



Aggiornamenti DataCloud

Claudio Grandi

Agenda



- Operazioni e gestione dell'infrastruttura
- Sviluppo software – Criticità per IAM e Orchestratore
- Mandato del GdL su Kubernetes
- Accesso alle risorse INFN per utenti ICSC/TeRABIT
- PoC della federazione cloud ICSC/TeRABIT
- Certificazione EPIC multisito
- Matrice funzionalità vs. backend
- Aggiornamento su attività del PO

Operazioni e gestione dell'infrastruttura



Consolidamento delle **cloud federate** dell'INFN

Attivazione dei progetti che operano su **rete privata**

Impatto minore sulla security, non richiesta la nomina a sysadmin

Organizzata la migrazione del backbone S3 da MINIO a **CEPH RGW**

A giugno – richiederà down di circa mezza giornata

Miglioramenti per monitoring e accounting

Risorse pledged per CYGNO: insostenibile continuare sul backbone

Non è mai stata questa la funzione del backbone

In attesa della migrazione del CNAF al Tecnopolo, si allocano le risorse a Bari

Parte delle risorse CPU su Grid → test di interoperabilità Grid-Cloud

Operazioni e gestione dell'infrastruttura



Iniziate le attività per il supporto agli utenti con risorse allocate dal RAC di ICSC (vedi oltre)

In corso il [survey delle risorse](#) di Tier-1 e Tier-2 (presentazione a Palau). Dopo la *fase-0*, da sviluppare meccanismi per l'aggiornamento automatico

In collegamento con le attività su monitoring e accounting già menzionate

Nuovi utenti INFN del sistema di trasferimento dati (CYGNO, DAMPE, DarkSide) – attenzione ad aspetti di AAI/security

Automazione dell'offloading – test su use case AI_INFNO

Sviluppo software



Miglioramento degli strumenti a supporto dello sviluppo

Continuous integration (Jenkins), consolidamento repository (NEXUS)

Nuova organizzazione infn-datacloud su github

Criticità per IAM e Orchestratore



IAM: Priorità alla transizione da VOMS a IAM sulle infrastrutture in produzione (WLCG e altri). Abbandono di CentOS7.

Fix di vulnerabilità. Eliminazione di dipendenze obsolete (Angular.js e MitreID). MFA

Orchestratore: Molte dipendenze obsolete

Proposto porting a Python per ampliare la base di sviluppatori

Supporto per Kubernetes

Maggiore integrazione con l'IM e collaborazione con UPV?

Serve definire la nuova architettura e il piano delle attività

Scarso effort, sia sul core sia sullo sviluppo dei servizi PaaS

Ci sono concrete possibilità di avere più sviluppatori ma da [chiarire l'impegno dei senior](#)



Mandato del GdL su Kubernetes



Presentato alla riunione del LNGS

Da approvare formalmente

Nelle riunioni del GdL sono state evidenziate attività su k8s anche in altri ambiti. Necessario far convergere le attività in un unico gruppo.

Accesso alle risorse INFN per utenti ICSC/TeRABIT



Primo passo: **registrazione degli utenti**. Opzioni:

- **INFN-AAI**

Deve essere configurata una «struttura» (ad es. DataCloud) per gestire le registrazioni. Per farlo sarebbe necessaria la definizione della catena di responsabilità (un Direttore di riferimento)

[aiuterebbe anche con una non conformità rilevata dagli auditor per DataCloud]

Gli utenti devono avere privilegi limitati, non come gli altri utenti INFN

Gli sviluppatori di AAI potrebbero produrre quanto serve per fine maggio. In gran parte è nel piano di lavoro

- **Servizio installato appositamente**

Non chiaro se ci sono le competenze e di conseguenza le tempistiche.

Potrebbe diventare un servizio di produzione per gli utenti ICSC senza identità federata

Accesso alle risorse INFN per utenti ICSC/TeRABIT



In entrambi i casi si devono rispettare le regole INFN (riconoscimento, formazione)

Per il riconoscimento possibilità di usare SPID per evitare la burocrazia

In preparazione un modulo per la certificazione da parte del datore di lavoro della formazione (utente/sysadmin)

Lo **user support di DataCloud** (WP2) avrà comunque la responsabilità di gestire gli utenti

Accesso Grid quando possibile, cloud quando richiesto, locale se indispensabile. Molte richieste per le HPC bubble, che saranno soddisfatte solo quando arriveranno

PoC della federazione Cloud ICSC/TeRABIT



Attività comune ICSC/TeRABIT. Deadline 30/6/2024

Coordinamento di WP6 e CTO

Riunione con CINECA e GARR 23/4

- **Federazione** di Galileo 100 con siti INFN. IAM, dashboard e orchestratore PaaS
- **Offloading** per accesso a Leonardo
- **Gestione dati** con Rucio e FTS

La risposta di CINECA è stata positiva ma ci sono alcuni punti da chiarire, in particolare per l'accesso allo storage su cui usano una soluzione S3 proprietaria DDN. Da programmare una riunione con i loro sysadmin

Noi comunque partiamo in parallelo per tutto quello di nostra competenza

Setup VO, servizi PaaS, offloading

Piano di lavoro INFN:

<https://docs.google.com/document/d/1O12SNX5loa31Yt1izCokl9IIQrqr9QDs0Az94T47gl/edit>

Certificazione EPIC multisito



Le attività procedono più lentamente del previsto. Chiesto dallo steering di rilasciare la versione 0 di tutti i processi entro il 15/5

Ridefiniti nell'ultimo steering i responsabili di alcuni processi

Il contratto non è ancora stato firmato anche se il consulente lavora dal luglio scorso!

Riunione in presenza 30/5/2024 a Frascati

Essenziale reperire una figura con profilo legale esperta nella gestione dei dati!

Perimetro di certificazione



Bisogna definire il **perimetro della certificazione** in tempi rapidi (maggio) per consentire la corretta pianificazione dei prossimi passi

Siti interi o solo porzioni (in particolare per il CNAF)?

Solo IaaS o anche PaaS e SaaS?

Tutto il processo software? Possiamo considerare IAM e orchestratore componenti esterne a EPIC?

Il processo di **sviluppo sicuro del software** si sovrappone alle difficoltà di cui abbiamo già parlato ma sarebbe strategico poter dimostrare di avere un approccio coerente per software e gestione

Documento di domande/risposte:

<https://istnazfisnucl.sharepoint.com/:w:/s/INFNCloud9/EZAF0qtAVJxPI-fZBLWsusYBmuBvX1z9vLtYmA8tZuAUZQ?e=LPo3Ug>

Matrice funzionalità vs. backend



Donvito

Aggiornamento su attività del PO



Spiga