

ML_INFNO – Cloud@CNAF

Diego Michelotto (diego.michelotto@cnafe.infn.it)

Cristina Duma (cristina.aiftimiei@cnafe.infn.it)

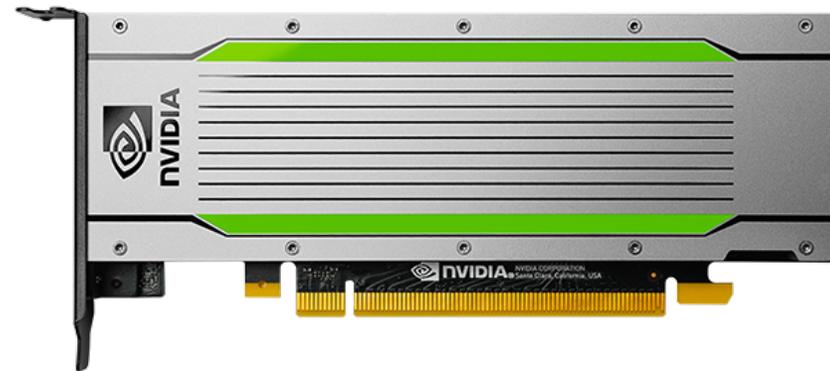
8/05/2020

Quest'opera è distribuita con Licenza Creative Commons
Attribuzione - Non commerciale - Condividi allo stesso modo 3.0 Italia.



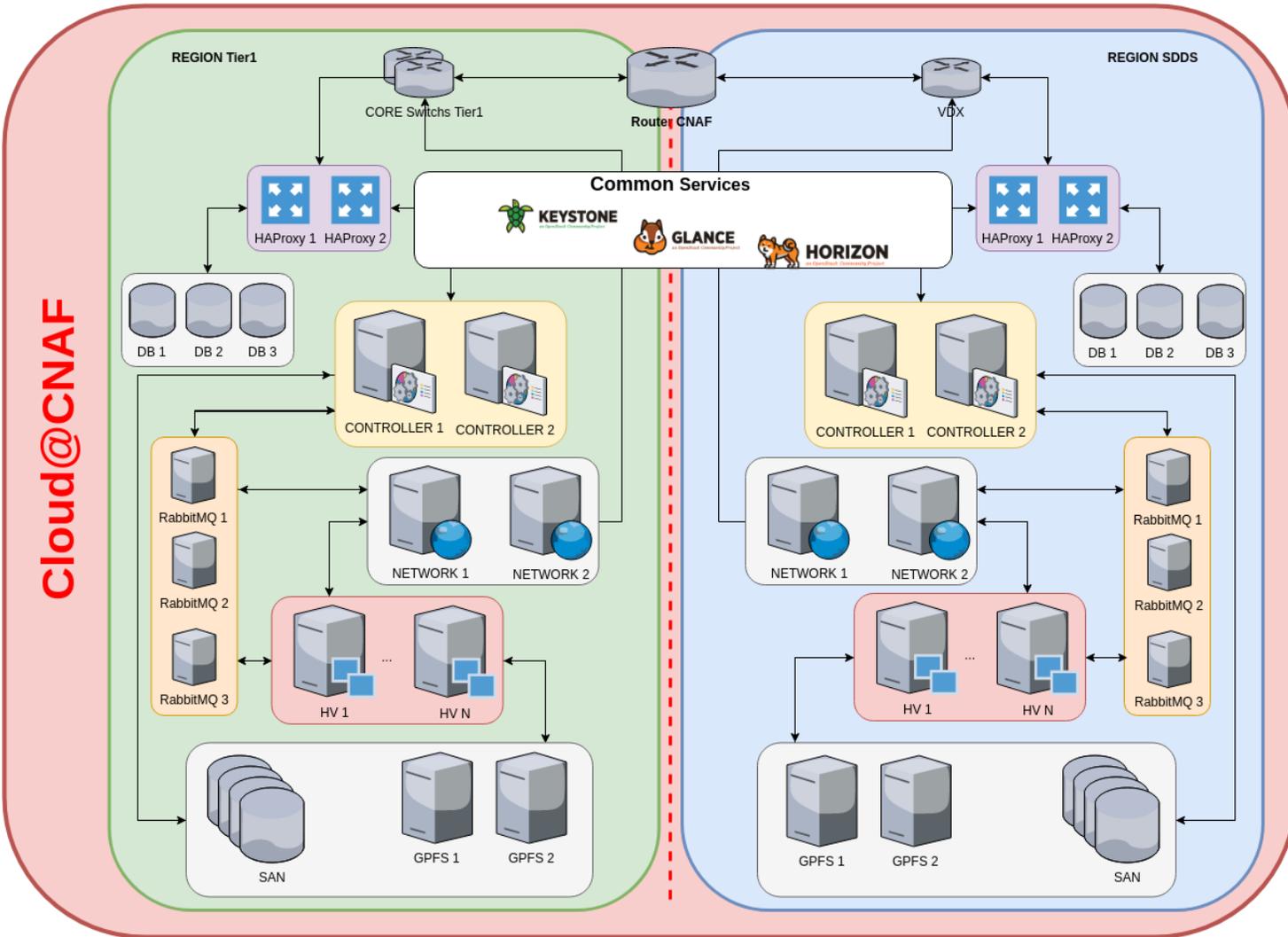
Indice

- Cloud@CNAF
- Hardware ML_INFN
- INFN Cloud
- Scenari d'uso
- Conclusioni



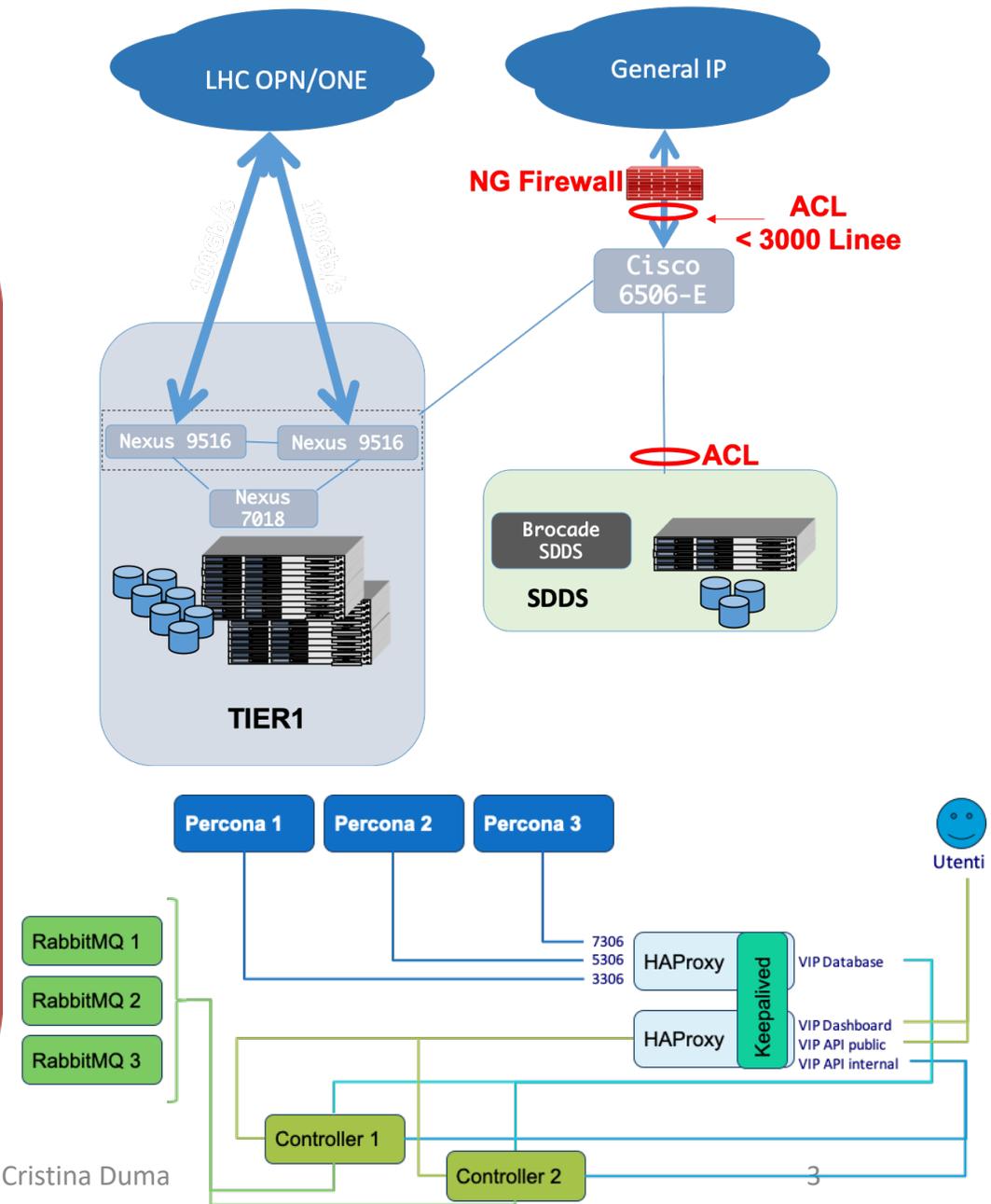
Cloud@CNAF

Cloud@CNAF



08/05/20

ML_INFN - Cloud@CNAF - Diego Michelotto, Cristina Duma



Hardware ML_INFN

Nuova macchina inserita in **Cloud@CNAF** nella regione **SDDS**

- Previsto l'utilizzo dei due dischi **NVMe** in **raid 0** da usare come **backend per Nova**
- Vista l'espandibilità di questa macchina se si aggiungono dischi di questo tipo si potrà migrare ad un altro livello di raid.
- Connettività **10Gb/s**
 - Possibilità di posizionamento vicino a switch Infiniband.
- Serve licenza «**NVIDIA Virtual GPU**» (prevista per GPU Tesla: **T4/V100/P100** ; "NVIDIA Education Pricing Program, for qualified Educational Institutions")
 - Virtualizzando la GPU, è possibile creare più VM per ogni GPU (Use case: scuola utenti, tante VM)
 - Altrimenti potrà essere creata al massimo 1 VM per ogni GPU
- Andranno creati gli opportuni **metadati, host_aggregate e flavour** per poter istanziare correttamente le VM che richiedono GPU sull'hardware ML_INFN
 - Possibilità di dedicare questi flavour a uno o più progetti, federati e non.

INFN Cloud

Infrastruttura cloud distribuita

- Backbone – basato su OpenStack
 - 2 regioni: Bari (!= RECAS Bari) e CNAF (!= Cloud@CNAF)
 - Per ogni regione: 80 CPU, 500GB RAM, ~ 500 TB RAW Ceph (Nova, Cinder, Glance, Swift)
 - Possibilità di volumi mirrored per DR
 - Rete GARR dedicata con L3VPN per interconnettere tra loro le reti private
- Cloud Federate
 - RECAS Bari
 - Cloud@CNAF è **già federata** con Cloud@INFN tramite <https://iam.cloud.infn.it> da fare:
 - **Verificare corretta integrazione** con PaaS Cloud INFN
 - Definizione delle risorse con GPU
 - Da **rivedere mapping utenti tra Keystone e IAM** (Da discutere)
 - Possibilità di creare Progetto/Tenant ad hoc per ML_INFNO con relativo gruppo dedicato su IAM

Scenari d'uso

- La VM istanziata sulla macchina ML_INFN avrà a disposizione
 - 1, più di 1, o una porzione di **GPU** (salvo disponibilità licenza)
 - Il disco della VM sarà una **porzione dei dischi NVMe**, quindi direttamente utilizzabile per lo spostamento in locale dei data set
 - L'hardware a disposizione ci permetterà di avere vm con **dischi grandi**.
 - Per esempio con 6 VM (1 per ciascuna GPU) si potranno avere a disposizione un disco da 4TB.
 - Avrà **connettività 10Gb/s** per il trasferimento dei data set da
 - Storage di esperimento dal T1 (Da studiare modalità di accesso. Es. NFS, WebDAV, altro)
 - Nuovo storage in arrivo al CNAF per progetti (Cinder, Swift, altro)
 - Cache locale (Root, altro)
 - Siti remoti (wget, curl, altro)

Scenari d'uso

1. Utente istanzia 1 VM con GPU su Cloud@CNAF facendo parte del progetto ML_INFN
2. Utente istanzia 1 VM con GPU da PaaS Cloud INFN selezionando le risorse di Cloud@CNAF
3. Utente istanzia 1 Jupyter Notebook con TensorFlow da PaaS Cloud INFN selezionando le risorse Cloud@CNAF
4. Dedicare parte delle risorse della macchina da per utilizzi alternativi, ad esempio per test di uso come nodo K8s/Mesos per istanziare container che utilizzino la GPU

Backup slides

Cloud@CNAF

Obbiettivi:

- Unica infrastruttura condivisa tra **Tier1, SDDS e Tier3**
- Riutilizzare **non replicare** il lavoro già fatto.
- Migliorare l'alta affidabilità
- Dare **risorse cloud a utenti/comunità del Tier1**
 - Region **Tier1**
- Dare **risorse a gruppi di utenti, progetti che non hanno pledge al Tier1**
 - Region **SDDS**

Servizi OpenStack

- Centrali: **Keystone, Glance, Horizon**
- Regioni: **Nova, Cinder, Neutron, Heat, Octavia**

Risorse

- **SDDS**: 792 Core, 2.9 TB RAM, 16 TB storage distribuito GPFS (Nova, Cinder), 16TB storage per local I/O su HDD (Nova), 1TB storage per local I/O su SSD (Nova), 510 FIP. **In arrivo nuovo storage per Ceph**
- **Tier1**: 496 Core, 1.8TB RAM, 500 TB storage distribuito GPFS (Nova, Cinder, Glance), 1022 FIP