Le HPC Bubble di TeRABIT

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

TeRABIT - Work-Package Leaders Meeting
Bari 25-26 Febbraio 2025

Gara "HPC Bubbles"

- Accordo Quadro Nazionale
  - Listino prezzi per nodi + accessori
  - Fondi Terabit, DARE e ICSC
  - 2 anni di validità
    - Indizione: 26/05/23
    - Aggiudicazione: 15/02/24
    - Contratto: 26/04/24
  - Lotto1 (base di gara: € 8.680.000,00 + IVA)
    - CPU, GPU, FPGA
    - Ordini
      - 2 Terabit (Nord/Sud): 4,850,685.90 € + IVA
      - 2 DARE (CNAF, BA) : 1,298,276.60 € + IVA
      - 11 ICSC: 2,818,056.60 € + IVA
  - Lotto2 (base di gara: € 2.459.000,00 + IVA)
    - Storage
    - Ordini
      - 2 Terabit (Nord/Sud): 1,344,708.30 €
      - 2 DARE (CNAF ,BA): 513,807.00 € + IVA
      - 5 ICSC: 1,003,105.00 € + IVA

Quantità nodi con fondi Terabit-ICSC-DARE

	Nodo CPU	Nodo GPU	Nodo FPGA Xilinx	Nodo FPGA Terasic	Nodo storage
BA	24	6	0	0	32
CNAF	26	30	2	2	52
MIB	0	0	2	2	0
NA	18	1	2	0	8
PD	6	6	0	0	0
PI	20	0	0	0	0
RM1	12	0	0	0	0
TO	14	6	0	0	0
LNGS	0	6	0	0	12
CT	12	0	0	0	8
LNF	12	0	0	0	0
LNFESA	8	6	0	0	6
LNL	4	0	0	0	0
MI	4	0	0	0	0
TOTALE	160	61	6	4	118

Core: 30 kcore fisici
Circa 34 HS/core

GPU: 244 NVIDIA H100
34 TFLOPS FP64 ➔ 8.3PFLOPS FP64
40 FPGA
InfiniBAnd 400Gbs

45 PB RAW

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

Offerta Lotto1

- RICCA-IT srl
- Ribasso: 42.87%

Codice	Voce di offerta economica
L1CPU	Nodo CPU
L1GPU	Nodo GPU 4Nvidia H100SXM5
L1FPGA_S	Nodo FPGA – Server
L1FPGA_X	Scheda FPGA Xilinx u55c
L1FPGA_T	Scheda FPGA Terasic p0701
L1SWIB	Switch Infiniband NDR
L1SWETH	Switch Ethernet 10/25
L1QSFPDD	Transceiver QSFP-DD
L1QSFP28	Transceiver QSFP-28
L1CBLMPO24	Cavo da MPO-24 a 8 LC
L1CBLMPO12	Cavo da MPO-12 a 4 LC
L1LCLC	Pannello LC-LC
L1CIB3	Cavo Infiniband NDR 3mt
L1CIB5	Cavo Infiniband NDR 5mt
L1IBTR400	Transceiver 400GB, NDR
L1IBTR800	Transceiver 800GB, 2xNDR
L1IBCBL	Cavo MPO12 to MPO12
L1UPCBL2	Cavo uplink switch 2m
L1UPCBL3	Cavo uplink switch 3m
L1UPCBL5	Cavo uplink switch 5m
L1UPCBL10	Cavo uplink switch 10m
L1ETHCBL2	Cavo LC-LC duplex 2m
L1ETHCBL3	Cavo LC-LC duplex 3m
L1ETHCBL5	Cavo LC-LC duplex 5m
L1ETHCBL10	Cavo LC-LC duplex 10m
L1GPULIC1	Licenza NVIDIA 1GPU 1 anno
L1GPULIC3	Licenza NVIDIA 1GPU 3 anno
L1GPULIC5	Licenza NVIDIA 1GPU 5 anno
L1CPURAM	Raddoppio ram di L1CPU
L1GPURAM	Raddoppio ram di L1GPU

Offerta Lotto2

- E4 Computer Engineering SpA
- Ribasso: 27.45%

Codice	Voce di offerta economica
L2STO	Nodo Storage
L2SWETH	Switch Ethernet 25Gbps
L2UPCBL2	Cavo uplink switch 2m
L2UPCBL3	Cavo uplink switch 3m
L2UPCBL5	Cavo uplink switch 5m
L2UPCBL10	Cavo uplink switch 10m
L2ETHCBL2	Cavo LC-LC duplex 2m
L2ETHCBL3	Cavo LC-LC duplex 3m
L2ETHCBL5	Cavo LC-LC duplex 5m
L2ETHCBL10	Cavo LC-LC duplex 10m
L2SWIB	Switch Infiniband NDR
L2CIB3	Cavo Infiniband NDR 3mt
L2CIB5	Cavo Infiniband NDR 5mt
L2IBTR400	Transceiver 400GB, NDR
L2IBTR800	Transceiver 800GB, 2xNDR
L2IBCBL	Cavo MPO12 to MPO12

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

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?

Stato Consegna e Installazioni - Terabit

- In grave ritardo le FPGA
  - Consegnate solo 2 server al CNAF
  - MIB zero consegne
- In produzione a
  - PD, primo use case (da ICSC) implementato – cluster SLURM
  - TO, RM1 in farm HTC
- In fase di configurazione e messa in produzione
  - CNAF
  - BARI
  - PISA
- A NA in attesa dell'upgrade dell'infrastruttura per poter accendere

		CONSEGNATO	INSTALLATO	COLLAUDATO	FATTURATO	PAGATO
TERABIT						
CNAF L1	4,201,671.30 €					
CNAF_L1_CNAF				mancano FPGA		
CNAF_L1_TO						
CNAF_L1_MIB						
CNAF_L1_PD						
CNAF_L1_PI						
CNAF_L1_RM1						
CNAF L2	888,756.00 €					
BA L1	649,014.60 €					
BARI_L1_BARI						
BARI_L1_NA				mancano fpga		
BA L2	455,952.30 €					
Totale ie	6,195,394.20 €					
Totale ic	**7,558,380.92 €**					
Totale L1 ie	4,850,685.90 €					
Totale L1 ic	5,917,836.80 €					
Totale L2 ie	1,344,708.30 €					
Totale L2 ic	1,640,544.13 €					

Stato Consegna e Installazioni - DARE

- Installate a rack sui siti
- Al CNAF questa settimana configurazione della rete per poter procedere alle installazioni
- Da decidere quanto rilassare la compliance con requirement ISO prima di passare in produzione

		CONSEGNATO	INSTALLATO	COLLAUDATO	FATTURATO	PAGATO
DARE						
CNAF L1	1,050,575.30 €					
CNAF L2	371,376.00 €					
BA L1	247,701.30 €					
BA L2	142,431.00 €					
Totale ie	1,812,083.60 €					
Totale ic	**2,210,741.99 €**					
Totale L1 ie	1,298,276.60 €					
Totale L1 ic	1,583,897.45 €					
Totale L2 ie	513,807.00 €					
Totale L2 ic	626,844.54 €					

Stato Consegna e Installazioni - ICSC

			CONSEGNATO	INSTALLATO	COLLAUDATO	FATTURATO	PAGATO
ICSC							
BA L1	377,994.00 €						
BA L2	189,908.00 €						
TO L1	146,028.00 €						
PI L1	219,042.00 €						
ROMA1 L1	73,014.00 €	6/5					
NA L1	224,399.00 €						
NA L2	185,752.00 €	6/5					
CT L1	214,344.00 €	6/5					
CT L2	186,200.00 €	6/5					
LNGS L1	510,639.30 €						
LNGS L2	288,209.00 €						
LNF L1	227,816.00 €						
LNFESA L1	656,116.30 €						
LNFESA L2	153,036.00 €	6/5					
LNL L1	84,332.00 €	6/5					
MI L1	84,332.00 €	6/5					
Totale ie	3,821,161.60 €						
Totale ic	**4,661,817.15 €**						
Totale L1 ie subito	2,362,034.60 €						
Totale L1 ic subito	**2,881,682.21 €**						
Totale L1 ie 6/5 ie	456,022.00 €						
Totale L1 ie 6/5 ic	**556,346.84 €**						

Modalità di accesso

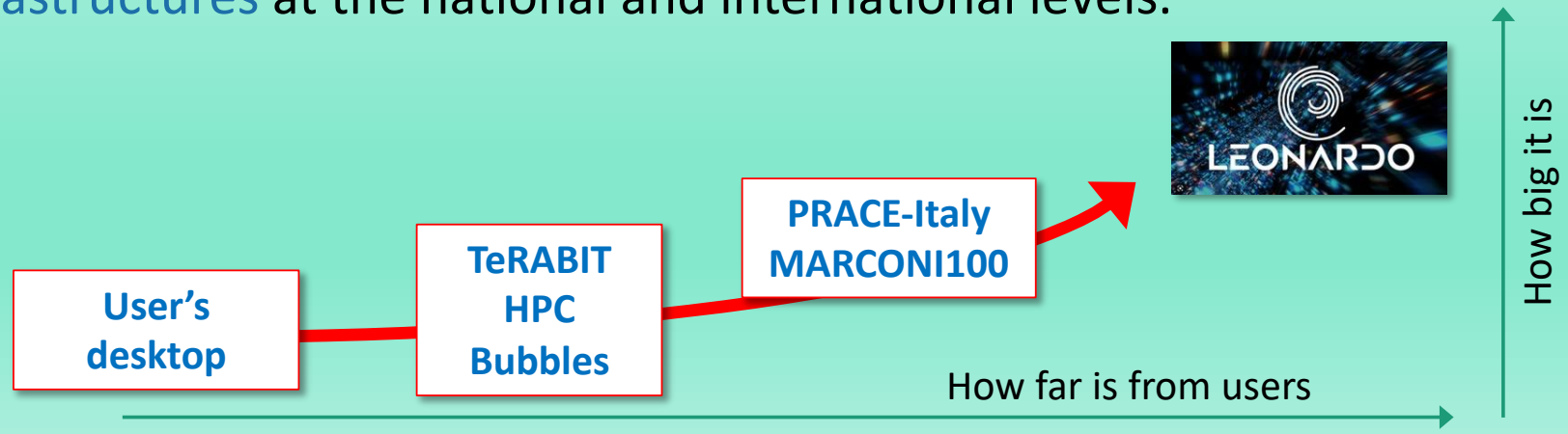
- Batch system (SLURM o HTCondor) – CNAF, TO, BARI, LNF, LNGS, PD, NA
  - sottomissione da macchina locale
    - User interface o jupyter notebook
    - AUTH su user interface via token IAM
    - Possibilità di job interattivi
- Integrazione in Cloud via OpenStack – CNAF, NA
  - AUTH via token IAM
 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

Federazione delle Bubble

HPC Bubbles: how fits the overall model

An example: the “HPC Bubbles”

- As part of the TeRABIT PNRR project, we are implementing several highly innovative concepts:
  - Availability of scalable HPC resources and services on INFN Cloud through Cloud-native interfaces at the IaaS, PaaS and SaaS levels.
  - Strong interaction between network, data and HPC/HTC resources.
  - Communication and federation between the HPC Bubbles and other HPC infrastructures at the national and international levels.

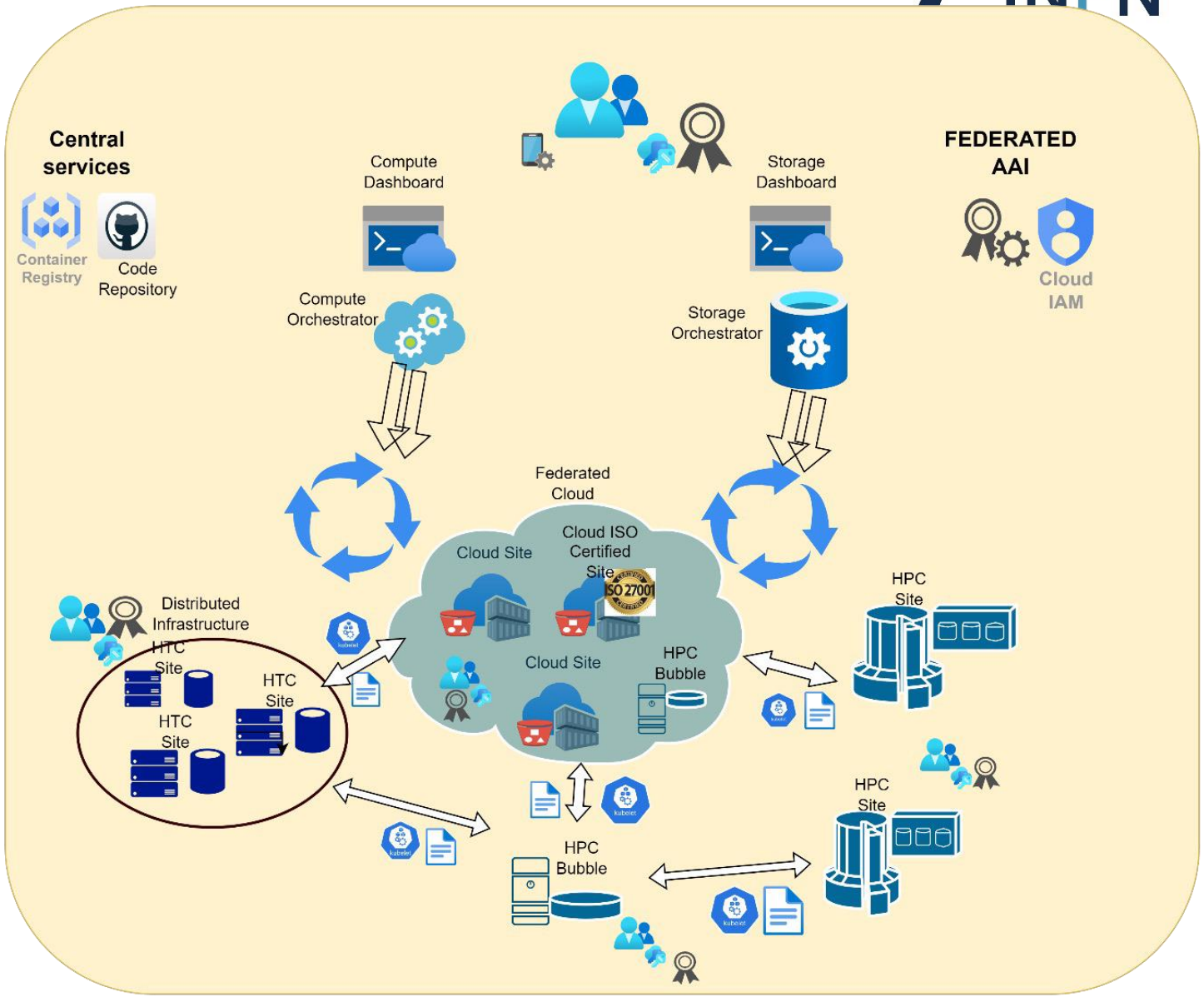


Etherogeneity

Integration of a diverse set of resources, providers, and solutions

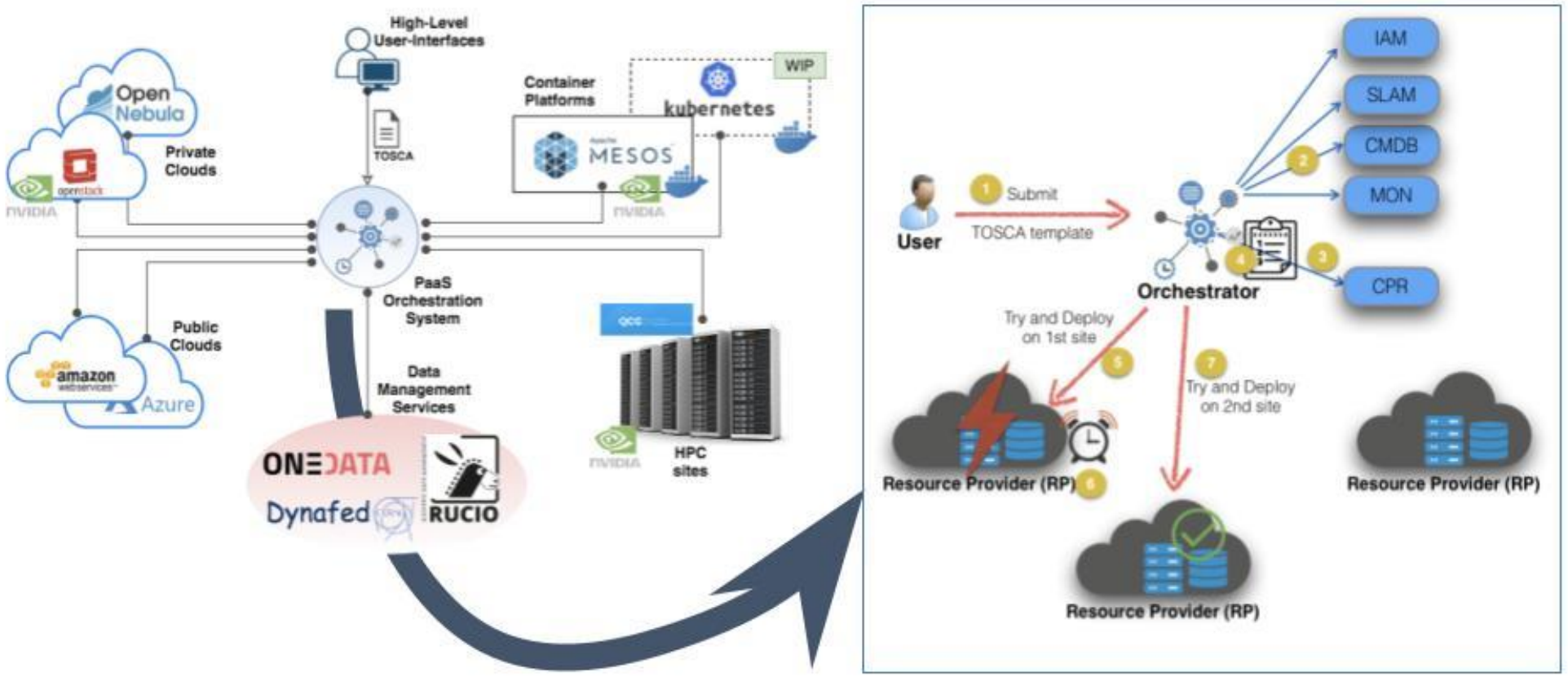
We call it:

Computing *continuum* from Edge, to Cloud, to HPC



Middleware

TOSCA Orchestration



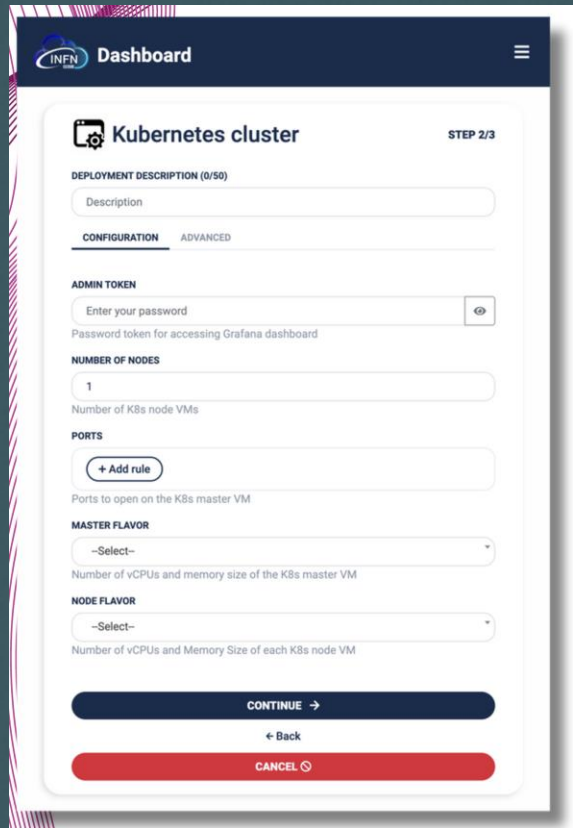
INFN-Cloud services

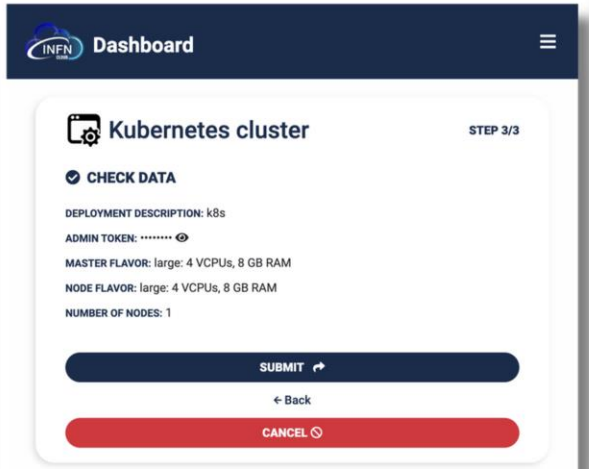
Centralized services (SaaS):
•INFN Cloud Registry service (Harbor)
•INFN Cloud object storage service (based on rados-Gateway)

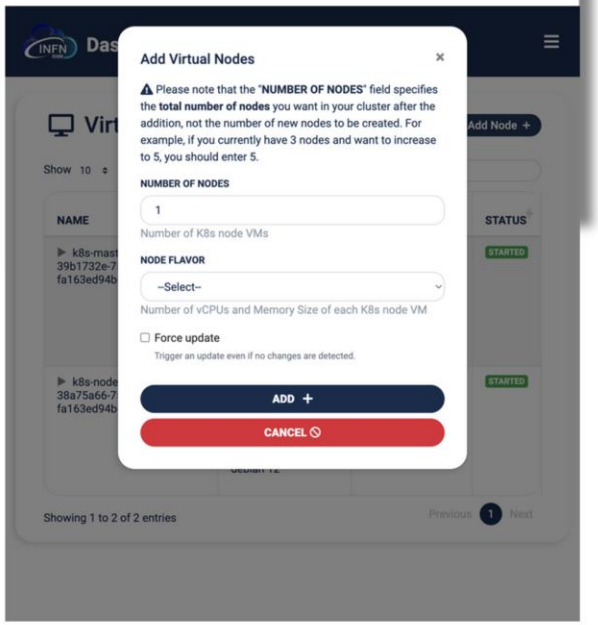
PaaS services:
•Virtual machine
•Docker run
•Docker compose
•Kubernetes cluster
•HTCondor mini
•HTCondor cluster
•Jupyter with persistence for Notebooks
•Jupyter + Matlab (with persistence for Notebooks)
•Spark + Jupyter cluster
•Working Station for CYGNO experiment
•Computational enviroment for AI_INFN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

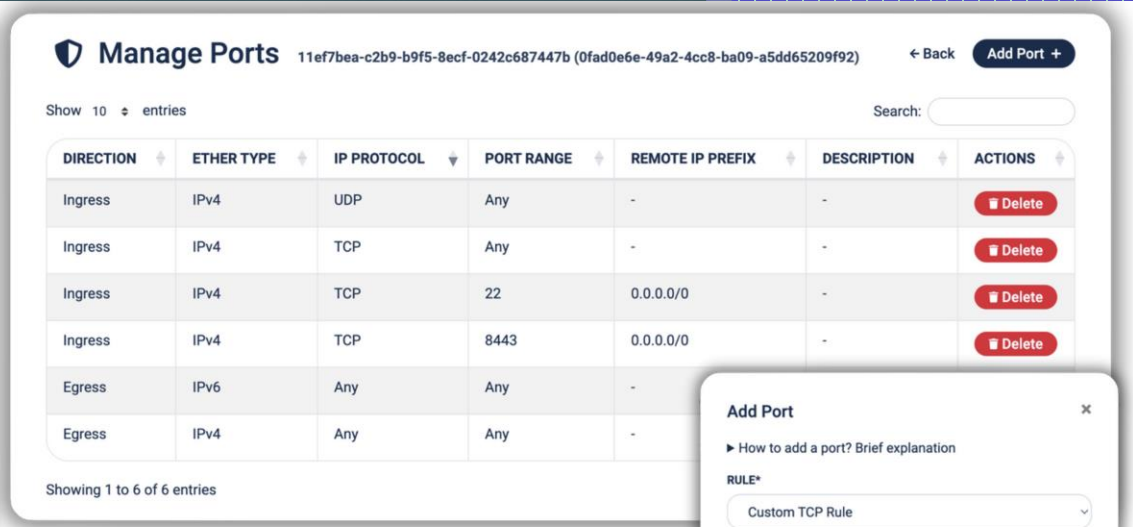
IaaS services:
•Start and Stop
•Hostname choice
•Manage VM ports

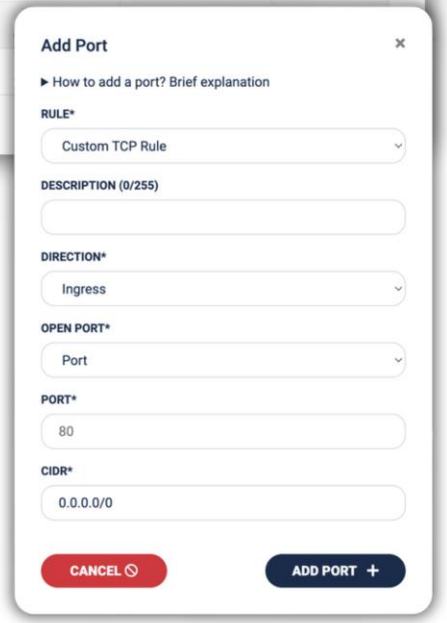
Provisioning di servizi





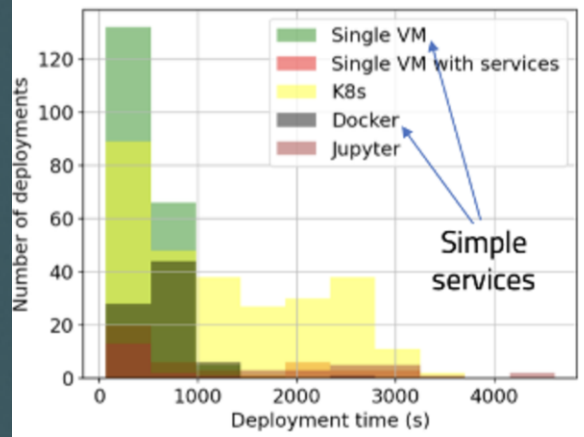


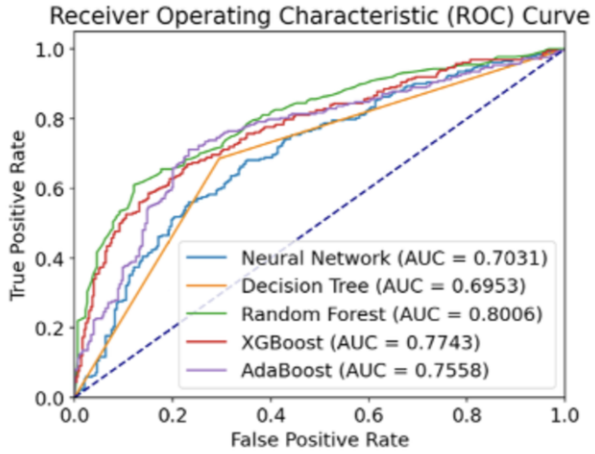




Dashboard
NEW
FUNCTIONALITIES

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)*

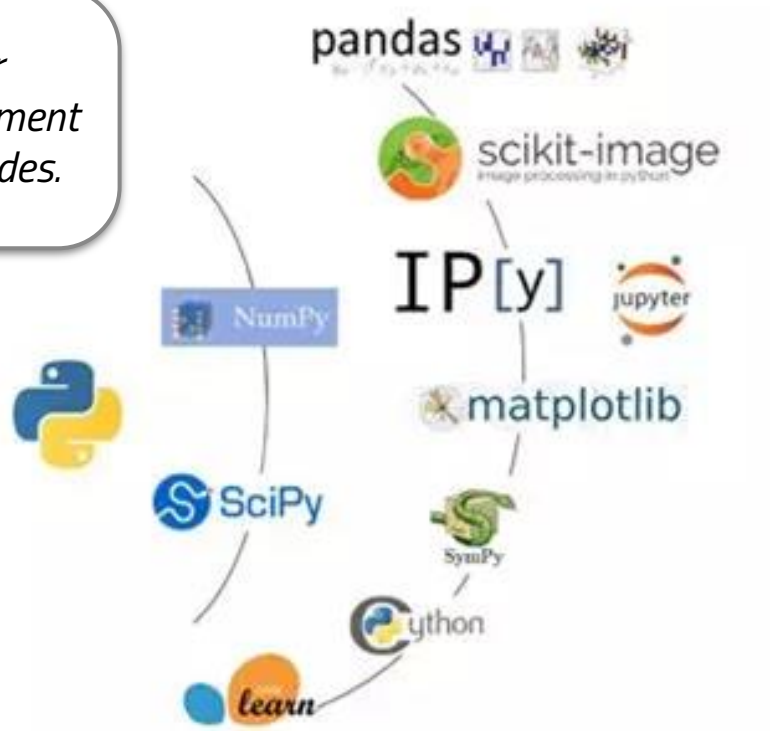
As Python+NumPy emerged as *de facto* standard for Data Science applications in early 2000s, the **Kubernetes APIs have been emerging as the standard for distributed applications**.

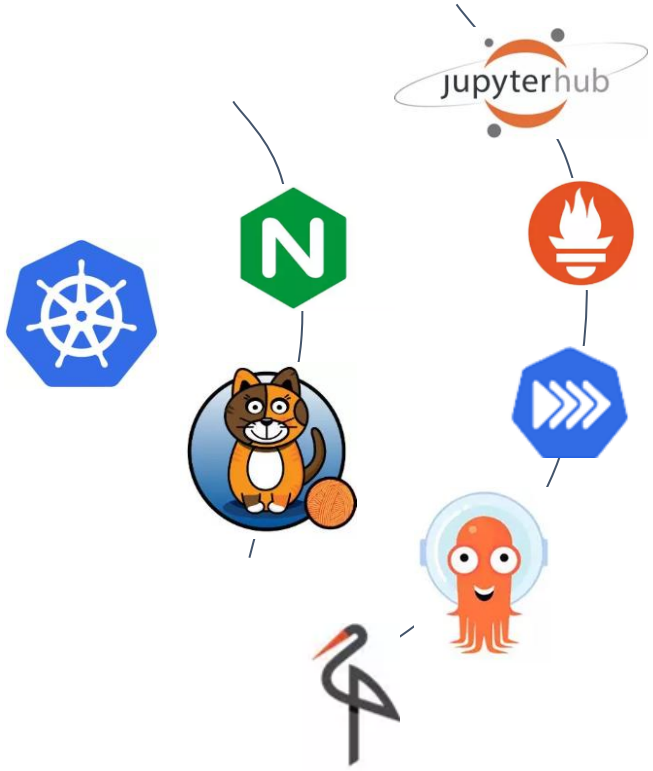
`pip install` → `helm install`

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

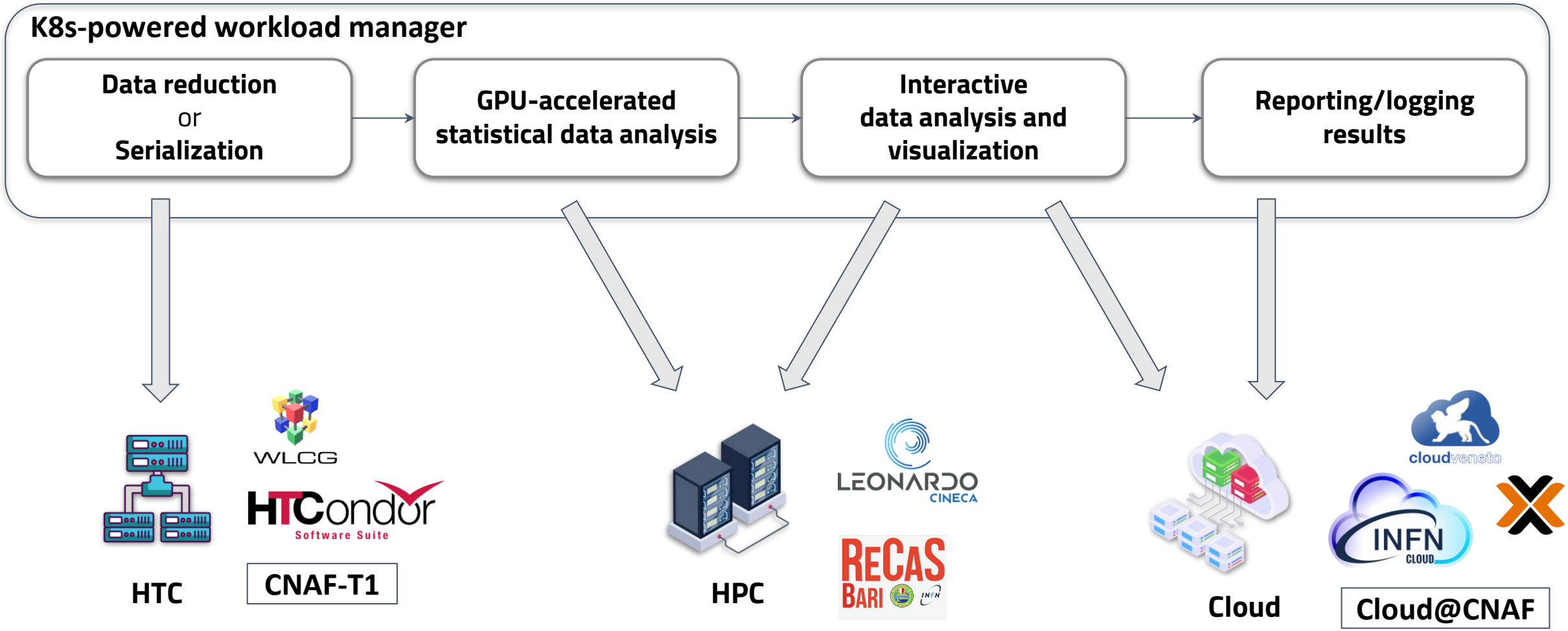
The [Cloud Native Computing Foundation](#) (CNCF) supports and coordinates the development of 200 k8s applications in more than 850 git repositories.

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.

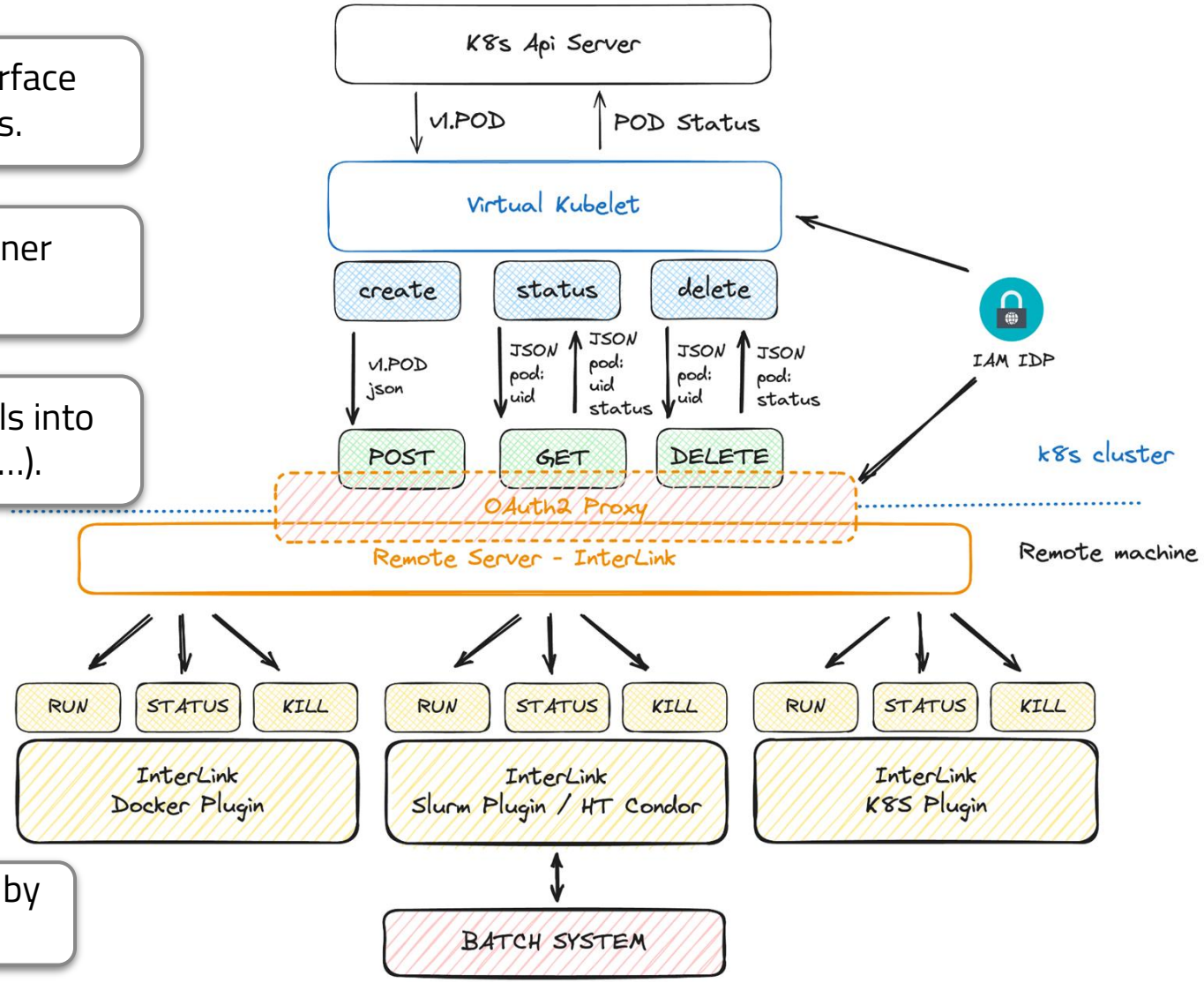
Interlink abstracts the interaction with the remote container runtime to few REST APIs: ***create, status, delete.***

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.

Summary and next steps

- Kubernetes is the standard for distributed applications, including workload managers.
- Offloading compute-intensive containerized workloads to remote dedicated compute backends with a serverless-computing approach enables complex workflows combining:
  - custom services (cloud)
  - High Throughput Computing (data reduction and grid-like jobs)
  - High Performance Computing (statistical data analysis and machine learning)
- Interlink is being deployed as part of the ***Integration PoC*** of ICSC and TERABit to integrate:
  - INFN Cloud instances (docker, Kubernetes)
  - CINECA Leonardo (slurm)
  - HTC resources at CNAF (HTCondor)
- Development is progressing fast and use-cases have been identified in the context of ICSC-Spoke 2 and ICSC-Spoke 3 as demonstrators

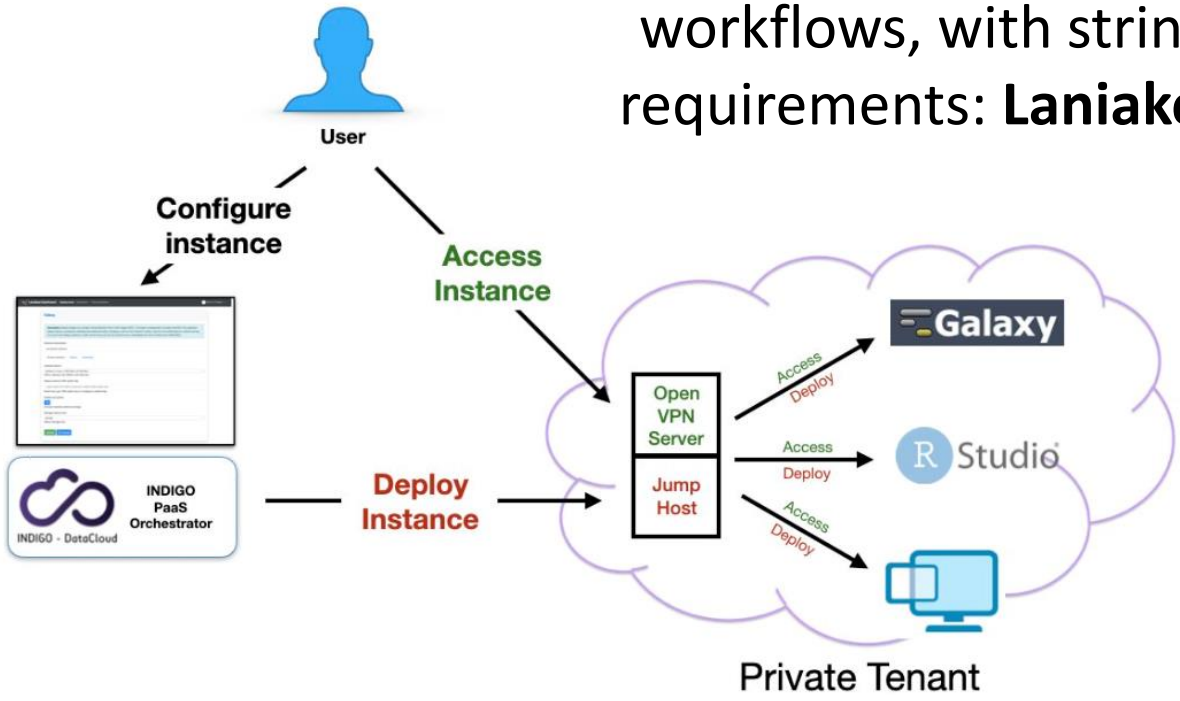
Security: multiple isolation levels

Deployments under VPN

VPN isolated environments - Automatic deployments of virtual environments on private networks.

Isolation is reached using Tenant and security groups properties, granting the access only through VPN authentication.

User authentication to the VPN using the same Laniakea credentials.

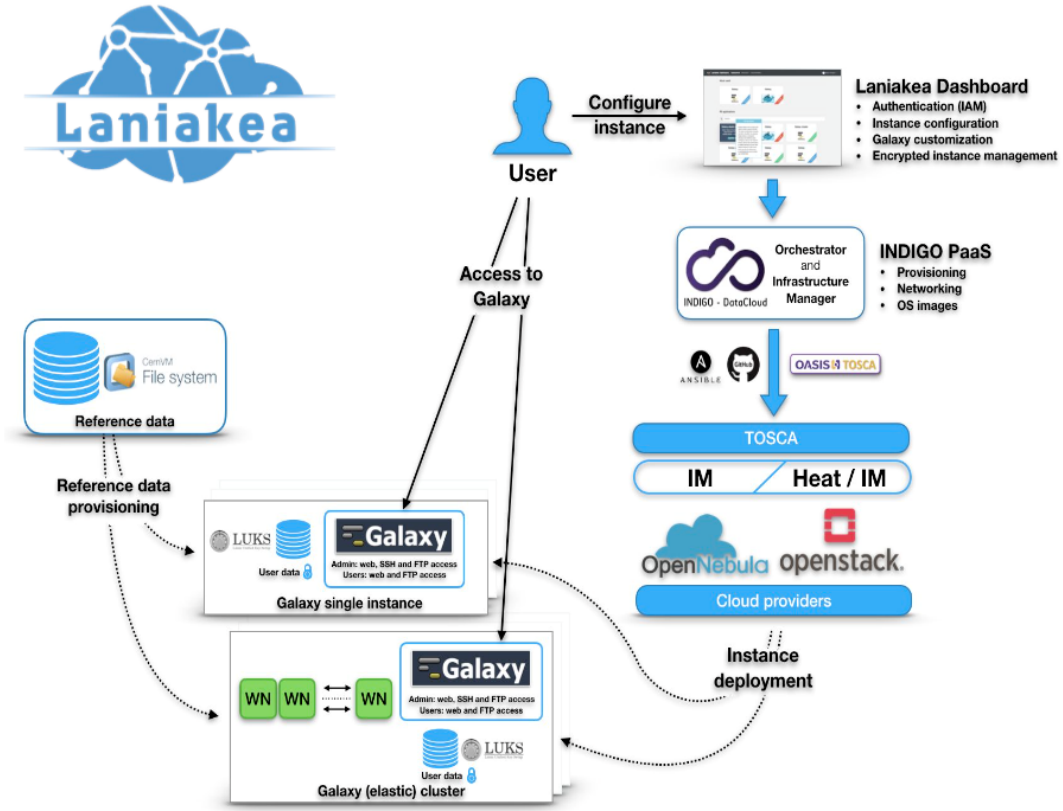


An *example* of the **Service Composition** approach: extending INFN Cloud to support complex workflows, with stringent security requirements: **Laniakea** (Elixir Italy)

41

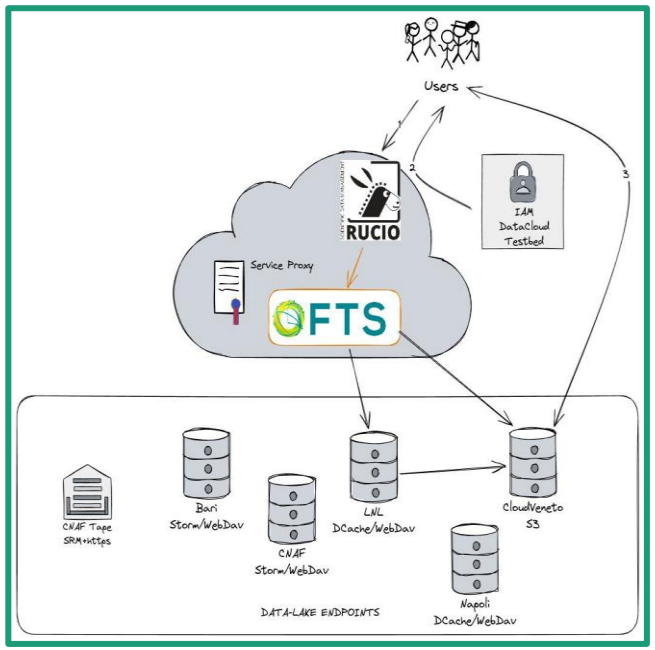
(Some of) The gory details

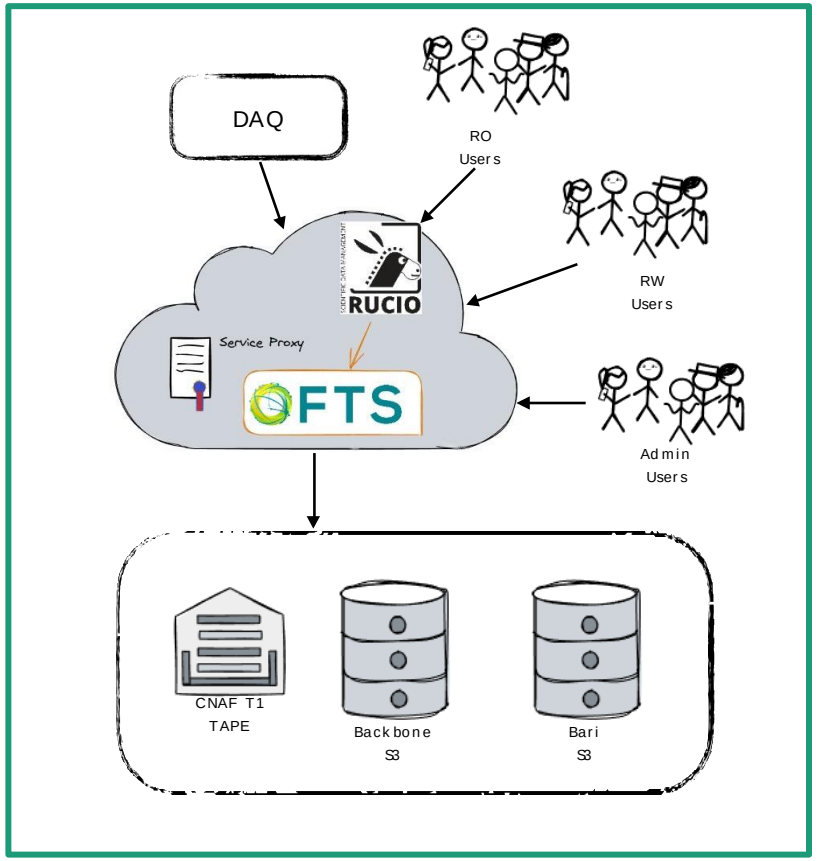
Laniakea architecture

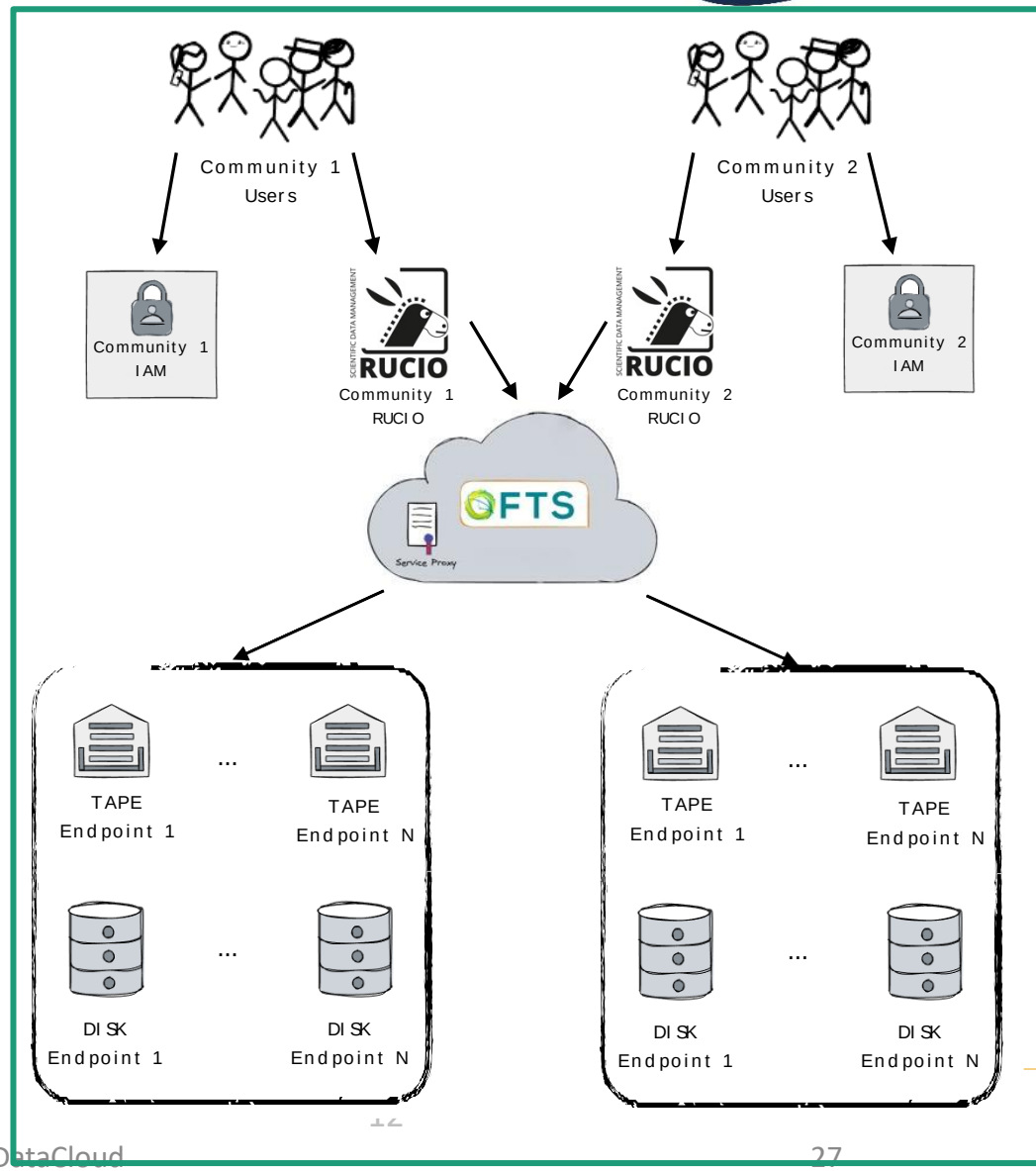


- **Dashboard -** User friendly access to configuration and and launch of a Galaxy instance.
- **IAM -** Authentication and Authorization system.
- **INDIGO PaaS -** Galaxy automatic deployment.
- **Cloud Providers -** (INFN) ReCaS-Bari and others.
- **Persistent storage** - With/without encryption.
- **Reference data availability** - With CERN-VM FileSystem.
- **CLUES -** Elasticity manager.

Data Lake: overview







Recap so far

INFN Cloud is the architectural foundation for the **evolution of the distributed infrastructure** managed and operated by INFN (HPC-BD-AI).

1. This is true for **all** our computing-related engagements with PNRR projects, including ICSC, TeRABIT, DARE, and others.
2. This covers **both hardware acquisitions and the Cloud service portfolio**, in accordance with our service composition architecture. Concretely, this means that we are:
   - Expanding hardware resources across the entire INFN DataCloud;
   - Extending the number of ISO-certified DataCloud regions in Italy;
   - Increasing the solutions offered by INFN Cloud.

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

Installazioni al CNAF

- CPU/GPU/FPGA:
  - Cluster SLURM
  - Metà dei nodi su bare metal
  - Metà dei nodi dentro openstack con schede pci-passthrough
    - Misuriamo le performance nelle due “partizioni”
- STORAGE
  - Fs scratch con gpfs usando gli NVMe sui nodi di calcolo
    - 3*8TB su CPU
    - 2*16 su GPU
    - Via IB
  - FS data via CEPH usando i nodi storage
    - 384*24TB RAW ⇒ 9PB RAW
    - Replica 3 ⇒ 3PB Net
    - Ogni nodo ha 2 nvme da 12TB
      - per cache HDD e per OSD veloci ceph
    - Via Ethernet
- Environment modules per sw
- Tempistiche
  - Manca la configurazione degli switch e siamo pronti ad installare
    - 2 settimane (stima ottimistica)

Use case non Terabit per le HPC Bubbles

Nome progetto	Progetto PNRR	Spoke	Ticket di supporto	Ticket Interni	Quantità di risorse Cloud	Core Cloud	GPU Cloud	Bubble (Y/N)	Storage disco Temporaneo Grid o Cloud (GB)	Storage Disco permanente Grid o Cloud (TB)	Storage Tape (TB)
ML4MMAP	ICSC	2	1603		2 nodi HPC Bubble con kubernetes e grafana presso INFN viene rimandata all'ultimo trimestre 2024	384	8	Y		100	100
ENIPIML	ICSC	2	1611		Il progetto presentato richiede l'allocazione di 1-2 nodi HPC Bubbles	384	8	Y			
HTQCD	ICSC	2			HPC Bubble			Y			4000
ENI-PRED	ICSC	2	1604		HPC Bubble			Y			
UC2.6.3) AI for satellites	ICSC	2	-	-	1 hpc bubble (4 gpu, 64 core, 128 GB ram), 3500 server/hours, 500GB temporary DISK, 1500GB disco permamente, 6 ore per un anno e mezzo	64	4	Y	0.5	1.5	
ANGELS - Smart steering whell for driver safety	ECOSISTER	6	-	-	8 server con 4 GPU, 128 core, 128GB per vcore, 20000 ore_gpu?, sync_and_share service	1024	32	Y		0.25	0.15
Big Data Analysis, Machine Learning and Visualization	ICSC	3	no ticket già in progress		3000 core, 2 GPU, 10TB disco - stesso o quasi progetto di federica Cuna - gestito a Bari - in realtà usano 100core	100	2	Y	1	10	0
Big Data Analysis, Machine Learning and Visualization	ICSC	8	no ticket già in progress		64core + 1 GPU (4TB RAM??? Da investigare)	64	1	Y	1	10	0

Claudio Grandi

Next steps / problemi aperti

- Implementazione della federazione
- Quante risorse «Terabit» possiamo assegnare a progetti «non-Terabit»?
- Quanto possiamo frammentare le HPC bubbles?
  - La modalità più richiesta sembra essere 1-2 nodi in interattivo
  - Modalità con cui partizionare ulteriormente
    - Virtualizzazione delle GPU
- Cosa facciamo se le FPGA continuano a non arrivare?

Special Thanks

- Financial Officer + Amministrazione CNAF
  - Martina Allegro, Giulia Grandi, Sara Haghshenas
- Infrastructure Manager
  - Silvia Calegari
- Ufficio RUP per l'accordo quadro
  - Giulia Grandi, Michela Pischedda, Marta Chiodi, Andrea Chierici
- Tutti i RUP e i DEC delle sedi INFN coinvolte
- Claudio Grandi