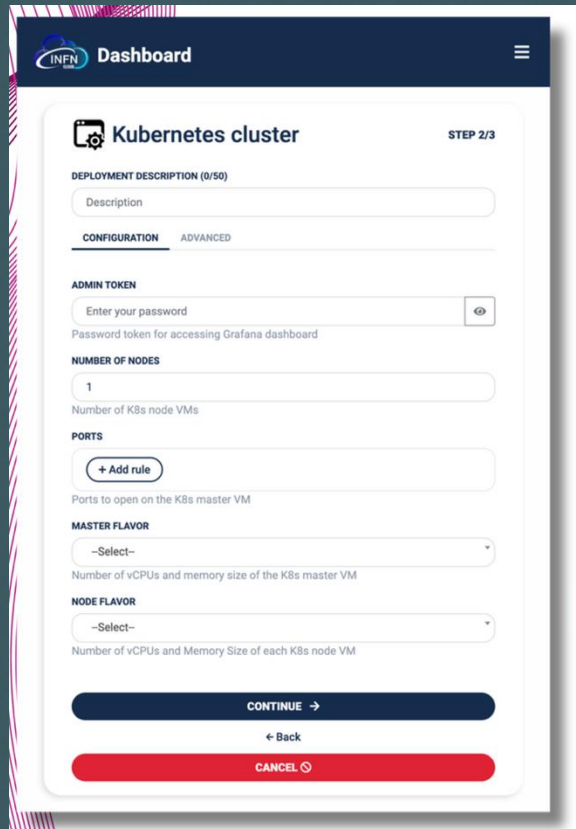
INFN Cloud Backbone dispone di circa 2000 cpu cores e 15TB di RAM distribuiti su due data center. Lo storage (block + object) ammonta, grazie alle nuove risorse ICSC, a circa 7.2PiB RAW.
Oggi circa 400 vm attive, più della metà delle quali usate per servizi, R&D, dal SOC ecc.

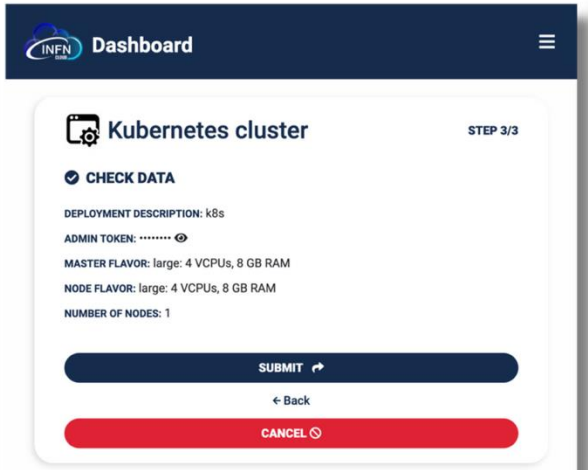
accounting v2.0 attivo

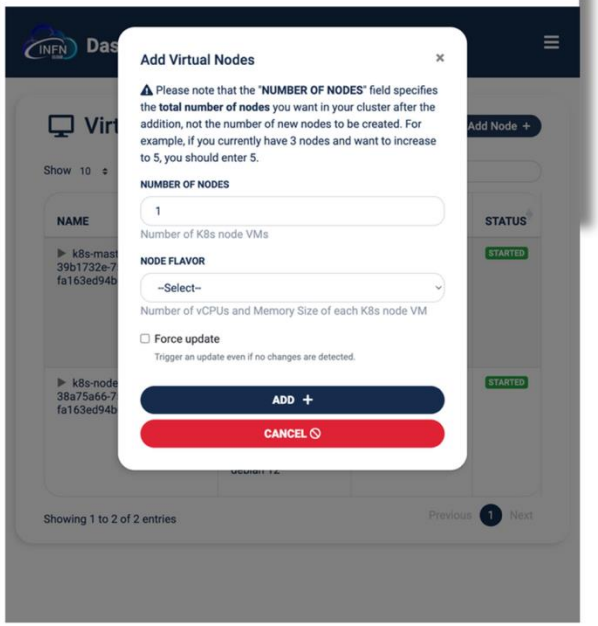
servizi SaaS attivi:

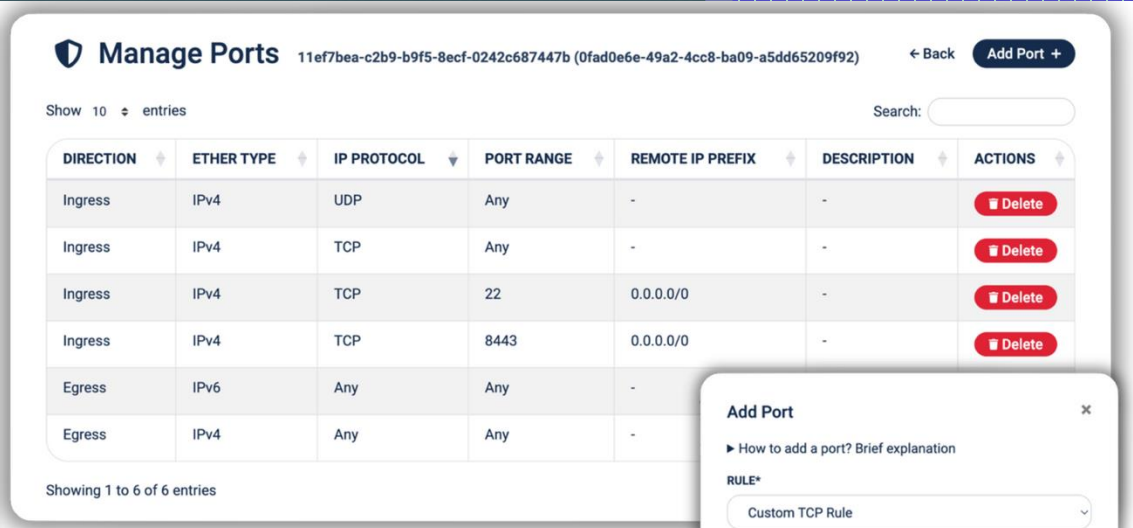
- S3
- docker image repository
- notebook aas
- monitoring aaS (healthchecks)
- iam e dashboard + orchestratore
- Cvmfs
- FTS
- servizio di backup per utenti su S3
- iam e dashboard per ICSC
- docker image repository per ICSC

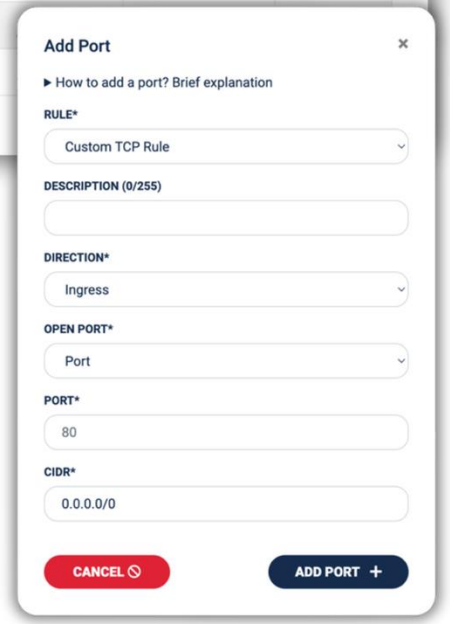
Provisioning di servizi





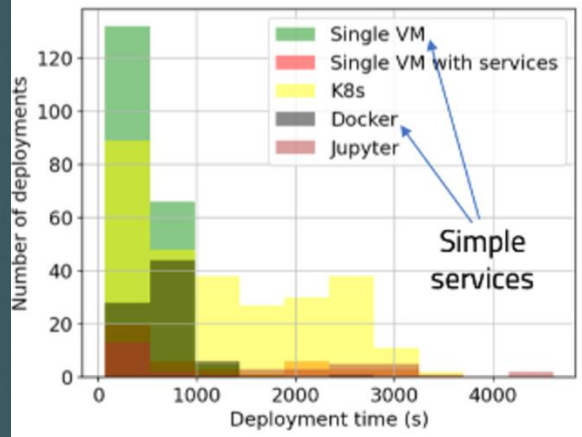


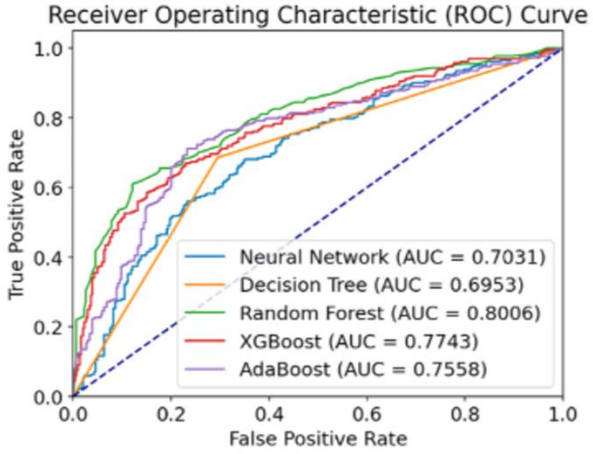




Dashboard
NEW
FUNCTIONALITIES

Smart use of resources





Preparatory Work

- Identified significant metrics and data sources.
- Prepared datasets to thoroughly analyze the problem.

Leveraging AI

FOR SMARTER ORCHESTRATION

interTwin

We are introducing AI-based techniques to **improve provider selection** and **resource allocation** in our orchestration system.
By incorporating artificial intelligence, we can make provider choices more **dynamic** and **efficient**, optimizing deployments across the cloud federation.

- **Predictive Model**: Forecasts deployment success by analyzing key metrics, improving provider selection.
- **Regression Model**: Predicts deployment times based on past data, optimizing scheduling and resource use.

Kubernetes *as the emerged standard for distributed applications (cloud-native)*

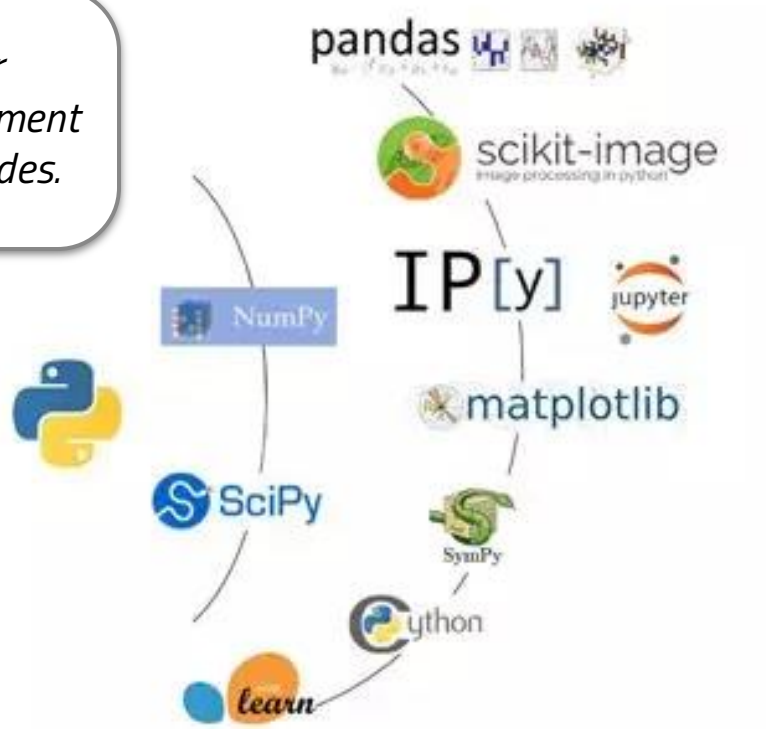
As Python+NumPy emerged as *de facto* standard for Data Science applications in early 2000s, the **Kubernetes APIs have been emerging as the standard for distributed applications**.

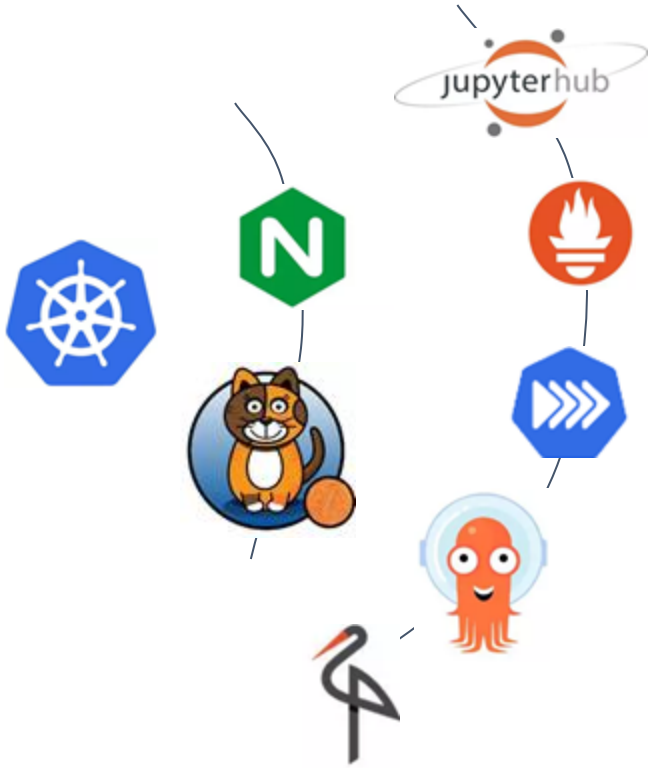
`pip install` ⇒ `helm install`

Kubernetes *is an open-source system for automating deployment, scaling and management of* ***containerized*** *applications on multiple nodes.*

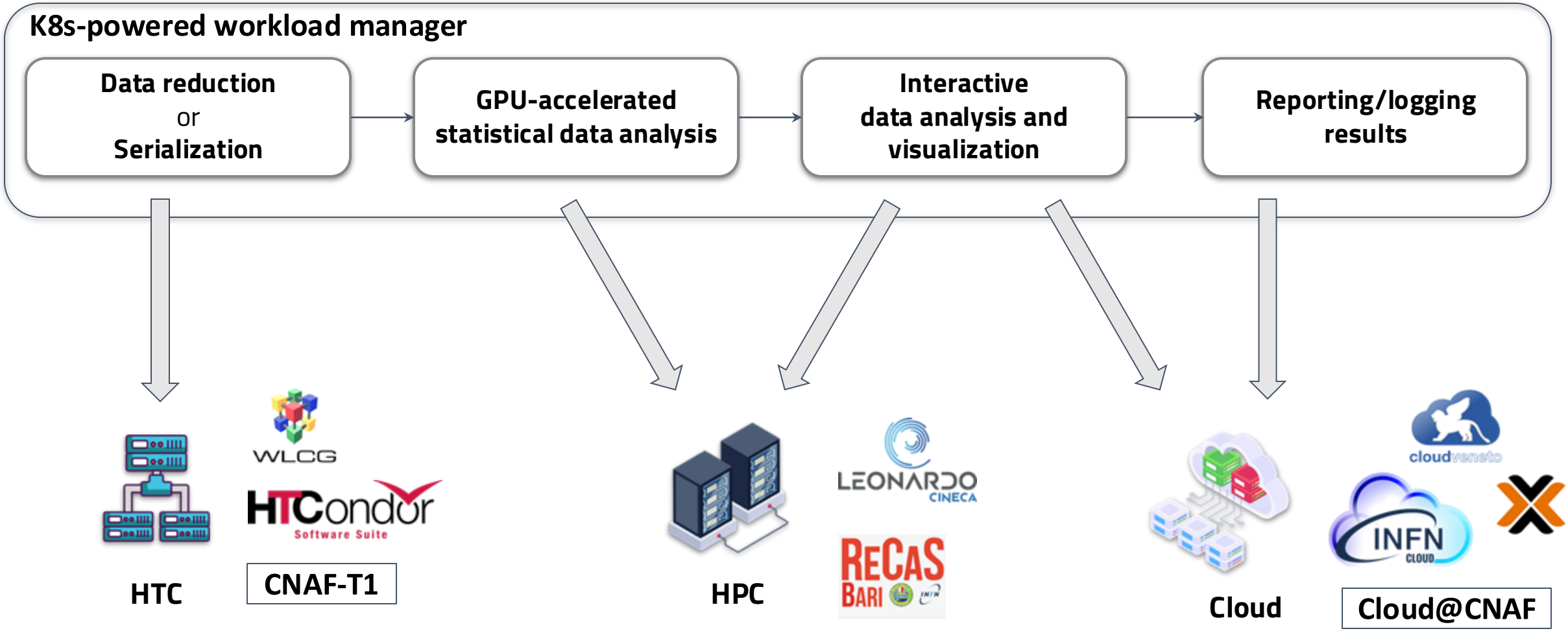
The [Cloud Native Computing Foundation](#) (CNCF) supports and coordinates the development of 200 k8s applications in more than 850 git repositories.

To benefit of K8s for compute-intensive applications, tasks can be organized in job queues and submitted to HPC/HTC centers.





A typical workflow, offloaded



How does it work?

Virtual Kubelet is an open-source project providing an interface between Kubernetes and serverless container platforms.

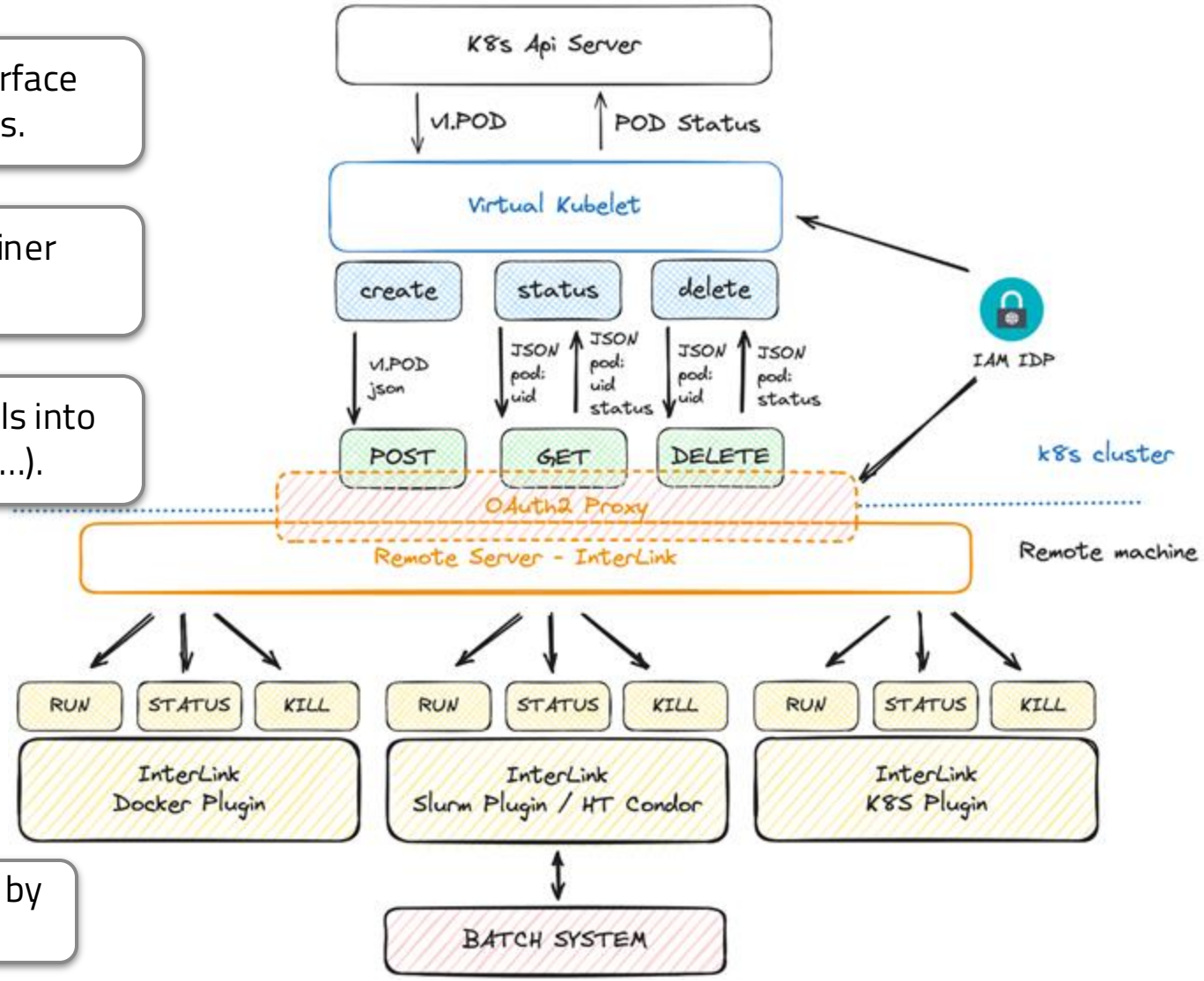
Interlink abstracts the interaction with the remote container runtime to few REST APIs: ***create, status, delete.***

Backend-specific plugins are developed to translate API calls into the "proper language" for each backend (CLI, REST, gRPC…).

OAuth2 and **OpenID Connect** are used to authenticate the Virtual Kubelet towards the remote backend.

The user submitting the job to the Kubernetes Workload manager is propagated as meta-data.

Data, accessed with remote protocols, can be **cached locally** by the compute backend (*e.g.* cvmfs)



Status of the development

	HTC	HPC	Cloud
Submission of self-contained applications	**Supported**	**Supported**	**Supported**
Remote volumes for I/O sandboxes	**PoC**	**PoC**	**Supported**
Network proxying for Jupyter access	**Supported**	**Supported**	**Supported**
Generic network proxying	**Ongoing**		
Multisite filesystem (*e.g. for scripts and notebooks*)	**Tech. Tracking**		
Monitoring and observability	**Supported**		

Scientific use-cases and their workflows

Flash Simulation for High Energy Physics experiments.
Workflow: Train models (HPC), Deploy models in grid-like simulation (HTC), Statistical validation (Cloud)

Quasi-interactive HEP data analysis for Analysis Facilities.
Workflow: An interactive application (Cloud) controls parallel, distributed data processing (HTC) and statistical interpretation (HPC)

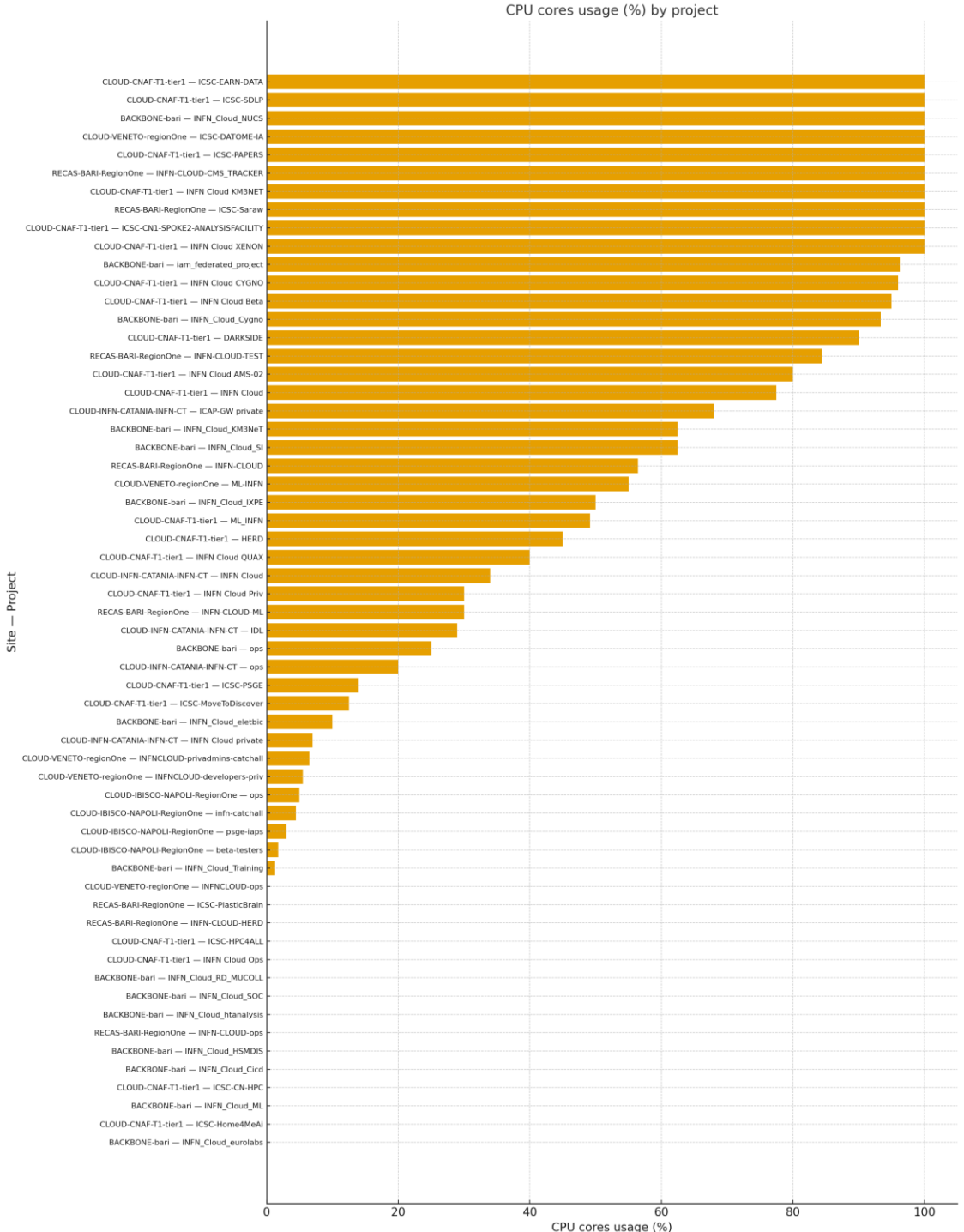
Interactive **GPU algorithm optimization for astrophysical data analysis**.
Workflow: optimize code for astrophysical tasks on GPUs directly through interactive notebook sessions hosted on HPC.

Punti interessanti per la discussione

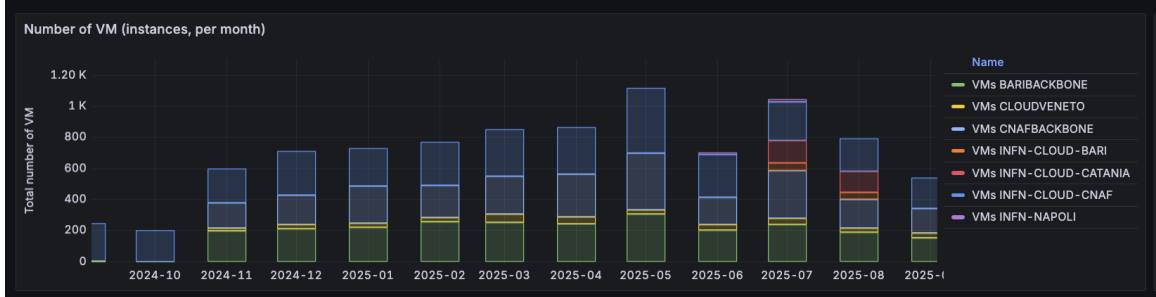
- Batch System:
  - HTCondor?
  - Slurm?
  - Grid?
- Cloud:
  - OpenStack
  - K8s
  - Others?

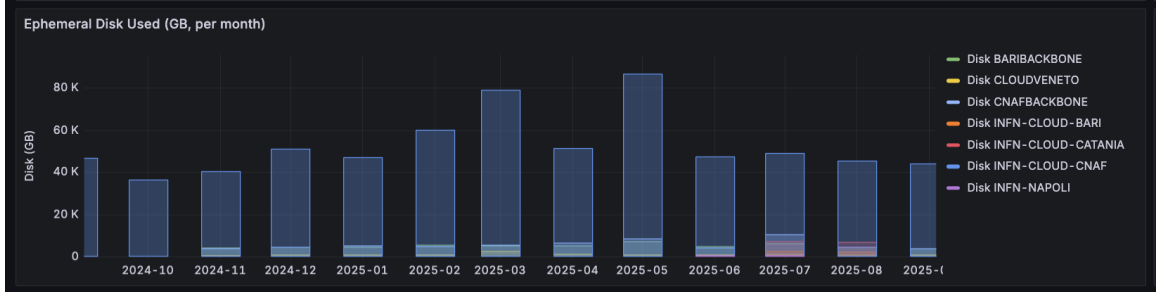
- Come si federano queste risorse?
  - Interlink
  - Gruppo di lavoro k8s
    - ARGOCD
    - OpenTofu/Terraform
  - OpenStack

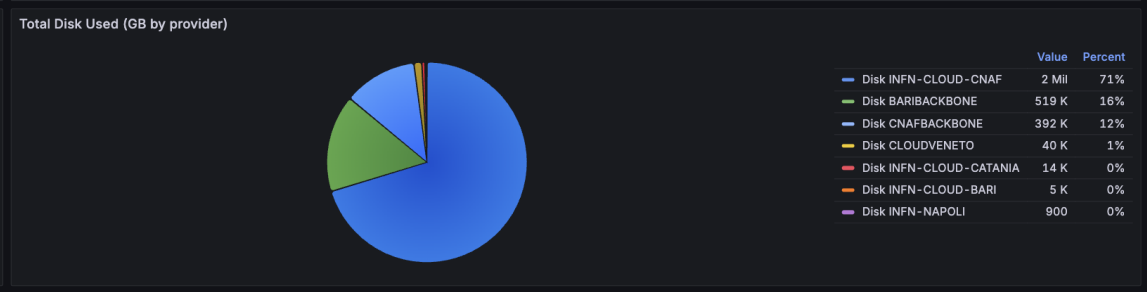
Uso risorse tramite PaaS

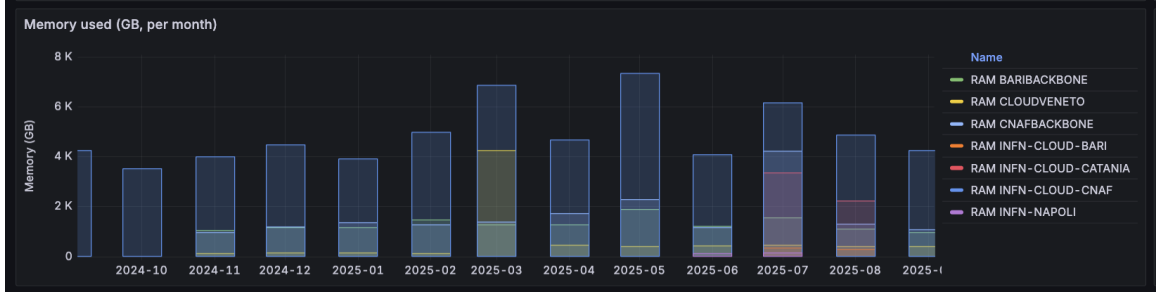


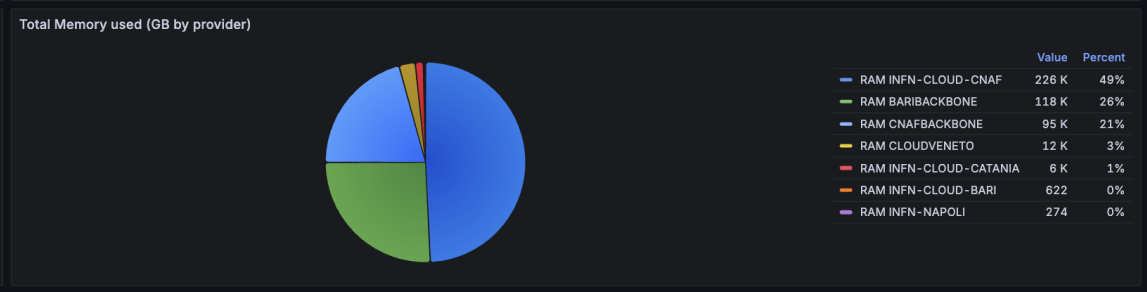
Uso risorse Cloud: Risorse per Sito

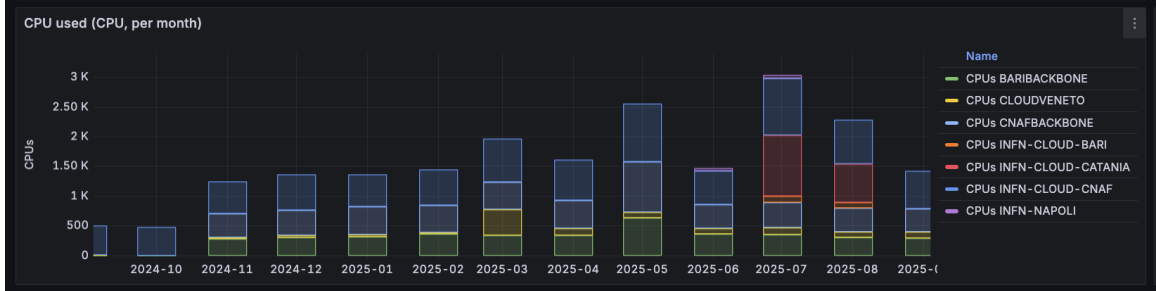


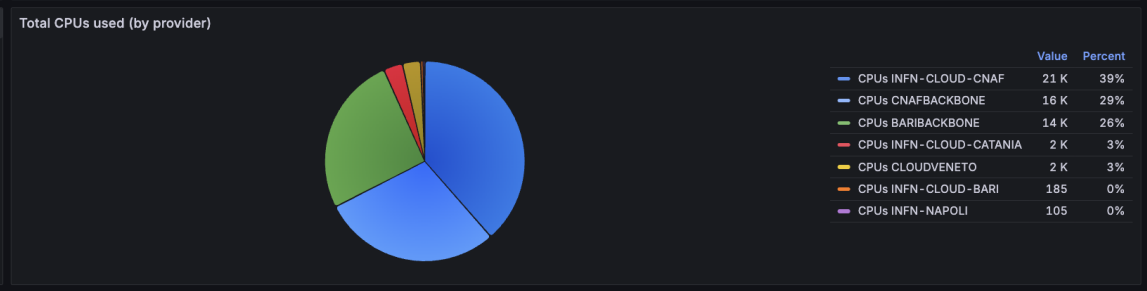


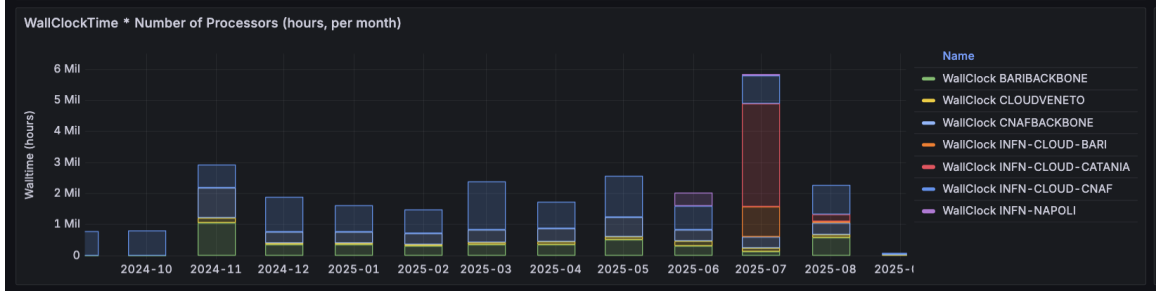


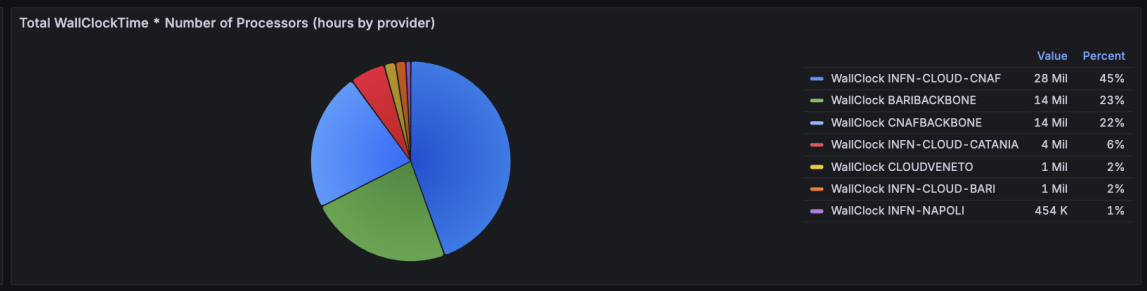




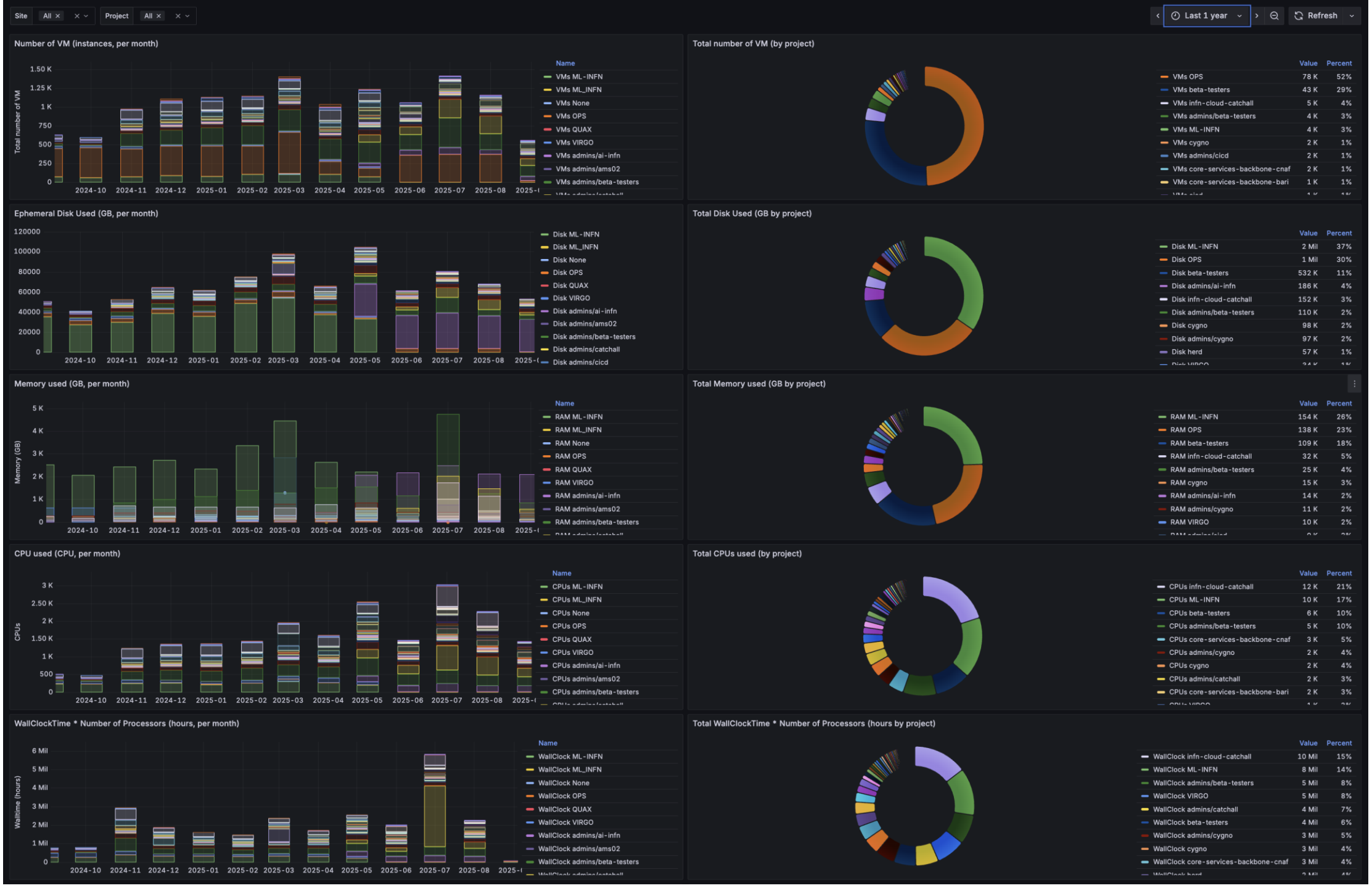








Uso risorse Cloud: Risorse per progetto



Richieste, assegnazioni e costi per Cloud@CNAF

2024

	Crescita netta		Costi		
	CPU (HS06)	Disk (TB-N)	CPU ((k€)	Disk (k€)	Totale (k€)
Cygno	2.800	125	28	13	41
Darkside	200	100	2	10	12
NEWS	50	10	1	1	2
QUAX	400	130	4	13	17
Totale	**3.450**	**365**	**35**	**37**	**72**
Totale effettivo	**2.760**	**365**	**28**	**37**	**65**

Tabella 11: Dettaglio delle risorse richieste su infrastruttura Cloud per gli esperimenti di CSN2.

+ vari esperimenti con pledge HTC migrato su cloud
es. AGATA/GAMMA e N_TOF CSN3
+ VIRGO low latency

2025

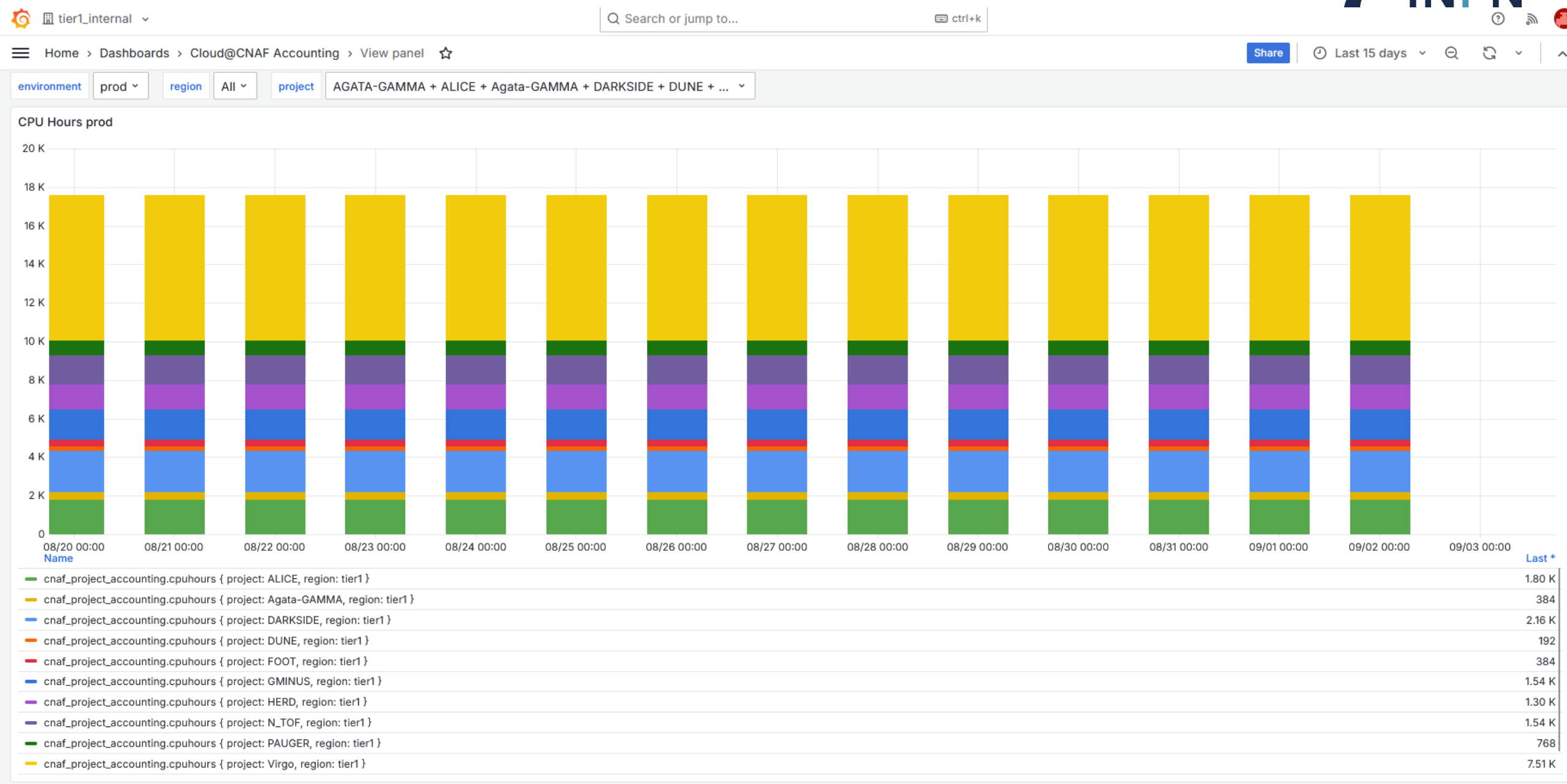
	Crescita netta		Costi (k€)		
	CPU (kHS23)	Disk (PB-N)	CPU	Disk	Totale
Darkside	1,00	0,10	10,0	10,0	20,0
LIMADOU	0,20	0,01	2,0	1,0	3,0
QUAX	0,30		3,0		3,0
Totale	**1.50**	**0,11**	**15,0**	**11,0**	**26,0**
Totale effettivo	**1,20**	**0,11**	**12,0**	**11,0**	**23,0**

Tabella 12: Risorse richieste su infrastruttura Cloud per gli esperimenti di CSN2.

Fino ad ora
costo risorsa cloud = costo risorsa HTC
E' corretto? Non dovremmo almeno togliere l'overlap factor?

Pledge HTC ridotto in caso di risorse assegnate anche su Cloud, es. VIRGO

Cloud@CNAF Usage per exp «pledgiati»



Richieste, assegnazioni e costi per HPC@CNAF (Bubble Terabit)

- Assegnato Pledge DUNE 2024 per GPU
- Assegnate risorse a JLAB12 per prove pledge 2026
  - 8760 GPU hrs/yr
  - 4400HS

- CSN4 2026: Si richiede la disponibilità di 2 Mch sulla partizione CPU delle HPC Bubbles. Lo scopo della richiesta e' di mettere a disposizione dei progetti di CSN4 delle risorse di calcolo su questi sistemi HPC per poter effettuare dei test di ambiente software, di prestazione e di connessione tra nodi.
  - 228 core per 1yr

CSN5

VVIP: - CPU: 3370 kcore-hour (13480 Eur) - GPU: 17500 GPU-hour, GPU H200 NVL (16625 Eur) - Storage: 24 TB, di cui 4 SSD (2400 Eur)
SPRITZ: - CPU: 150 kcore-hour (600 Eur) - GPU: 50 GPU-hour, GPU NVIDIA/AMD (OpenCL support on the cluster) VRAM > 12 GB (4750 Eur)
SEGNAR: - CPU: 150 kcore-hour (600 Eur) - Storage: 10 TB (1000 Eur)
AI_INFN: - CPU: 35 kcore-hour (140 Eur) - GPU: 1000 GPU-hour (950 Eur)
AIM_MIA: - GPU: 9000 GPU-hour, GPU NVIDIA A100 o superiore (8550 Eur)
PLASMA4BEAM2: - CPU: 6000 kcore-hour (2400 Eur)
GEANT4INFN: - CPU: 2000 kcore-hour (8000 Eur) - GPU: 10000 GPU-hour, H200 con 64GB di memoria (9500 Eur) - Storage: 10 TB (1000 Eur)
VITA_5: - CPU: 1400 kcore-hour (5760 Eur) - GPU: 3000 GPU-hour (2850 Eur) - Storage: 300 TB, di cui 100TB per dati sensibili su storage INFN (30000 Eur)
GAP: - CPU: 200 kcore-hour (800 Eur) - Storage: 10 TB (1000 Eur)
BRAINSTAIN: - CPU: 400 kcore-hour (1600 Eur) - GPU: 60000 GPU-hour (57000 Eur)
ARTEMIS: - GPU: 2000 GPU-hour, GPU VRAM GPU >32 GB (1900 Eur) - Storage: 5 TB (500 Eur)
MIRO: - CPU:1500 kcore-hour (6000 Eur) - GPU: 2700 GPU-hour, di cui 2200 su GPU NVIDIA A100 > 32 GB di VRAM (2565 Eur)
SPHINX: - CPU: 100 kcore-hour (400 Eur) - Storage: 0.2 TB (20 Eur)

- Costi «con ammortamento» per CPU/GPU
  - costo per il core-hour di 15 * 8/(24*365*4)=0,0034 euro
  - costo per GPU-hour che è 100keur/(4x365x24)/4= 0,713 euro