



Massimo Sgaravatto
INFN Padova
On behalf of the “Cloud dell'area
Padovana” team

Cloud dell'area Padovana Status update

- Obiettivi
 - Uso piu` efficiente delle risorse
 - Gestione sistemistica meno onerosa
 - Evitare la proliferazione di cluster dei diversi gruppi/esperimenti
- → Implementazione di una Cloud tra la Sezione di Padova e i Laboratori Nazionali di Legnaro
 - Per sfruttare le infrastrutture, il manpower, le competenze in entrambi i siti
 - C'e` gia` l'ottima esperienza del Tier-2 in tal senso
- Integrata in una Cloud INFN Nazionale
- Sinergie con altri progetti Cloud (v. Progetto Cloud dell'Universita` di Padova)

- Uso di OpenStack come middleware Cloud
- Unica installazione OpenStack
- Risorse distribuite nei 2 siti
- “Servizi” a Padova, Compute node sia a Padova che a Legnaro
 - Situazione simmetrica wrt il Tier-2, dove i servizi sono a LNL e i Worker Node sono distribuiti nei 2 siti

- Neutron: OpenvSwitch driver con GRE tunneling
- Configurazione “Provider router with private networks”
- Use case: permettere di accedere alle VM cloud dalle LAN di Padova e Legnaro senza necessita` di avere floating IP
- Implementato attraverso
 - Ethernet “estesa” tra Padova e LNL usando le VLAN
 - 2 istanze L3
- Prima istanza L3 per interfaccia esterna “classica” con indirizzo pubblico e NAT
 - Queste VM possono avere floating ip sulla rete pubblica, oppure escono con source nattedo

- Seconda istanza per una seconda interfaccia esterna senza NAT
 - Esce sulla rete usando come gateway una macchina esterna (cld-nat)
 - Gestisce le VM sulla rete 10.62.0.0
 - VM mantengono nelle LAN i loro indirizzi IP 10.62.x.y privati
 - Escono in internet con IP nattato (quello di cld-nat)
 - VM sono raggiungibili direttamente dalle LAN di Padova e Legnaro
 - Route statica sul router di sezione per raggiungere la rete 10.62 attraverso cld-nat
 - L'accesso dalle VM alle LAN di Padova e Legnaro regolato da un firewall via iptables
- V.
http://wiki.infn.it/progetti/cloud-areapd/best_practices/neutron_multiple_external_interfaces

- Configurazione “Off Compute Node Storage – Shared File System” per lo storage delle istanze
- Uso di Cinder per block storage persistente
- Uso di glusterFS come backend
 - In particolare per gestire facilmente l'aggiunta di nuovo storage
- Non abbiamo previsto il deployment di Swift, almeno per il momento
- Possibilita` di accedere a storage “esterno” alla Cloud
 - Tunnel GRE tra storage systems e compute nodes (in modo da non “passare” per il network node)
 - V.
http://wiki.infn.it/progetti/cloud-areapd/best_practices/storage_external

- Per l'implementazione di questa Cloud sono stati fatti i seguenti acquisti:
- Padova (10 Keuro da CCR, resto [~ 60 Keuro] dalla Sezione)
 - 1 Blade enclosure DELL M1000e
 - 4 lame DELL M620 ciascuna con 1 processore E52609, 32 GB RAM
 - 5 lame DELL M620 ciascuna con 2 processori E52670v2, 80 GB RAM
 - 2 switch DELL Force 10 MXL
 - 1 server iSCSI DELL MD3620i, con 23 dischi SAS da 900 GB
- Legnaro (10 Keuro da CCR, resto [~ 17 Keuro] dal Laboratorio)
 - 4 Fujitsu Primergy RX300S7 (2 processori XEON E52650v2, 96 GB RAM)
 - Server storage DELL PowerVault MD3600f (FC) con 12 dischi da 4 TB
 - Ottica 10 Gbps per upgrade link dedicato con Padova

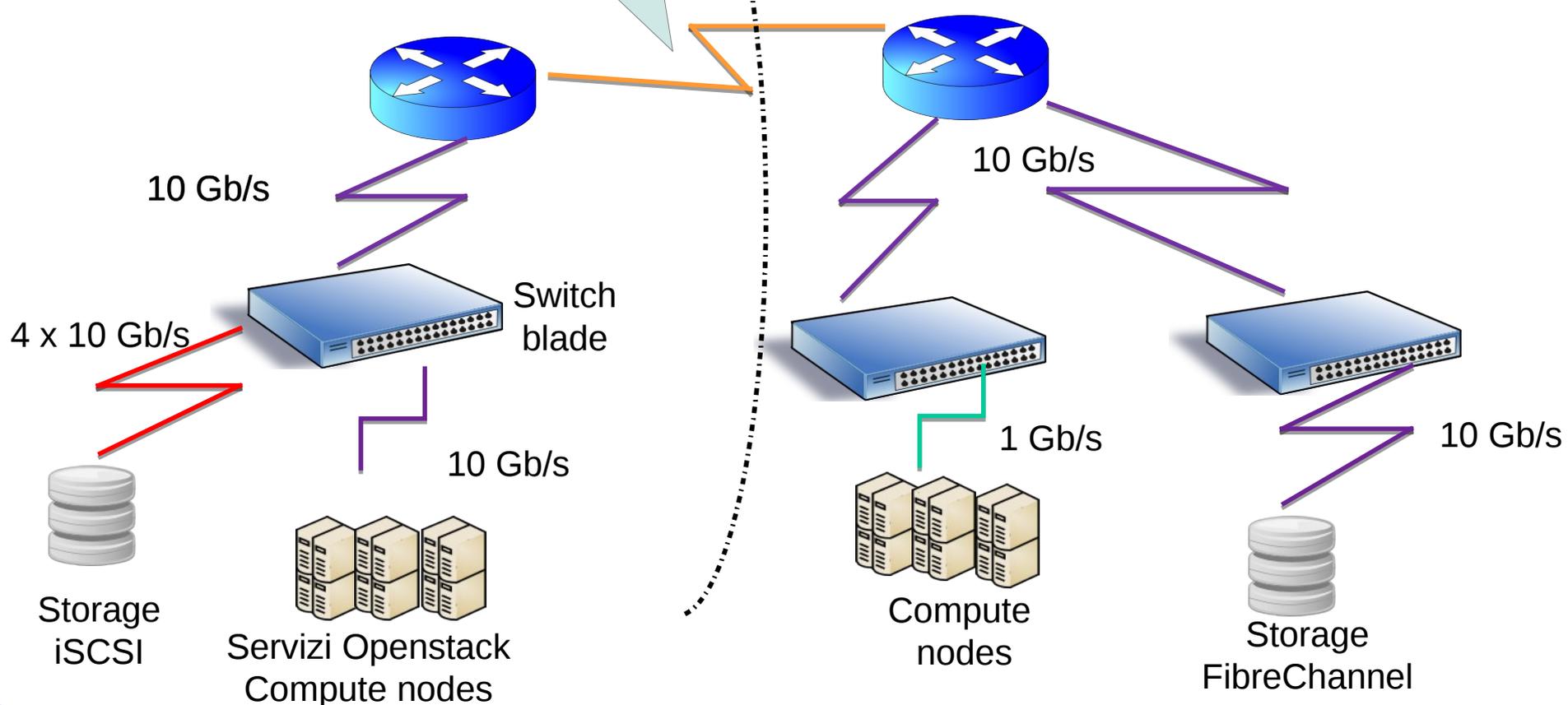
- 4 lame di Padova con processore E52609 usate per i servizi Openstack, in modalita` High Availability (HA)
- 5 lame di Padova rimanenti e i nodi di Legnaro usati come “compute node” (dove vengono istanziate le VM)
- Storage usato per le immagini (Glance), lo storage delle VM, lo storage persistente (Cinder), altro storage Posix (via GlusterFS)
 - Storage non pre-allocato tutto: vedremo in base alle necessita` dove serve
- Oltre a queste vengono usate altre risorse
 - Per servizi per l’installazione e configurazione delle macchine (Foreman, Puppet), servizi di monitoring, database in HA, EGI-Cloud, testing, attivita` di R&D, etc.
 - Stiamo usando al momento vecchio hardware

Layout

Padova

LNL

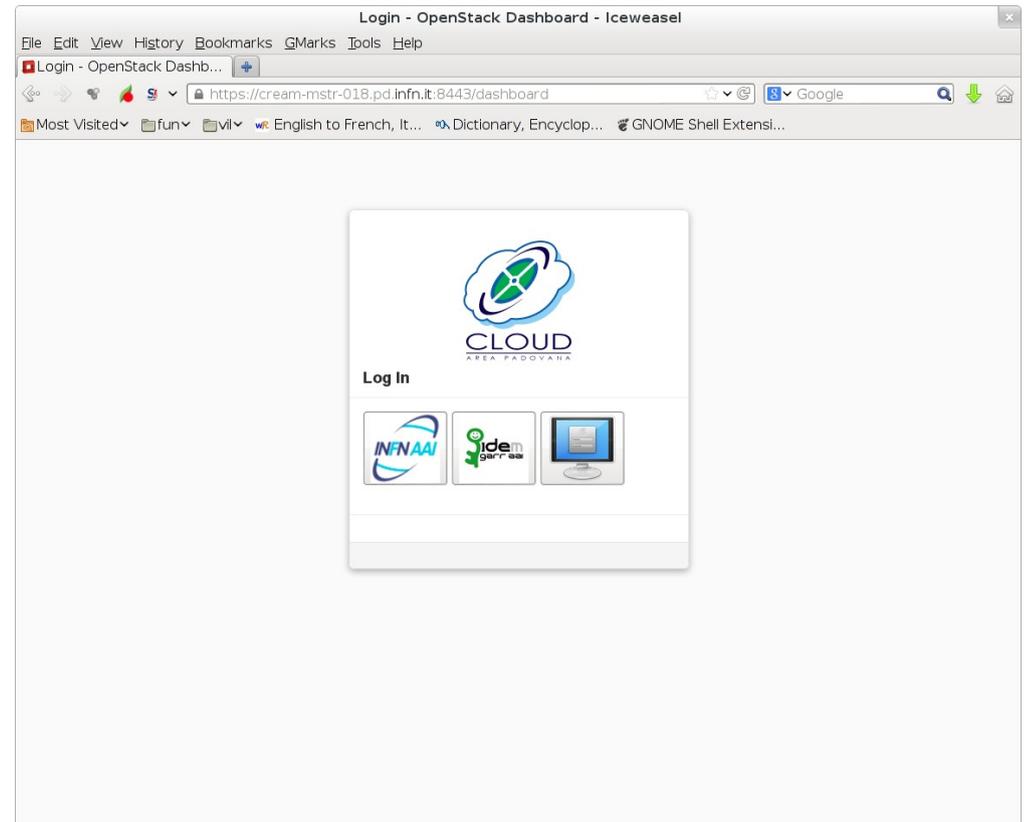
10 Gbps
(dal 20/3/2014)
Qui non passa
Il traffico del Tier2



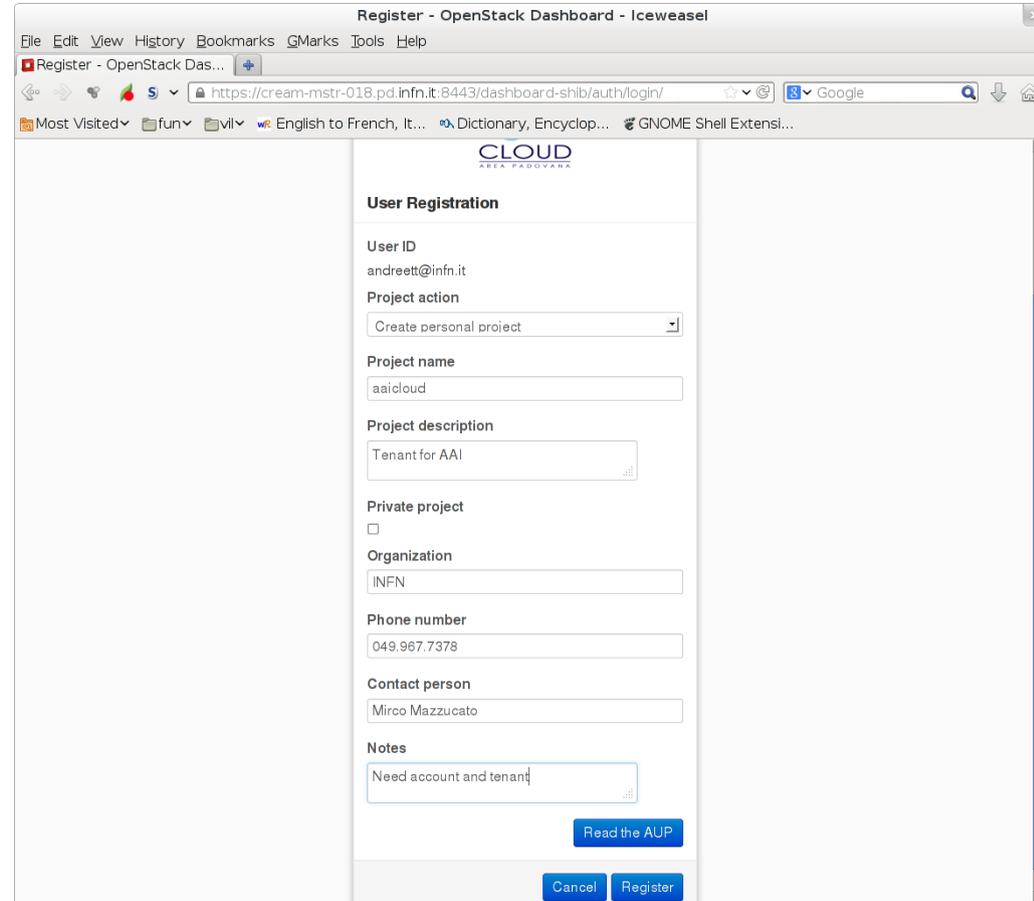
- Implementata una cloud di preproduzione che stiamo aprendo ai primi utenti
 - <https://cloudpp-areapd.pd.infn.it:8443/dashboard>
- Usa la versione Havana di OpenStack
- Risorse limitate
 - 1 controller node a Padova, 3 compute node (2 a Padova e 1 a Legnaro)
 - 112 cores, 282 GB RAM disponibili per le VM degli utenti

- Evoluzione della cloud di preproduzione in una cloud di produzione, includendo le altre risorse e progressivamente nuovi servizi e funzionalita`
- Servizi aggiuntivi
 - Integrazione con INFN-AAI e “funzionalita` avanzate” per la gestione degli utenti e dei progetti
 - Nella cloud di preproduzione si usa la soluzione “standard” di Openstack
 - High Availability (HA)
 - Qualche servizio in HA gia` disponibile nella Cloud di preproduzione
 - Scheduling efficiente wrt quello “standard” di Openstack
 - V. prossima presentazione

- Nel portale di Openstack sono stati integrati i protocolli di autenticazione:
 - SAML2 (Shibboleth)
 - OpenID 2.0
- Possibilita` di autenticarsi alla Cloud usando INFN-AAI
- Possibilita` anche di autenticare tutti gli utenti riconosciuti dalla federazione IDEM
- Possibilita` di autenticarsi con username e password



- Per ogni utente e' richiesta una registrazione
- Nella registrazione puo' richiedere la creazione di un nuovo progetto o l'inserimento in uno o piu' progetti esistenti
- Al momento della registrazione l'utente deve accettare le AUP del servizio (http://wiki.infn.it/progetti/cloud-areapd/documents/acceptable_use_policy)

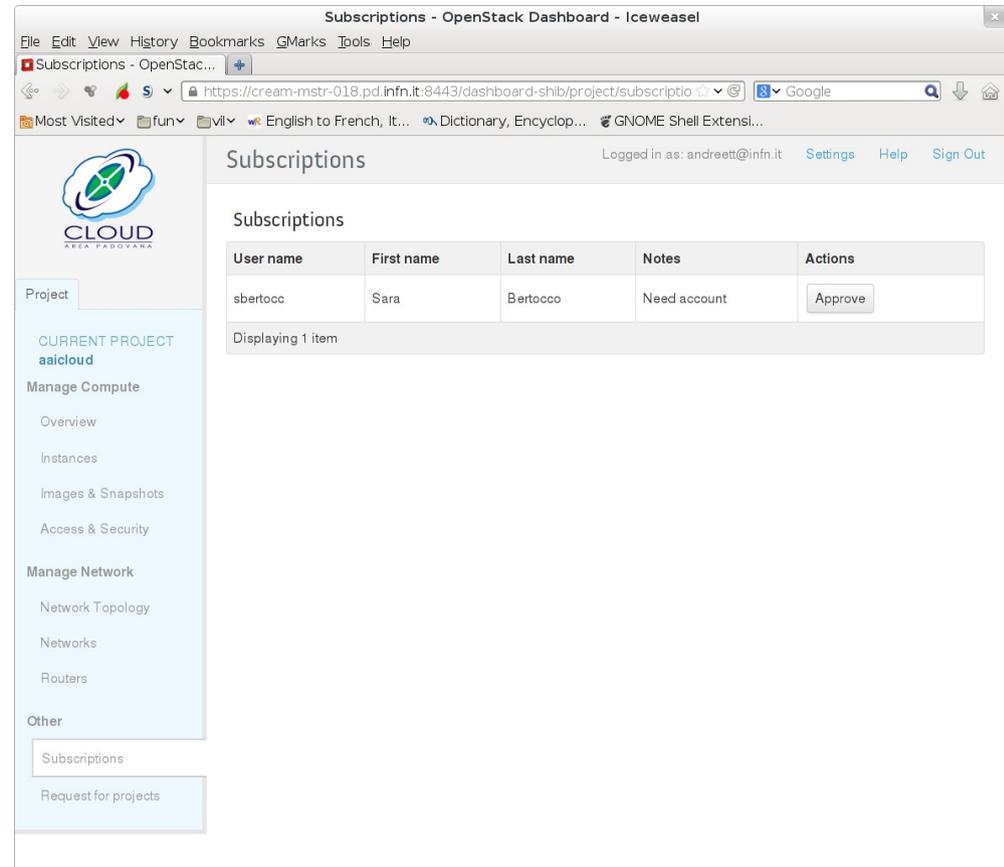


The screenshot shows a web browser window titled "Register - OpenStack Dashboard - Iceweasel". The address bar shows the URL: <https://cream-mstr-018.pd.infn.it:8443/dashboard-d-shib/auth/login/>. The page content is a registration form for "CLOUD AREA PADOVANA". The form fields are as follows:

- User ID:** andreett@infn.it
- Project action:** Create personal project (dropdown menu)
- Project name:** asicloud
- Project description:** Tenant for AAI
- Private project:**
- Organization:** INFN
- Phone number:** 049.967.7378
- Contact person:** Mirco Mazzucato
- Notes:** Need account and tenant

At the bottom of the form, there are three buttons: "Read the AUP", "Cancel", and "Register".

- Ogni utente autenticato deve essere autorizzato ad accedere al servizio
- Oltre all'autorizzazione da parte dell'amministratore della Cloud, un utente che richiede l'ingresso ad un progetto deve essere approvato dall'amministratore di quel progetto
- Modificato il portale in modo da gestire un flusso di approvazione della richiesta di registrazione
 - Notifiche per ogni azione intrapresa nel flusso di approvazione delle registrazioni



Subscriptions - OpenStack Dashboard - Iceweasel

Subscriptions

User name	First name	Last name	Notes	Actions
sbertocc	Sara	Bertocco	Need account	Approve

Displaying 1 item

- Già implementato nella cloud di pre-produzione
 - Cluster Mysql HA usando Percona XtraDB MultiMaster formato da 3 host
 - Cluster HAProxy/Keepalived formato da 3 host per i servizi che richiedono load balancing
- Da mettere in produzione la configurazione dei servizi Openstack in HA
 - 2 Controller Node
 - Servizi: Keystone, Glance, RabbitMQ, Nova API, Nova Schedd, Nova Console, Neutron Server, Cinder API, Cinder Schedd
 - RabbitMQ in cluster mode “nativo” (senza l'uso di HAProxy) con ha_queues
 - 2 Network Node
 - Servizi: Neutron Agents (L2, L3, Metadata, DHCP), Cinder Volume
 - HA in modalita` Active/Active
 - Script per migrazione router fra L3 agents (unico componente non stateless)

- Alcuni tra gli esperimenti/applicazioni che ci aspettiamo proveranno per primi a usare la Cloud:
 - Virtual Analysis Facility (VAF) di ALICE
 - Attivita` gia` partita, fatta nel contesto di un progetto PRIN
 - Altre farm di analisi di esperimento (in particolare CMS, LHCb e Belle-II)
 - Tomografia Muonica

- Sinergia con il working group Cloud della Commissione Calcolo e Reti dell'INFN
 - Partecipazione Padovana in diverse attività, in particolare:
 - Cloud INFN nazionale
 - Istanze keystone e swift a Padova, e cluster openstack a Padova che usa i servizi nazionali
 - V. presentazione di S. Stalio
 - QUACK (Queues in Openstack)
 - Lo scheduler per Openstack efficiente (v. presentazione di E. Frizziero) ne è un elemento

- Partecipiamo alle attività dell'EGI Federated Cloud
- Alcune soluzioni possono essere utili anche per la Cloud dell'Area Padovana
 - Es. automazione della distribuzione e contestualizzazione di virtual appliance pubblicate in un marketplace tramite vmcatcher
- RC a Padova quasi pronto per entrare in fase di certificazione
 - E' un'istanza diversa da quelle di (pre)produzione
- Uno dei due progetti H2020 ICT-07 sottomessi ad Aprile in cui INFN-Padova e' coinvolto ha come riferimento la EGI FedCloud



- Progetto recentemente approvato, promosso da A. Garfagnini
- Coinvolge diversi dipartimenti (tra cui Fisica e Astronomia)
- Le risorse hardware saranno ospitate nella prima fase c/o il dipartimento di Fisica e Astronomia
- E` in corso la gara per l'acquisto dell'hardware
- MoU tra INFN e Universita` in fase di discussione
 - L'idea e` di mettere in qualche modo insieme queste 2 cloud

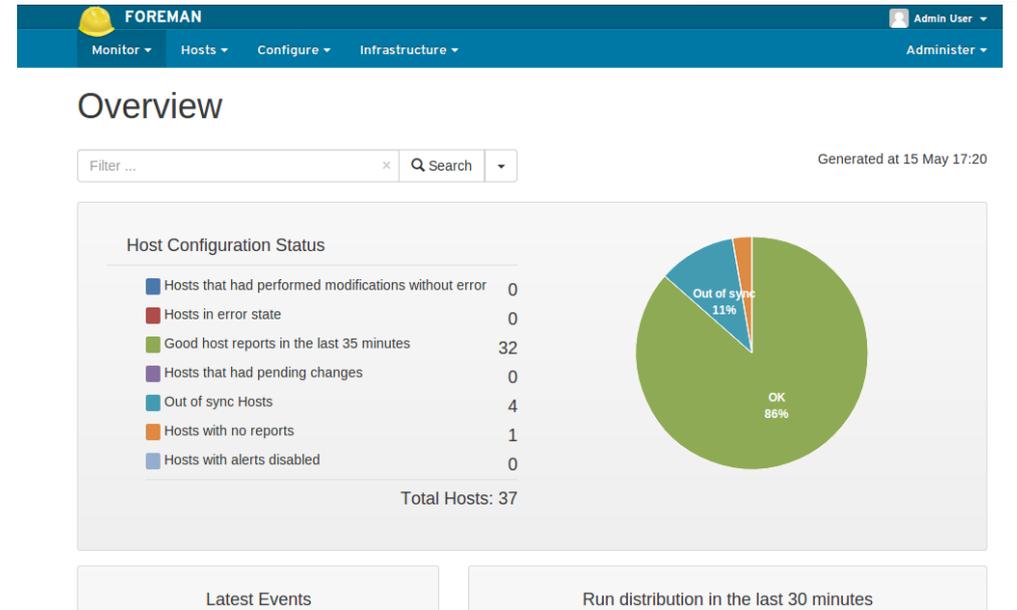
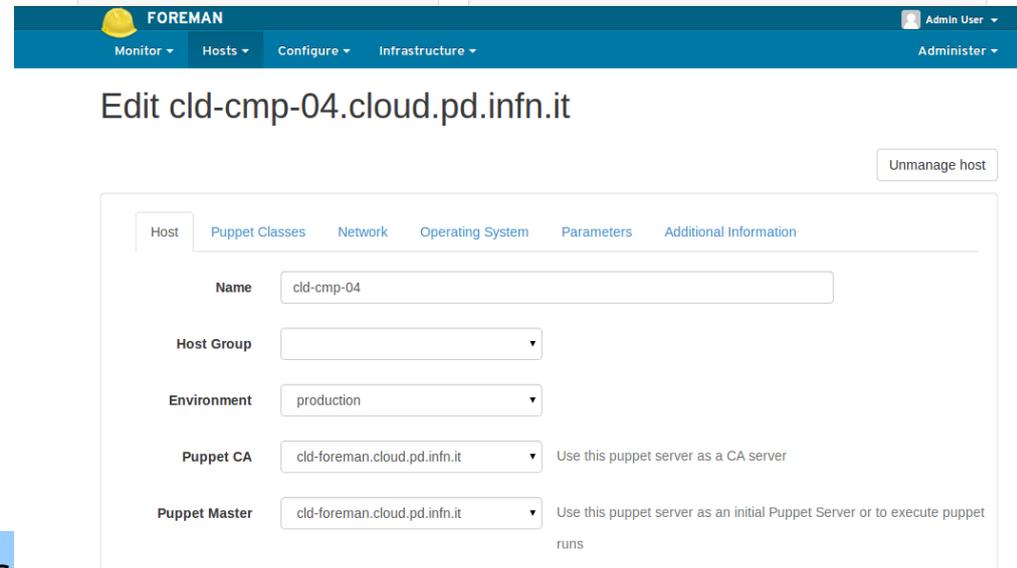
- Padova:
 - Cristina Aiftimiei, Paolo Andreetto, Sara Bertocco, Alberto Crescente, Fulvia Costa, Alvisè Dorigo, Eric Frizziero, Federica Fanzago (inizia ora), Michele Michelotto, Massimo Sgaravatto, Sergio Traldi, Marco Verlato, Lisa Zangrando
- Legnaro
 - Massimo Biasotto , Sergio Fantinel, Michele Gulmini, Massimo Venaruzzo
- Lunga lista di persone ma nessuno lo fa a tempo pieno, e le varie persone continuano a fare anche quello che facevano prima (Servizio Calcolo, Grid middleware, Grid operations, Tier-2, attività` negli esperimenti)
- Molto positiva la condivisione delle diverse competenze
- Contributo **fondamentale** da parte del personale precario

- Servizio di preproduzione disponibile a utenti skillati e selezionati che vogliono iniziare a provare la cloud
- Altre funzionalità e risorse verranno integrate progressivamente
- Sinergia con le attività del Cloud Working Group della CCR e altre iniziative Cloud related
- Contatti: cloud@lists.pd.infn.it
- <http://wiki.infn.it/progetti/cloud-areapd/home>

THE END

Backup slides

- Gestione centralizzata e automatizzata dei nodi dell'infrastruttura Cloud via Foreman e Puppet
- Puppettizzazione dei diversi servizi graduale e progressiva
- Alcuni moduli presi in rete (in particolare Puppet Forge), altri sviluppati in casa

FOREMAN Edit cld-cmp-04.cloud.pd.infn.it

Unmanage host

Host | Puppet Classes | Network | Operating System | Parameters | Additional Information

Name: cld-cmp-04

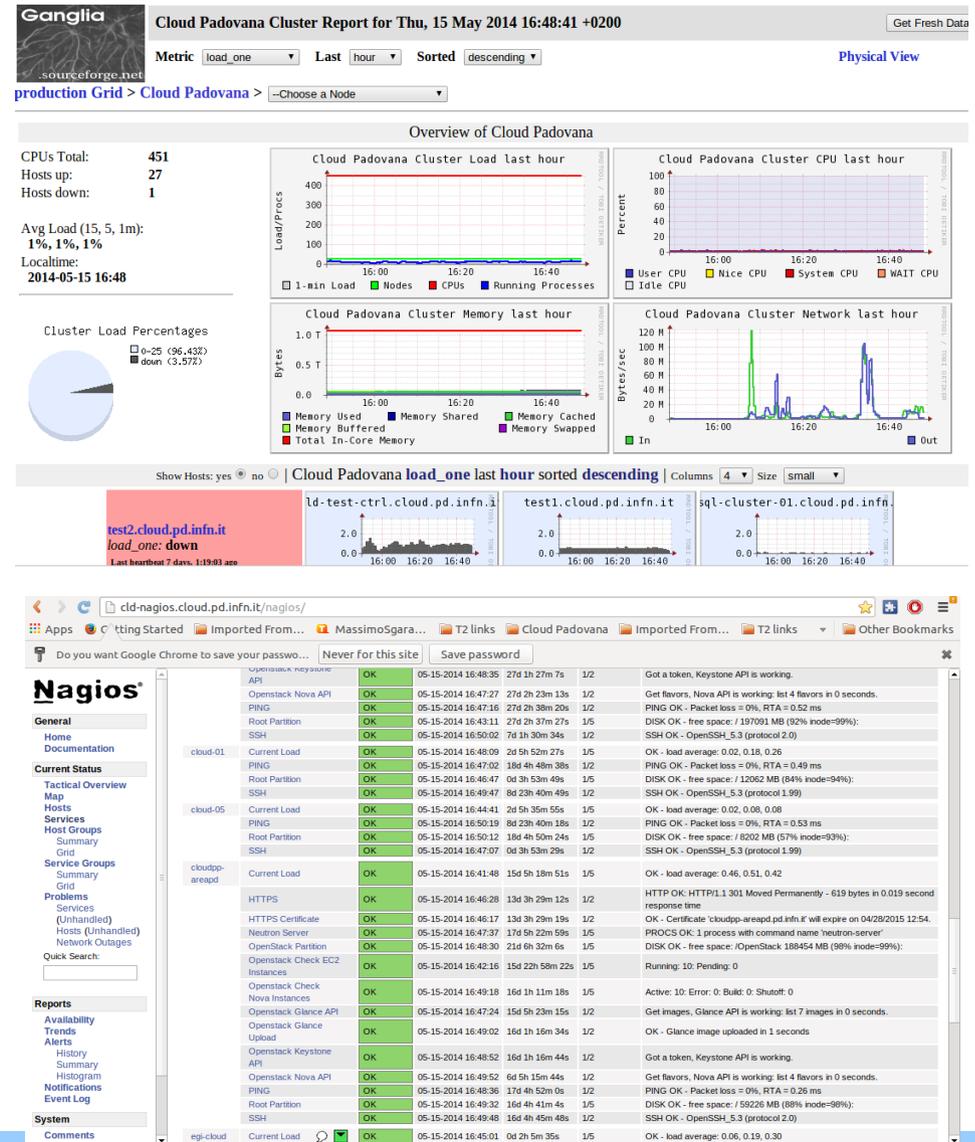
Host Group: [dropdown]

Environment: production

Puppet CA: cld-foreman.cloud.pd.infn.it Use this puppet server as a CA server

Puppet Master: cld-foreman.cloud.pd.infn.it Use this puppet server as an initial Puppet Server or to execute puppet runs

- Implementata una infrastruttura di monitoring basata su Ganglia e Nagios
- Sensori Nagios: alcuni presi in rete, altri sviluppati in casa



- La documentazione nella wiki dell'INFN (<http://wiki.infn.it/progetti/cloud-ar-eapd/home>)
- Le informazioni relative alle risorse hardware disponibili e il logbook delle operazioni effettuate vengono mantenute in DOCET
 - Tool sviluppato a Padova che permette la gestione di tutte le informazioni relative ad un data centre
 - Viene usato anche per le attività del Tier2



Cloud - Area Padovana

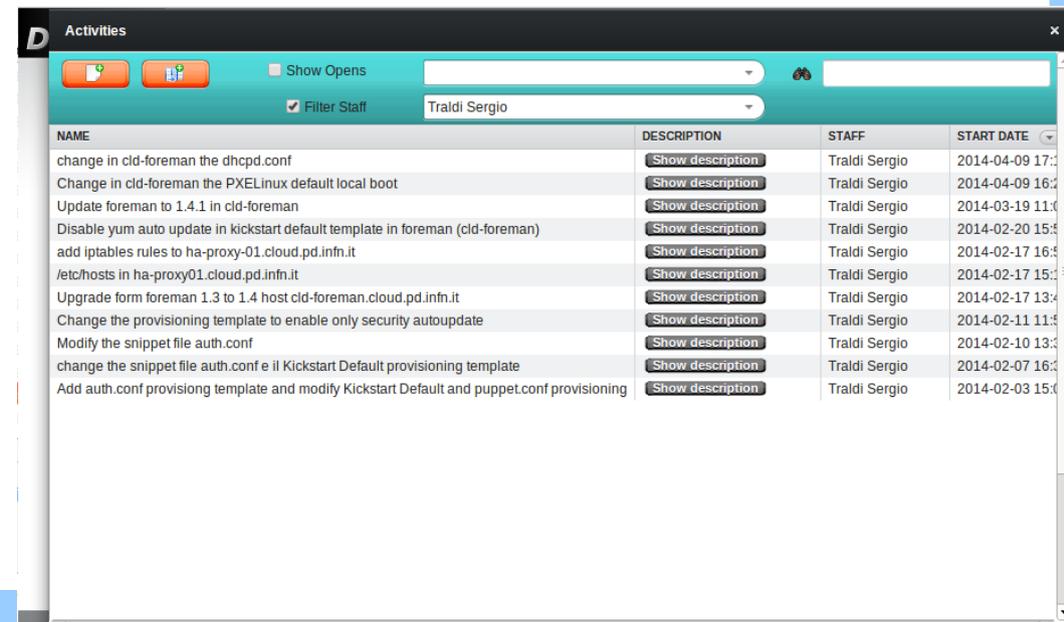
Questa è la wiki page per il progetto di Cloud dell'Area Padovana, il cui obiettivo è dotare la sezione INFN di Padova e i Laboratori Nazionali di Legnaro di un servizio di Cloud computing e storage.

Obiettivo è quello di implementare una stessa infrastruttura che possa essere utile per diversi aspetti del calcolo nelle due strutture INFN: supporto ad esperimenti, servizi e attività di progettazione, sviluppo e conduzione degli attuali e futuri acceleratori (ALPI, SPES).

A tal fine si utilizzerà in particolare il software OpenStack, uno dei middleware Cloud open-source più usati. Componenti innovative o particolari customizzazioni, sviluppate in particolare all'interno del Cloud Working Group della CCR dell'INFN (con cui si collabora attivamente) verranno messe in produzione quando ritenute sufficientemente mature.

Fisicamente le risorse di questa Cloud saranno distribuite su entrambe le sedi.

Una infrastruttura Cloud comune INFN LNL-Sezione di Padova perseguirà la storia di collaborazione tra le due strutture (il Tier-2 Legnaro-Padova degli esperimenti ALICE e CMS ne è un esempio) e potrà beneficiare dell'esperienza già acquisita, nonché di parte dell'infrastruttura già esistente.



NAME	DESCRIPTION	STAFF	START DATE
change in cld-foreman the dhcpd.conf	Show description	Traldi Sergio	2014-04-09 17:00
Change in cld-foreman the PXELinux default local boot	Show description	Traldi Sergio	2014-04-09 16:00
Update foreman to 1.4.1 in cld-foreman	Show description	Traldi Sergio	2014-03-19 11:00
Disable yum auto update in kickstart default template in foreman (cld-foreman)	Show description	Traldi Sergio	2014-02-20 15:00
add iptables rules to ha-proxy-01.cloud.pd.infn.it	Show description	Traldi Sergio	2014-02-17 16:00
/etc/hosts in ha-proxy01.cloud.pd.infn.it	Show description	Traldi Sergio	2014-02-17 15:00
Upgrade form foreman 1.3 to 1.4 host cld-foreman.cloud.pd.infn.it	Show description	Traldi Sergio	2014-02-17 13:00
Change the provisioning template to enable only security autoupdate	Show description	Traldi Sergio	2014-02-11 11:00
Modify the snippet file auth.conf	Show description	Traldi Sergio	2014-02-10 13:00
change the snippet file auth.conf e il Kickstart Default provisioning template	Show description	Traldi Sergio	2014-02-07 16:00
Add auth.conf provisioning template and modify Kickstart Default and puppet.conf provisioning	Show description	Traldi Sergio	2014-02-03 15:00

- Accesso alla Cloud consentito a tutti gli utenti INFN di Padova e Legnaro
- Accesso anche ad altri utenti INFN, con un numero limitato di risorse disponibili
- Tutti gli utenti devono accettare le Acceptable Usage Policy (AUP) del servizio (http://wiki.infn.it/progetti/cloud-areapd/documents/acceptable_use_policy)
- Openstack organizza gli utenti in Progetti
 - Un progetto rappresenta un gruppo/esperimento
 - Ad ogni progetto e` associata una quota sull'uso delle risorse
 - Un utente puo` appartenere anche a piu` progetti
- Possibilita` di provare la cloud (considerando un numero limitato di risorse) riservata a tutti gli utenti registrati
 - Progetto "Guests" con un numero limitato di risorse
- Progetti per singoli esperimenti/gruppi che hanno messo risorse dedicate nella Cloud

- Policy analoga a quella del Cern
 - Immagini pubbliche create dagli admin]
 - Ogni utente puo` registrare la propria immagine (ma non puo` essere pubblica)
 - Possibilita` di share di immagini tra diversi progetti
- Stiamo predisponendo una macchina che possa essere usata per costruirsi immagini (usando il software OZ) usabili per il proprio use case
- Posso anche creare una VM specificando uno snapshot
 - Si ` crea una VM usando una certa immagine esistente
 - Su questa VM si fanno le modifiche che interessano (es. installazione di pacchetti software supplementari)
 - Si salva uno "snapshot"
 - Si puo` usare questa snapshot per creare altre VMs