

RMLab : Gestione Agile di un data center distribuito

Il progetto RMLab nasce nel 2015 grazie alla collaborazione delle sezioni INFN di Roma Tor Vergata, Roma Tre e dei Laboratori Nazionali Di Frascati, con il supporto fondamentale del GARR per il networking.

L'idea è nata dall'esigenza di voler creare un'infrastruttura distribuita che fornisse una piattaforma di Cloud Computing di tipo IaaS basata su Openstack, mettendo in comune le competenze e le risorse di diverse sezioni.

Alla base di questa soluzione c'era la necessità di avere un layer di rete che permettesse di far comunicare le tre reti delle sedi INFN in maniera sicura, trasparente ed efficiente.

La soluzione tecnica offerta dal GARR per la realizzazione di tale layer è stata quella di creare tre reti private L3 all'interno di ogni sede e farle comunicare tra di loro tramite protocollo MPLS.

Fondamentale per poter gestire efficacemente un piattaforma di questo tipo era però definire un nuovo modello di gestione di data center distribuiti: modello che permettesse un management agile dell'intero sistema in maniera trasparente, snella e funzionale rispetto ad ogni amministratore delle varie sedi, rimuovendo i confini fisici della collocazione dell'hardware.

Abbiamo adottato un sistema di installazione, provisioning e configurazione delle macchine e dei servizi basato su Foreman/Puppet e ci siamo appoggiati a Git per gestire il versioning del codice che definisce lo stato dell'infrastruttura.

L'implementazione dei vari servizi è stata effettuata su tre diverse tecnologie, non legandoci mai ad una unica soluzione predefinita: abbiamo servizi su bare metal, su macchine virtuali e su Docker container, considerando volta per volta la tecnologia migliore rispetto alle caratteristiche del servizio che andava implementato.

Il punto chiave, indipendentemente dall'approccio utilizzato, è che il servizio possa essere spostato il più rapidamente possibile da una sede all'altra, riducendo al minimo il downtime.

Un servizio di DNS interno, aggiornabile in autonomia dai sistemi stessi tramite nsupdate in fase di avvio, permette di cambiare velocemente la disposizione dei servizi a livello geografico senza che l'infrastruttura ne subisca conseguenze.

Il DNS viene usato anche per gestire l'ha di alcuni servizi tramite script che vanno a modificare i record in round robin, dopo aver controllato quali degli endpoint sono attivi.

Il monitoraggio, l'analisi dei failure e dei log e la verifica delle metriche vengono eseguite dagli amministratori con strumenti open source come zabbix ed ELK.

Attualmente forniamo agli utenti un totale di circa 350 cores, 750GB di RAM, 3,5 TB di ephemeral storage e 12TB di object storage Ceph in replica 3 - una replica per sede.

Il vero valore aggiunto rimane comunque la scalabilità orizzontale di ogni componente dell'infrastruttura.