

Stato dell'infrastruttura per il calcolo scientifico dell'INFN

DataCloud Team

Agenda

- Stato delle gare
 - Definizione dell'architettura per usare le nuove risorse
- Stato delle risorse in produzione
 - Uso delle risorse (Accounting)
 - Evoluzione degli utenti e delle comunità
- Stato dei servizi di Piattaforma
- Stato degli user services
- Stato delle attività di sviluppo
- Stato del PoC con il CINECA (ICSC&Terabit)
- Stato dell'infrastruttura EPIC
- Conclusioni e prospettive

Stato delle gare

- Le risorse dell'accordo quadro Terabit, sono quasi tutte consegnate
 - Le FPGA sono state piuttosto problematiche
- Le bubble nei siti hanno fornito la possibilità di creare cluster HPC in molti siti INFN
 - Diversi contributi fra presentazioni e poster che raccontano l'esperienza fatta
 - Crediamo sia necessario continuare in tale direzione cercando di mettere a fattor comune expertise e know-how dei vari gruppi

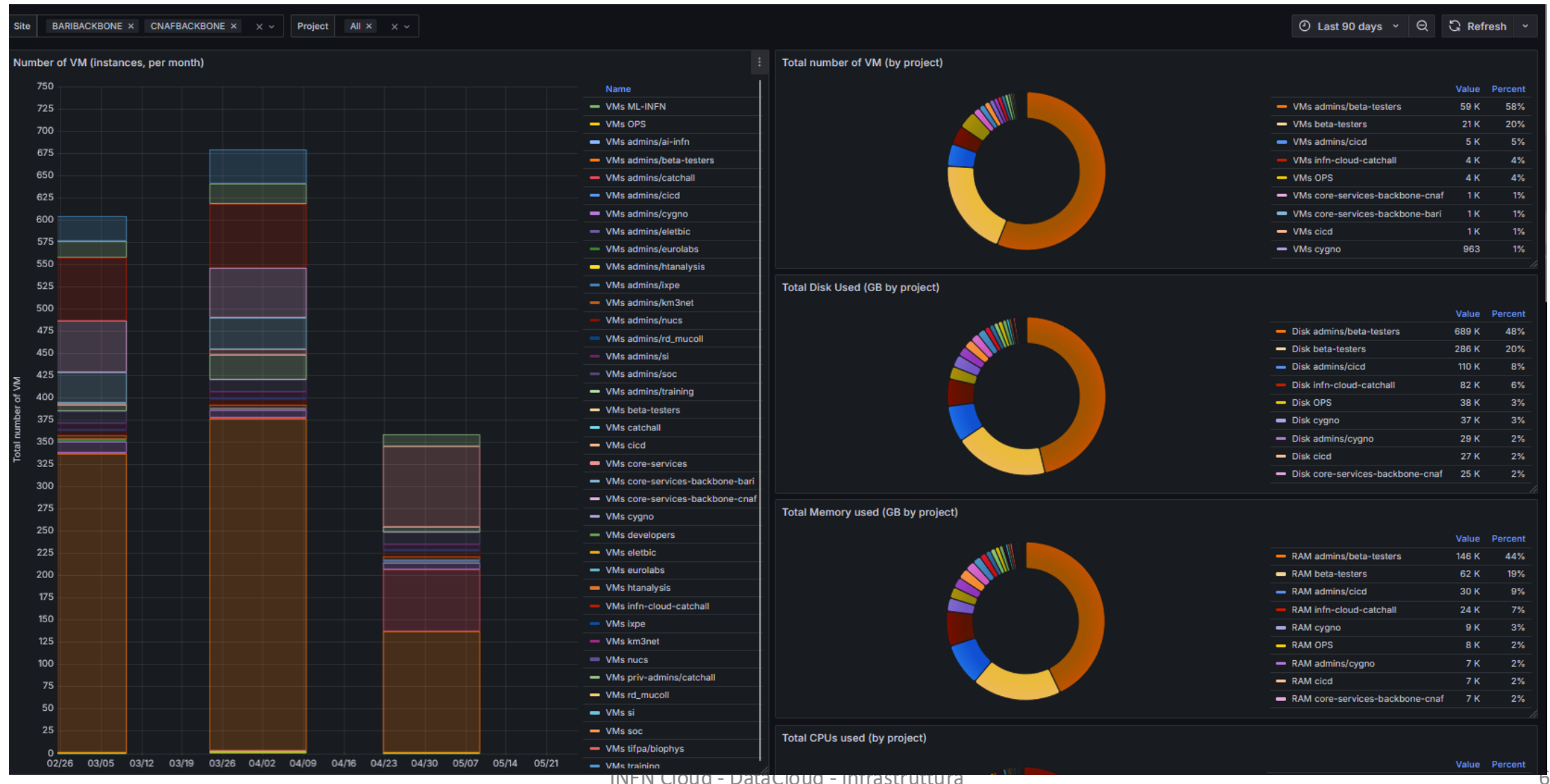
Stato delle gare (2)

- È in corso di approvazione la gara dell'Accordo Quadro sullo storage:
 - Lotto1 => «Tier2» storage (Low & High density)
 - Max € 4.514.671,00 (IVA Compresa)
 - Circa 60PByte
 - Lotto2 => «CEPH» storage per il cloud
 - Con dischi sia SSD che magnetici
 - Max € 2.196.671,00 (IVA Compresa)
 - Si prevede la possibilità di acquistare un cluster compatibile con una configurazione di CEPH con raid-sw

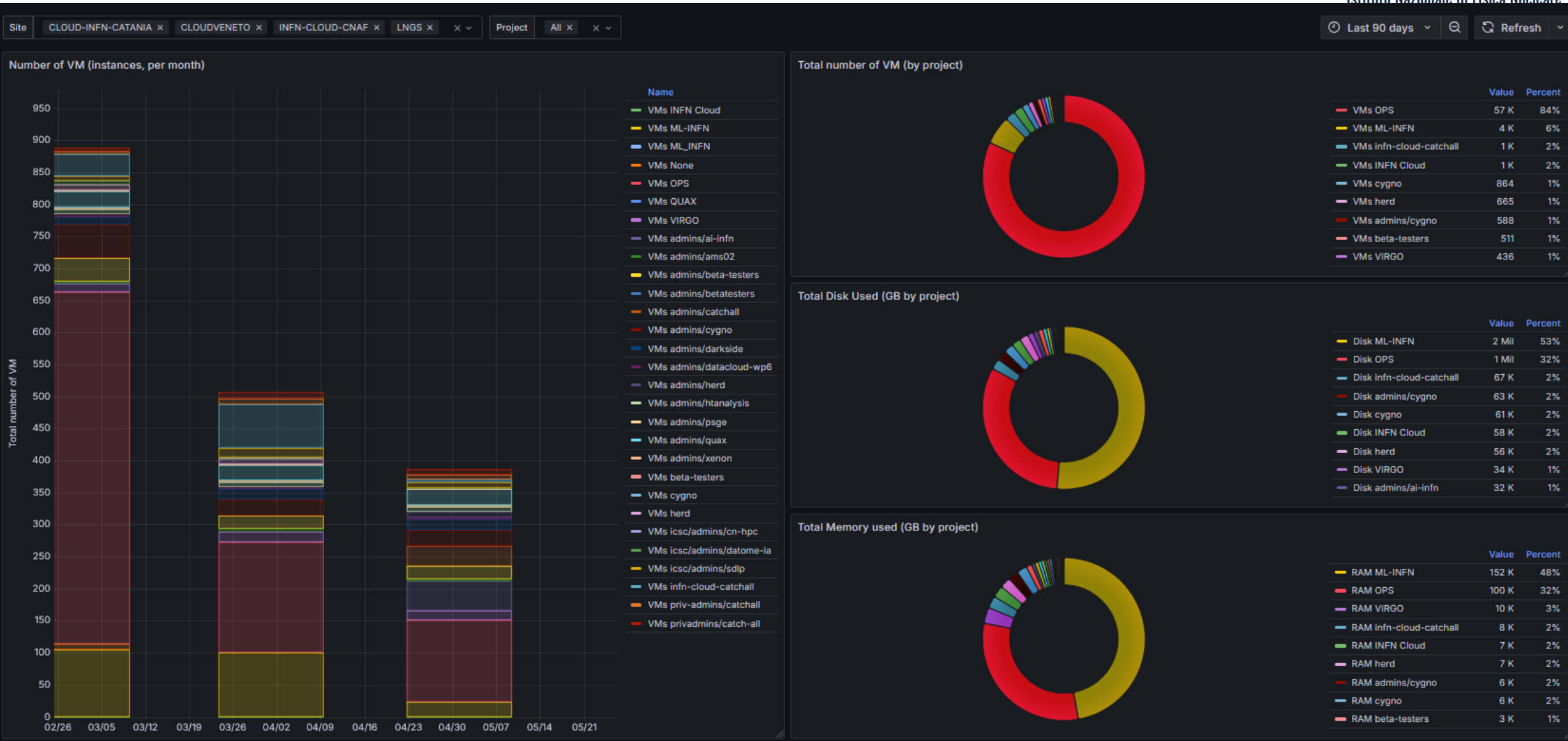
Stato dell'infrastruttura in produzione

- Backbone (Bari/CNAF)
 - 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 375 vm attive, più della metà delle quali usate per servizi, R&D, dal SOC ecc.

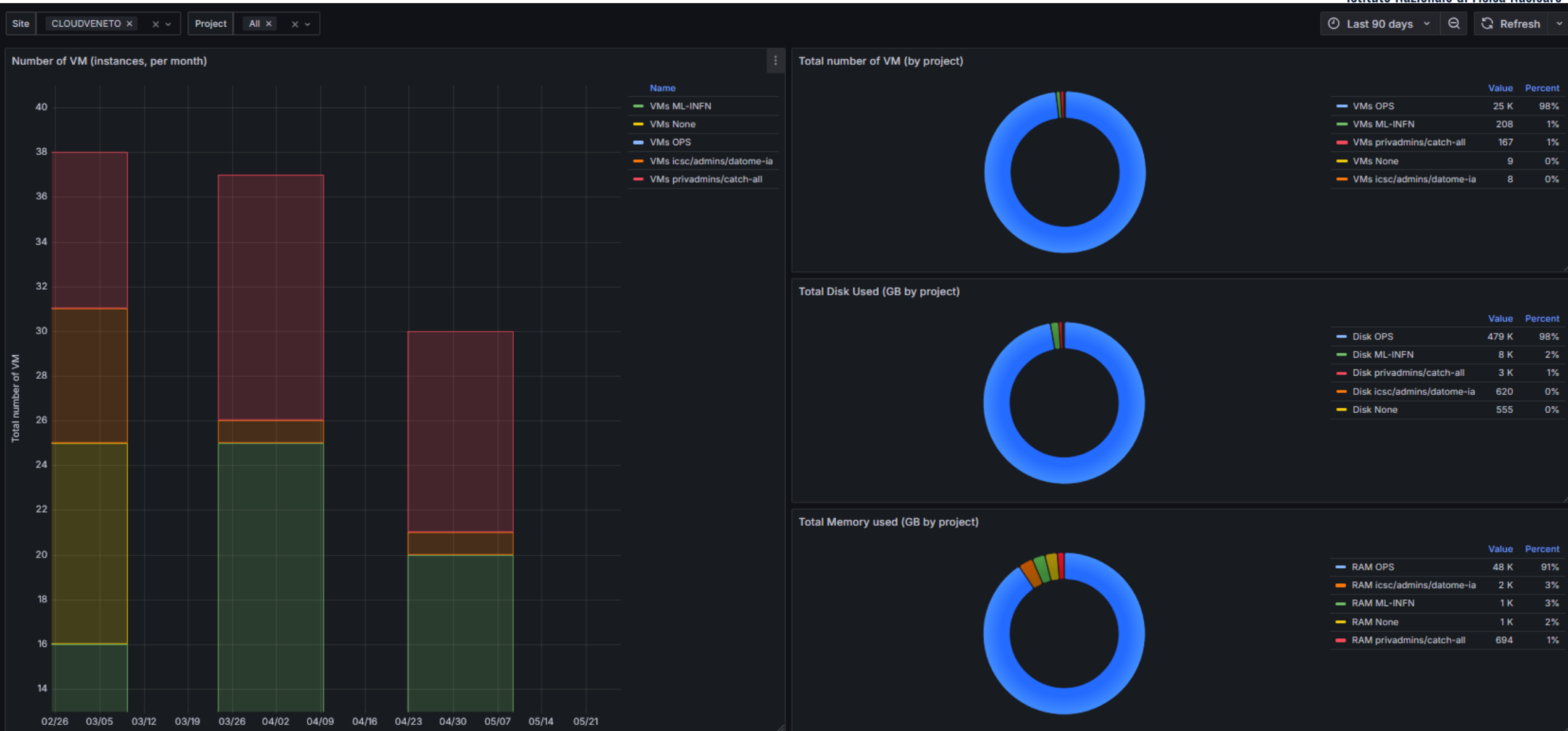
Stato dell'infrastruttura in produzione



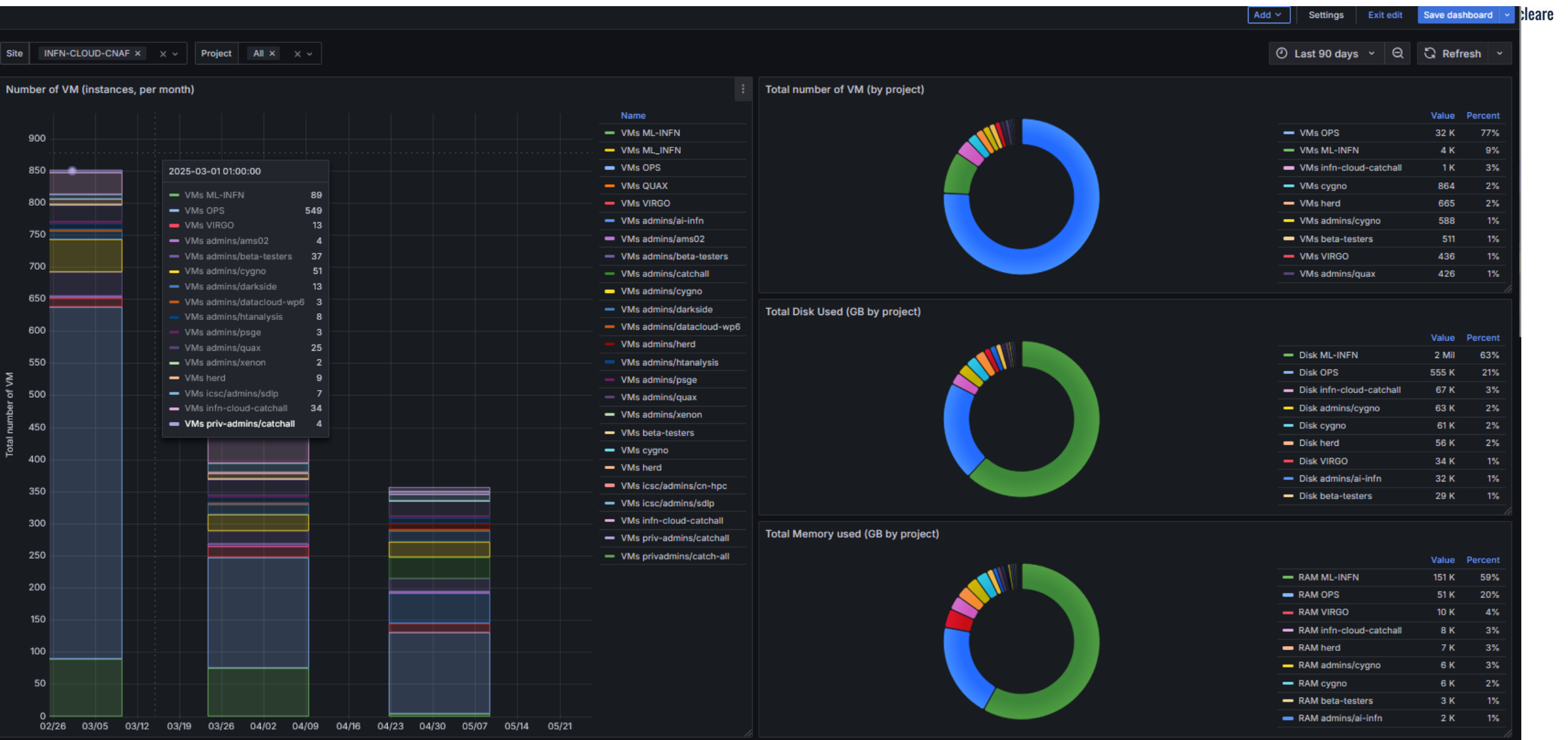
Stato dell'infrastruttura in produzione



Stato dell'infrastruttura in produzione



Stato dell'infrastruttura in produzione



Stato dell'infrastruttura in produzione

- **"name": "CLOUD-IBISCO-NAPOLI",**
- "idp_name": "icsc-cloud",
- "name": "icsc/admins/saraw",
- "name": "icsc/priv-admins/psge-iaps",
- "name": "icsc/admins/datome-ia",
- "name": "icsc/priv-admins/icap-gw",
- "name": "icsc/admins/idl",
- "name": "icsc/priv-admins/plasticbrain",
- "name": "icsc/admins/sdlp",
- "name": "icsc/admins/cn-hpc",
- "idp_name": "infn-cloud",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygn0",
- "name": "admins/herd",
- "name": "admins/ai-infn",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "name": "infn-catchall",
- "name": "beta-testers",
- "name": "ops",
- "name": "psge-iaps",

Stato dell'infrastruttura in produzione

- **"name": "CLOUD-CNAF-T1",**
- "idp_name": "icsc",
- "name": "icsc/admins/saraw",
- "name": "icsc/priv-admins/psge-iaps",
- "name": "icsc/admins/datome-ia",
- "name": "icsc/priv-admins/icap-gw",
- "name": "icsc/admins/idl",
- "name": "icsc/priv-admins/plasticbrain",
- "name": "icsc/admins/sdlp",
- "name": "icsc/admins/cn-hpc",
- "idp_name": "inf-n-cc",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygno",
- "name": "admins/herd",
- "name": "admins/ai-inf-n",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "name": "ICSC-CN1-SPOKE2-ANALYSISFACILITY",
- "name": "HERD",
- "name": "ICSC-SDLP",
- "name": "INFN Cloud AMS-02",
- "name": "INFN Cloud XENON",
- "name": "ICSC-PSGE",
- "name": "INFN Cloud Ops",
- "name": "DARKSIDE",
- "name": "ICSC-CN-HPC",
- "name": "INFN Cloud CYGNO",
- "name": "INFN Cloud Priv",
- "name": "ML-INFN",
- "name": "INFN Cloud QUAX",
- "name": "INFN Cloud",
- "name": "INFN Cloud Beta",
- "name": "tier1",

Stato dell'infrastruttura in produzione

- "name": "BACKBONE",
- "idp_name": "inf-n-cloud",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygno",
- "name": "admins/herd",
- "name": "admins/ai-inf-n",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "name": "INFN_Cloud_eurolabs",
- "name": "INFN_Cloud_KM3NeT",
- "name": "INFN_Cloud_ML",
- "name": "ops",
- "name": "INFN_Cloud_HSMDIS",
- "name": "INFN_Cloud_IXPE",
- "name": "INFN_Cloud_SI",
- "name": "INFN_Cloud_SOC",
- "name": "INFN_Cloud_NUCS",
- "name": "INFN_Cloud_eletbic",
- "name": "INFN_Cloud_Cicd",
- "name": "INFN_Cloud_RD_MUCOLL",
- "name": "iam_federated_project",
- "name": "INFN_Cloud_Cygno",
- "name": "INFN_Cloud_Training",
- "name": "INFN_Cloud_htanalysis",
- "name": "bari",

Stato dell'infrastruttura in produzione

- "name": "RECAS-BARI",
- "idp_name": "inf-n-cloud",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygn0",
- "name": "admins/herd",
- "name": "admins/ai-inf-n",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "idp_name": "icsc",
- "name": "icsc/admins/saraw",
- "name": "icsc/priv-admins/psge-iaps",
- "name": "icsc/admins/datome-ia",
- "name": "icsc/priv-admins/icap-gw",
- "name": "icsc/admins/idl",
- "name": "icsc/priv-admins/plasticbrain",
- "name": "icsc/admins/sdlp",
- "name": "icsc/admins/cn-hpc",
- "name": "INFN-CLOUD-ops",
- "name": "INFN-CLOUD-HERD",
- "name": "INFN-CLOUD-TEST",
- "name": "INFN-CLOUD",
- "name": "INFN-CLOUD-ML",
- "name": "INFN-CLOUD-CMS_TRACKER",
- "name": "ICSC-Saraw",
- "name": "ICSC-PlasticBrain",
- "name": "RegionOne",

Stato dell'infrastruttura in produzione

- "name": "CLOUD-INFN-CATANIA",
- "idp_name": "inf-n-cloud",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygn0",
- "name": "admins/herd",
- "name": "admins/ai-inf-n",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "idp_name": "icsc",
- "name": "icsc/admins/saraw",
- "name": "icsc/priv-admins/psge-iaps",
- "name": "icsc/admins/datome-ia",
- "name": "icsc/priv-admins/icap-gw",
- "name": "icsc/admins/idl",
- "name": "icsc/priv-admins/plasticbrain",
- "name": "icsc/admins/sdlp",
- "name": "icsc/admins/cn-hpc",
- "name": "ops",
- "name": "INFN Cloud private",
- "name": "INFN Cloud",
- "name": "ICAP-GW private",
- "name": "IDL",
- "name": "INFN-CT",

Stato dell'infrastruttura in produzione

- "name": "CLOUD-VENETO",
- "idp_name": "infn-cc",
- "name": "priv-admins/beta-testers",
- "name": "developers",
- "name": "admins/cms-tracker",
- "name": "admins/cicd",
- "name": "admins/training",
- "name": "admins/ixpe",
- "name": "admins/km3net",
- "name": "admins/eurolabs",
- "name": "admins/eletbic",
- "name": "admins/nucs",
- "name": "admins/rd_mucoll",
- "name": "admins/soc",
- "name": "admins/si",
- "name": "admins/hsmdis",
- "name": "priv-admins/catchall",
- "name": "admins/xenon",
- "name": "admins/psge",
- "name": "admins/darkside",
- "name": "admins/ams02",
- "name": "admins/quax",
- "name": "admins/htanalysis",
- "name": "admins/cygno",
- "name": "admins/herd",
- "name": "admins/ai-infn",
- "name": "admins/beta-testers",
- "name": "ops",
- "name": "admins/catchall",
- "idp_name": "icsc",
- "name": "icsc/admins/saraw",
- "name": "icsc/priv-admins/psge-iaps",
- "name": "icsc/admins/datome-ia",
- "name": "icsc/priv-admins/icap-gw",
- "name": "icsc/admins/idl",
- "name": "icsc/priv-admins/plasticbrain",
- "name": "icsc/admins/sdlp",
- "name": "icsc/admins/cn-hpc",
- "name": "INFNCLOUD-developers-priv",
- "name": "INFNCLOUD-privadmins-catchall",
- "name": "ICSC-DATOME-IA",
- "name": "INFNCLOUD-ops",
- "name": "ML-INFN",
- "name": "regionOne",

Stato dei Servizi

- servizi SaaS attivi:
 - s3
 - docker image repository
 - notebook aaS,
 - monitoring aaS (healthchecks)
 - ovviamente iam e dashboard + orchestratore (con supporto a ICSC/terabit/etc)
 - cvmfs
 - fts
 - servizio di backup per utenti su S3
 - iam e dashboard per ICSC
 - docker image repository per ICSC

INDIGO IAM

- 6 nuove release IAM fino a 1.11.1 (ultima rilasciata) con focus in particolare su: miglioramento interfaccia e funzionalità di gestione di una VO, aumento delle informazioni contenute nell' AUDIT log, primo supporto MFA.
- Principale topic prossima release: supporto completo a MFA anche in caso di login con provider remoti o certificati x509
- Continuano gli sviluppi in parallelo: non salvataggio degli accesso token nel database, abbandono dipendenze obsolete (nuova dashboard React già in fase avanzata di realizzazione e migrazione entro 2026 alle ultime librerie Spring Security)

StoRM WebDAV e StoRM Tape REST API

- 8 nuove release StoRM WebDAV da 1.4.2 a 1.8.1. Già in uso al T1 la versione RC di 1.9.0 vicina al rilascio definitivo.
- Focus di queste release: aggiornamenti dipendenze, miglioramento performance, bug fixing e full support di storage area collegate a tape system
 - aggiornamento dipendenze a Spring Boot (v3.4.5) e Jetty (v12) usato come web server, Bouncy Castle (v1.80) e CANL (v2.8.3)
 - re-direzione trasferimenti GET a NGINX server per migliorare performance
 - riconoscimento e blocco accessi ai file che sono presenti solo su tape e non su buffer disco (stub files)
- StoRM Tape: miglioramenti policy di accesso con supporto a gruppi OIDC tramite regole OPA (Open Policy Agent)

PaaS – Maintenance (1)

- Passaggio da SLAT + CMDB + CIP a Federation Registry + Feeder
- Integrazione dell'Orchestratore con nuovo Federation Registry
- Aggiunti servizi e migliorata usabilità della dashboard:
 - Per gli utenti
 - Scelta del provider inserito come primo step nella procedura di sottomissione di un deployment.
 - Possibilità di retry se deployment è fallito
 - Force delete di un deployment
 - Per gli admin
 - Reset dei deployment inconsistenti
 - Accesso a tutte le informazioni del deployment, log compresi, filtrabili per gruppo e stato
 - Cancellazione deployment di altri utenti
 - Accesso al numero dei servizi PaaS istanziati per tipologia, per gruppo e per provider
 - Integrazione con Federation Registry
 - Flavor e sistemi operativi mostrati sono recuperati direttamente dai siti
 - Aggiunta una nuova variante del servizio PaaS "Kubernetes cluster" con un InterLink Virtual Node
 - Aggiunto un nuovo servizio centralizzato basato su Healthchecks.io

PaaS – Maintenance (2)

- Aggiornati Tosca Templates
 - Passaggio a Debian 12 e integrazione con Federation Registry
 - Revisione e aggiornamento dei servizi basati su Kubernetes
 - Aggiornamento versione owncloud e nexcloud del servizio sync&share per risolvere vulnerabilità
- Ripristinato e aggiornato il servizio centralizzato Jupyter Notebook as a Service (NaaS)
 - Nuove docker images basate su immagini ufficiali
- Creata nuova dashboard per ICSC + Vault dedicato
- Testata con successo la sottomissione tramite IM di un deployment su cluster Kubernetes integrato con IAM
- Aggiornamento ricette Ansible alla versione più recente
 - Introduzione di Ansible Vault: password e certificati criptati
 - Aggiunta ricetta per Federation Registry + Feeder
- Revisione procedura per la messa in produzione
 - Abilitato il backup dei DB sulla PaaS di Pre-produzione
 - Aggiornata ricetta Ansible per fare il deployment dei servizi della PaaS in contemporaneo sia su BARI che su CNAF

PaaS - Nuovi sviluppi

- Federation-Registry
 - Migliorati test automatici
 - Supporto configurazione reti di CloudVeneto
- Piattaforma streaming dati
 - Passaggio a un sistema message-driven (Kafka)
 - Le componenti che recuperano e producono le informazioni sono state dockerizzate e in pre-produzione.
- AI-Ranker
 - Implementati i servizi di training e inferenza che recuperano le metriche da Kafka
 - Modificato l'Orchestratore attuale per interagire con questa componente
- Federation Manager
 - Documento per le REST API esposte dal servizio
 - Scrittura della componente che esegue i test sui provider e li inoltra a Kafka
 - Design dell'interfaccia grafica
- Orchestratore Python
 - Documento riassuntivo dell'architettura e dei messaggi scambiati
 - Progettata architettura, REST API, interazione tra componenti e studiata integrazione con Rucio

Altri sviluppi/attività

- S3WebUI messa in produzione ed in continuo miglioramento
- Riorganizzazione e aggiornamento documentazione interna
- Realizzato inventario sulle VM che ospitano servizi PaaS di produzione e pre-produzione
- Lavoro di revisione di tutti i ticket aperti WP5
- Organizzati seminari interni su utilizzo Jenkins, sulle tecnologie OpenID-Connect/OAuth, su Federation-Registry API e su Rucio
- Partecipazione con contributi a conferenze quali ad esempio CHEP2024 e ISGC2025

Interlink offloading

Extend the container orchestration layer (Kubernetes) to support «offloading» to external, possibly HPC, systems

Based on Virtual Kublet (virtual k8s node)



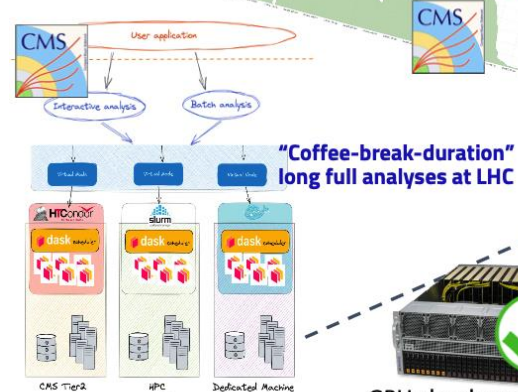
HOW IT'S GOING? Not too bad!

GPU accelerated HEP simulations and triggers

Grid-friendly workflows

10k cores = 1k pods

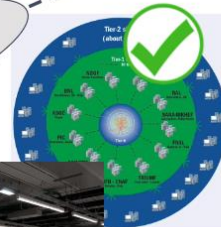
ML training and inference of particle tagging



interLink

We started here in the middle, remember?

GPU cloud machine

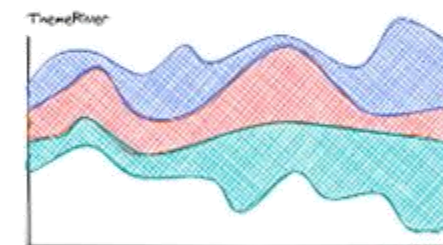


My analysis pipeline



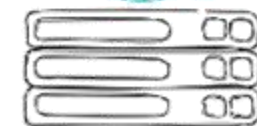
My k8s cluster

Virtual K8s node



Analysis results

Workflow task



HPC GPU Node

Architettura e il Middleware di federazione

L'obiettivo del PoC è integrare l'infrastruttura di calcolo dell'INFN e di CINECA usando il middleware di federazione, dimostrando la fattibilità del DataLake Nazionale previsto dall'architettura di ICSC (e TeRABIT).

Basato su 4 pillars

- Federazione di identità attraverso il servizio **IAM**
- Federazione di cloud attraverso il servizio **PaaS Orchestrator**
- Federazione di risorse eterogenee (cloud-HPC) attraverso il servizio **interLink**
- Federazione di storage per gestione dati attraverso **RUCIO (+FTS)**

Tra i servizi ancillari ha un ruolo di rilievo il sistema per la distribuzione del software basato su CVMFS

Stato dell'implementazione

Tutti i servizi middleware per la federazione sono stati deployati

- L'integrazione con le risorse INFN è completata mentre con le risorse CINECA è ancora work-in-progress.

Federazione di identità seppur tecnicamente fattibile per CINECA ha un forte impatto sulla gestione dell'utenza e quindi sul processo di certificazione.

Federazione cloud è stata dimostrata ma è limitata (lato CINECA) dal problema relativo alla federazione di identità

Federazione risorse eterogenee via interLink (INFN-Cloud & Leonardo) è completa

Federazione di storage è in fase di realizzazione.

L'infrastruttura CVMFS può considerarsi di produzione sia per INFN che CINECA

Fase di Validazione

Per la validazione del PoC sono stati selezionati alcuni use case (quelli più visibili dalla nostra prospettiva)

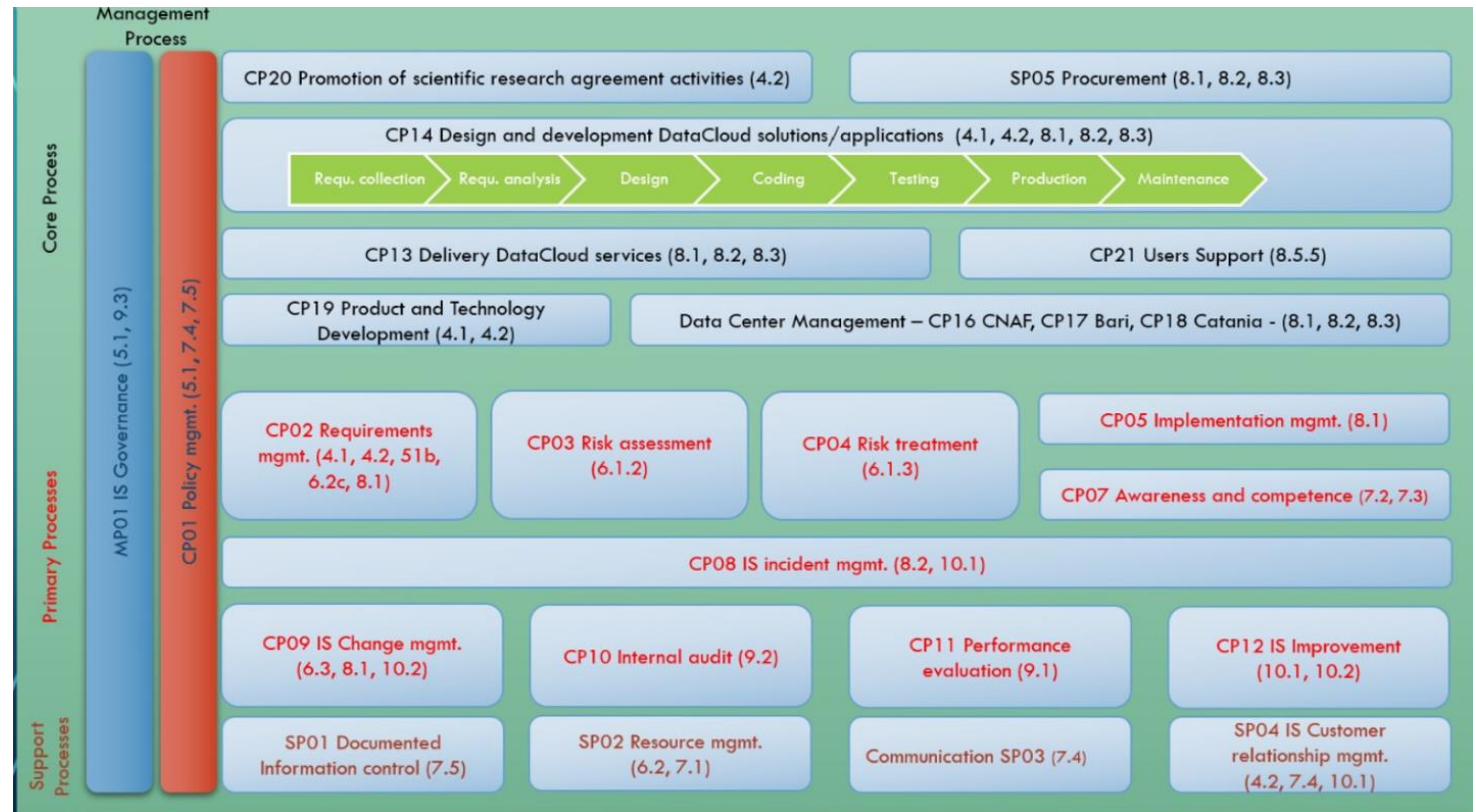
- **Spoke2** (L.Anderlini)
 - Advanced Machine Learning. Flash Simulation and bleeding edge applications
- **Spoke3** (D.Ciangottini)
 - Scientific Hub for interactive processing over GPU equipped node
- **TeRABIT** (G.Donvito)
 - Urgent Computing

Appena la federazione storage sarà completata vari use cases (l.e. Spoke8) saranno utilizzati da dimostratori

EPIC

Riconfermato lo scope: **Co-design**, development and maintenance of DataCloud software solutions for the research sector. Provision of DataCloud IaaS, SaaS and PaaS services with **Community Cloud** deployment model. Services provided by CNAF, Bari and Catania sites.

L'audit esterno per la certificazione multisito CNAF-BA-CT + alcuni processi in AC e' stato schedulato a dicembre '25



Conclusioni e considerazioni per il futuro

- In quest'ultimo anno si sono visti diversi risultati tangibili del lavoro di collaborazione:
 - Si è allungata la lista di siti federati
 - Si è allungata la lista dei SaaS forniti
 - Lo sviluppo della PaaS sta continuando in modo stabile e con una roadmap chiara, nonostante diversi scossoni
- La sinergia fra i progetti PNRR e quelli EU e le attività core dell'ente si dimostra una strada percorribile in modo stabile
- Le risorse (IT) finanziate dai progetti PNRR stanno andando in produzione (anche nella regione EPIC)
- stiamo supportando oltre alle comunità interne anche quelle del PNRR
- stiamo lavorando (con fatica) con CINECA per creare la federazione italiana
- stiamo cercando di rendere sostenibili le componenti che sviluppiamo e che gestiamo sull'infrastruttura.

➔ Per continuare con queste attività abbiamo bisogno che tutte le persone che hanno contribuito possano continuare a lavorare nel nostro ente con la giusta motivazione