

Cluster GPU basato su docker

Gioacchino Vino (INFN Bari)

ARPA meeting, Bari, 12Apr2021

INDICE

- **Motivazioni**
- **Dotazione Hardware**
- **Caratteristiche**
- **Tecnologie**
- **Come funziona**
- **Sviluppi futuri**

MOTIVAZIONI

- L'incremento della quantità e della eterogeneità dei dati, sta portando a una maggiore richiesta di hardware performante che riesca ad abbattere i sempre crescenti tempi di esecuzioni degli algoritmi.
- La diffusione di algoritmi di Intelligenza Artificiale (e quindi di Deep Learning) ha dimostrato di poter raggiungere delle prestazioni di alto livello a discapito di un fasi di addestramento dello stesso molto lunghi.
- La parallelizzazione e' la soluzione che si preferisce adottare in centri di calcolo, su vari livelli:
 - Esecuzione di applicazioni a singolo core su un gruppo di macchine
 - Esecuzione di applicazioni multi core su un gruppo di macchine
 - **Esecuzione di applicazioni su GPU su un gruppo di macchine**

DOTAZIONE HARDWARE

- In tutto si posseggono 5 macchine fisiche, ognuna con le seguenti caratteristiche:
 - CPU: 96x Intel(R) Xeon(R) Gold 6252 CPU @ 2.10GHz
 - GPU: 4x Nvidia V100 32GB
 - RAM: 754 GB
 - DISCO: 4TB SSD e 500GB HDD

- Il cluster complessivamente possiede:
 - CPU: 480x Intel(R) Xeon(R) Gold 6252 CPU @ 2.10GHz
 - GPU: 32x Nvidia V100 32GB
 - RAM: 3.68TB
 - DISCO: 20TB SSD e 2.5TB HDD

CARATTERISTICHE

- Condivisione delle risorse di calcolo tra più utenti contemporaneamente
- Concessione delle risorse in maniera atomica (CPU=4, MEM=500MB, DISCO=20GB, GPU=2)
- Isolamento tra le risorse concesse tra gli utenti
- Gestione dei guasti e migrazione dell'applicazione su una macchina funzionante
- Utilizzo di un file system distribuito
- Schedulazione di applicazioni batch
- Fornitura di applicazioni interattive per la scrittura di codice
- Applicazione di politiche di limitazione delle risorse che ogni utente può richiedere (quota)
- Sicurezza

TECNOLOGIE

Apache Mesos



Apache
MESOS[™]

- unifica tutte le risorse presenti in un cluster di macchine creando una singola entità da cui e' in grado di discretizzare le risorse da concedere alle singole applicazioni
- Supporta containers
- Possibilità di gestire un grande numero di nodi
- Caratteristiche offerte:
 - Condivisione delle risorse di calcolo tra più utenti contemporaneamente
 - Concessione delle risorse in maniera atomica (CPU=4, MEM=500MB, DISCO=20GB, GPU=2)

TECNOLOGIE

Marathon



- Interfaccia grafica per la sottomissione delle applicazioni in Apache Mesos
- Alta affidabilità delle applicazioni eseguite
- Verifica periodica dello stato di salute di una applicazione
- Supporta container
- Load Balancing e Service Discovery
- Sicurezza: permette l'accesso alle applicazioni interattive tramite HTTPS
- Caratteristiche offerte:
 - Gestione dei guasti e migrazione dell'applicazione su una macchina funzionante

TECNOLOGIE

Chronos

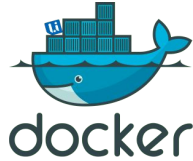


- Scheduler ad alta affidabilità di applicazioni da eseguire su un cluster Mesos
- E' l'equivalente di 'cron' per Apache Mesos
- Supporta container

- Caratteristiche offerte:
 - Schedulazione di applicazioni batch

TECNOLOGIE

Docker



- Un container e' una unita' software che permette di impacchettare codice, librerie, dipendenze e configurazione in modo che l'applicazione possa essere eseguita velocemente e in modo affidabile su macchine con ambienti o sistemi operativi differenti.
- Il container isola l'applicazione dalla configurazione della macchina su cui e' eseguito.
- L'immagine di un container e' leggera, standalone e contiene tutto il necessario per essere eseguita.
- Immagini ufficiali disponibili (Nvidia, TensorFlow, SQL, ...)
- Caratteristiche offerte:
 - Isolamento tra le risorse concesse tra gli utenti

TECNOLOGIE

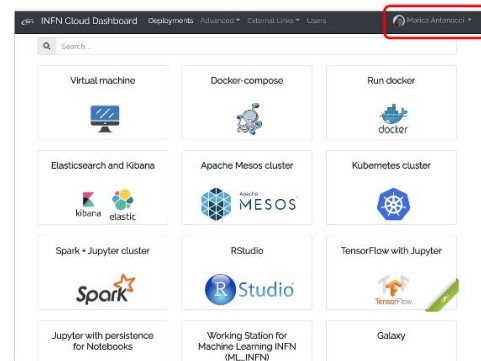
Dashboard/PaaS

- Offre una semplice e intuitiva interfaccia grafica tramite la quale l'utente può effettuare le operazioni di richiesta di avvio/rimozione di container e raccogliere tutte le informazioni per accedervi
- Interagisce con Marathon e Chronos
- Autenticazione via LDAP o IAM
- Caratteristiche offerte:
 - Applicazione di politiche di limitazione delle risorse che ogni utente può richiedere (quota)
 - Interfaccia grafica con cui l'utente può interagire facilmente

COME FUZIONA

Richiesta di un nuovo container

- Accesso alla dashboard con le credenziali ReCaS (LDAP ReCaS)
- Selezione del container di cui si ha bisogno (gli utenti hanno la possibilità di poter utilizzare un proprio container)

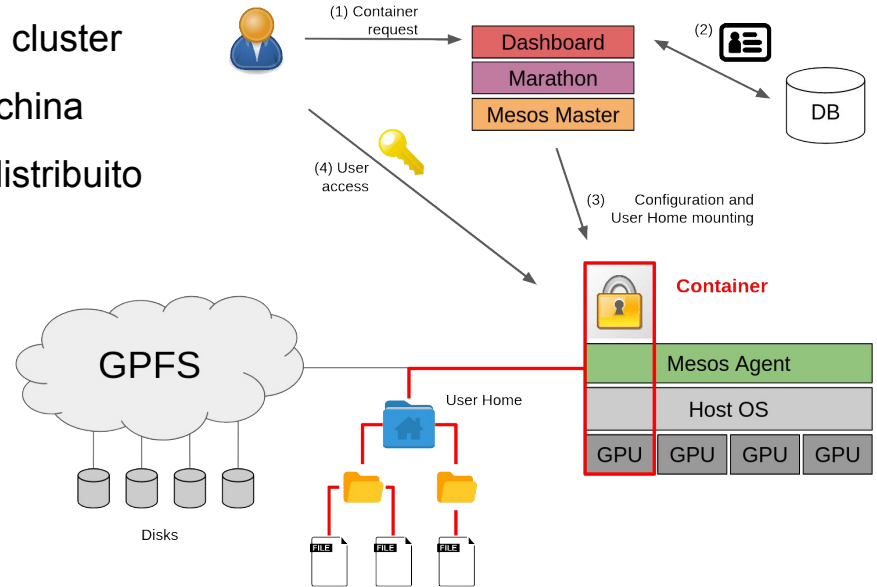
A screenshot of the 'Virtual machine' configuration form. The form has a title 'Virtual machine' and a description: 'Launch a compute node getting the IP and SSH credentials to access via ssh'. Below the description is a 'Deployment description' field. The 'Configuration' tab is selected, and the 'Advanced' sub-tab is active. The 'services_ports' section has an 'Add rule' button. The 'Ports to open on the host' section is empty. The 'flavor' dropdown is set to '--Select--', with a note: 'Number of vCPUs and memory size of the Virtual Machine'. The 'operating_system' dropdown is also set to '--Select--', with a note: 'Operating System for the Virtual Machine'. At the bottom are 'Submit' and 'Cancel' buttons.

- Inserimento di informazioni riguardo il container selezionato:
 - Risorse richieste (CPU, GPU, RAM, Disco)
 - Eventuali librerie da installare nel container
 - Altre informazioni (password, chiave SSH,...)

COME FUZIONA

Avvio del container richiesto

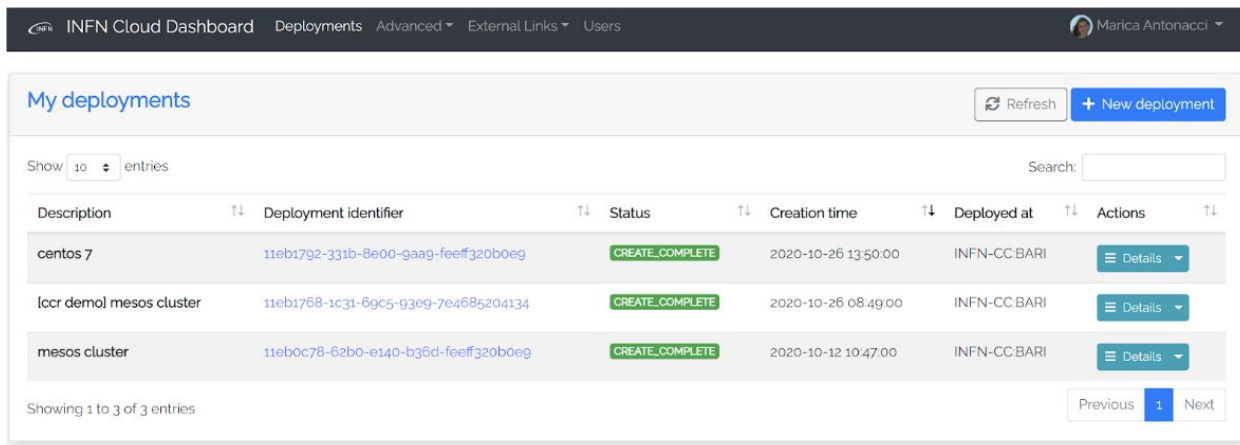
- Verifica se l'utente abbia superato la propria quota
- Verifica se le risorse richieste sono disponibili nel cluster
- Avvio e configurazione del container su una macchina
- Montaggio della propria \$HOME del file system distribuito di ReCaS all'interno del container
- Consegna all'utente di tutte le informazioni necessarie per poter accedere in sicurezza al container



COME FUZIONA

Lista dei propri container

- Informazioni presenti:
 - Nome, ID, Stato, data di creazione
 - La colonna azioni permette di cancellare il proprio container o di recuperare informazioni più dettagliate



The screenshot shows the 'My deployments' section of the INFN Cloud Dashboard. The dashboard header includes 'INFN Cloud Dashboard', navigation links for 'Deployments', 'Advanced', 'External Links', and 'Users', and a user profile for 'Marica Antonacci'. The 'My deployments' section has a 'Refresh' button and a '+ New deployment' button. Below this is a search bar and a 'Show 10 entries' dropdown. The main content is a table with columns for Description, Deployment identifier, Status, Creation time, Deployed at, and Actions. Three deployments are listed, all with a status of 'CREATE_COMPLETE'. The 'Actions' column for each entry contains a 'Details' button.

Description	Deployment identifier	Status	Creation time	Deployed at	Actions
centos 7	11eb1792-331b-8e00-9aag-feeff320b0e9	CREATE_COMPLETE	2020-10-26 13:50:00	INFN-CC.BARI	Details
lccr demo mesos cluster	11eb1768-1c31-69c5-93e9-7e4685204134	CREATE_COMPLETE	2020-10-26 08:49:00	INFN-CC.BARI	Details
mesos cluster	11eb0c78-62b0-e140-b36d-feeff320b0e9	CREATE_COMPLETE	2020-10-12 10:47:00	INFN-CC.BARI	Details

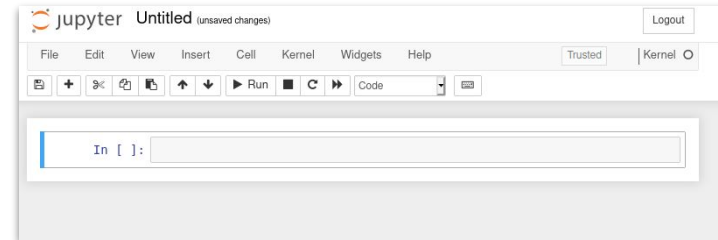
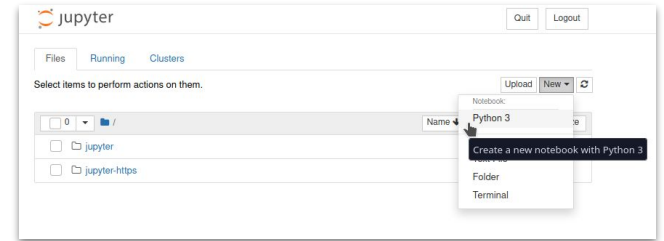
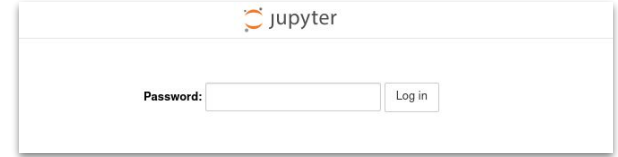
Showing 1 to 3 of 3 entries

Previous 1 Next

COME FUZIONA

Caso d'uso: richiesta di un container con Jupyter Notebook

- Alla conclusione della fase di istanziazione del container sarà restituito un URL tramite il quale accedere al proprio Jupyter Notebook tramite un browser
- L'inserimento della propria password permette di visualizzare il contenuto della propria HOME e creare script Python3
- Cliccando sul file creato, un IDE (Integrated Development Environment) e' reso disponibile all'utente tramite il quale può scrivere codice che utilizzi le risorse assegnate.



SVILUPPI FUTURI

- Aggiunta di nuovi nodi al cluster con GPU più performanti (Nvidia A100 40GB)
- Aumentare l'offerta di container da presentare all'utente
- A breve sarà disponibile un'interfaccia analoga a quella presentata per la sottomissione di job batch e/o periodici che richiedono l'utilizzo di GPU
- Costruzione automatica di container e container registry interno a ReCaS