



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

Cloud dell'area Padovana Status update

- Progetto partito circa un anno fa
- Obiettivo: implementazione di un servizio IaaS tra la Sezione di Padova e i Laboratori Nazionali di Legnaro
 - Per evitare la proliferazione dei cluster dei diversi gruppi/esperimenti
 - Uso più efficiente delle risorse
 - Gestione sistemistica meno onerosa
- Progetto comune tra Padova e LNL per sfruttare al meglio le infrastrutture, il manpower, le competenze in entrambi i siti
 - Vista anche l'ottima esperienza del Tier-2

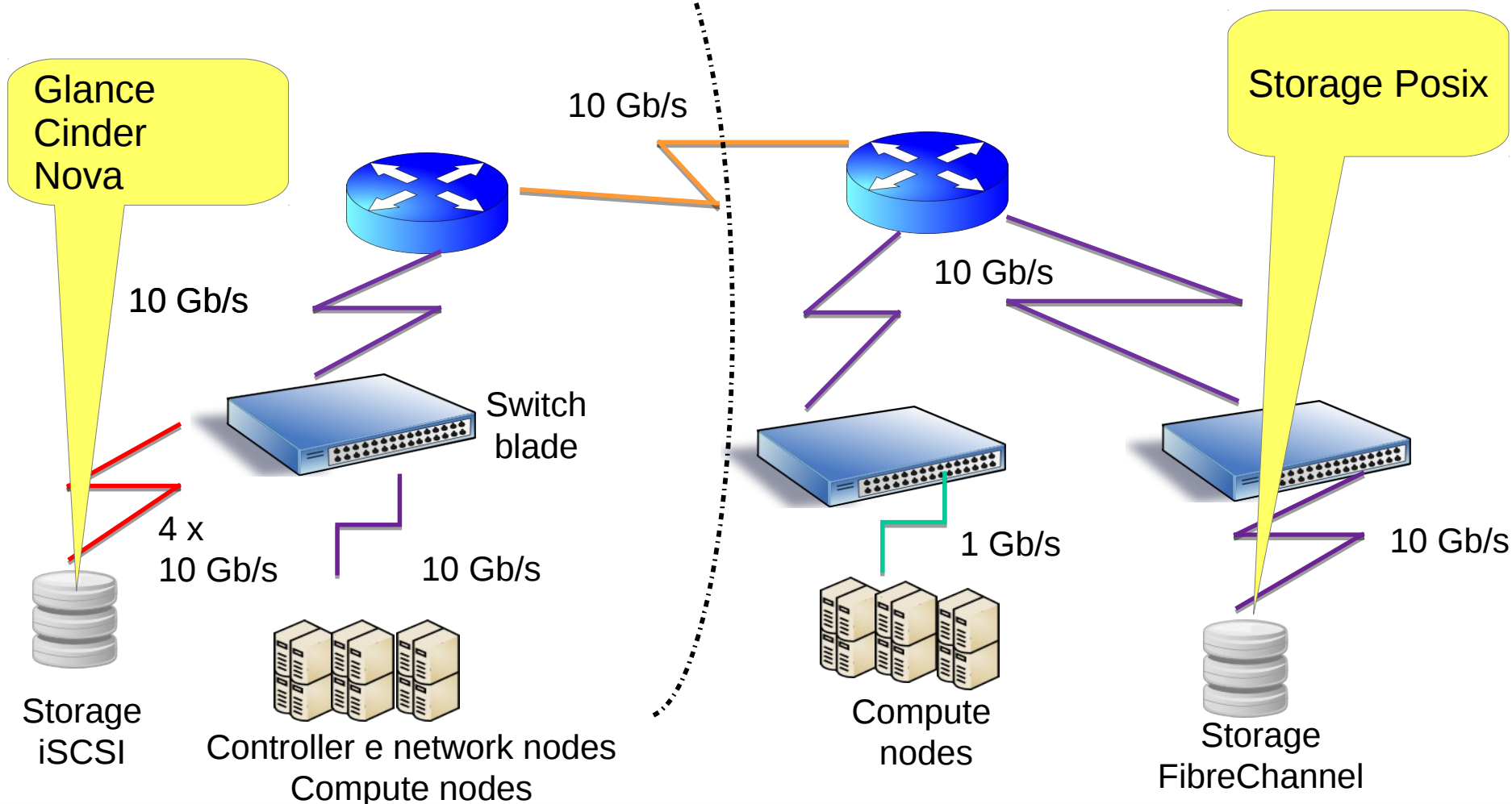
Preproduzione --> Produzione

- Attivo da Maggio un servizio di pre-produzione con funzionalita` e risorse limitate
 - <https://cloudpp-areapd.pd.infn.it:8443/dashboard>
 - Versione Havana di OpenStack su SL6
 - 1 controller node a Padova, 3 compute node (2 a Padova e 1 a Legnaro)
 - E` stato utile sia a noi che ai primi utenti
- E` ora disponibile un servizio di produzione
 - <https://cloud-areapd.pd.infn.it:443/dashboard>
 - Il servizio di pre-produzione verra` dismesso quando tutti gli utenti avranno completato la migrazione

- Rispetto al servizio di pre-produzione
 - Servizi configurati in high availability
 - “Secure” dei servizi via SSL
 - Accesso sicuro ai servizi anche via CLI e non solo via Dashboard/HTTPS
 - Implementato a livello di HAProxy
 - Autenticazione possibile sia attraverso INFN-AAI che attraverso username/password
 - Procedura per la registrazione degli utenti e dei progetti integrata nella Dashboard
 - Supporto per live migration
- OpenStack release Havana su SL6
 - Pianificata migrazione ad Icehouse

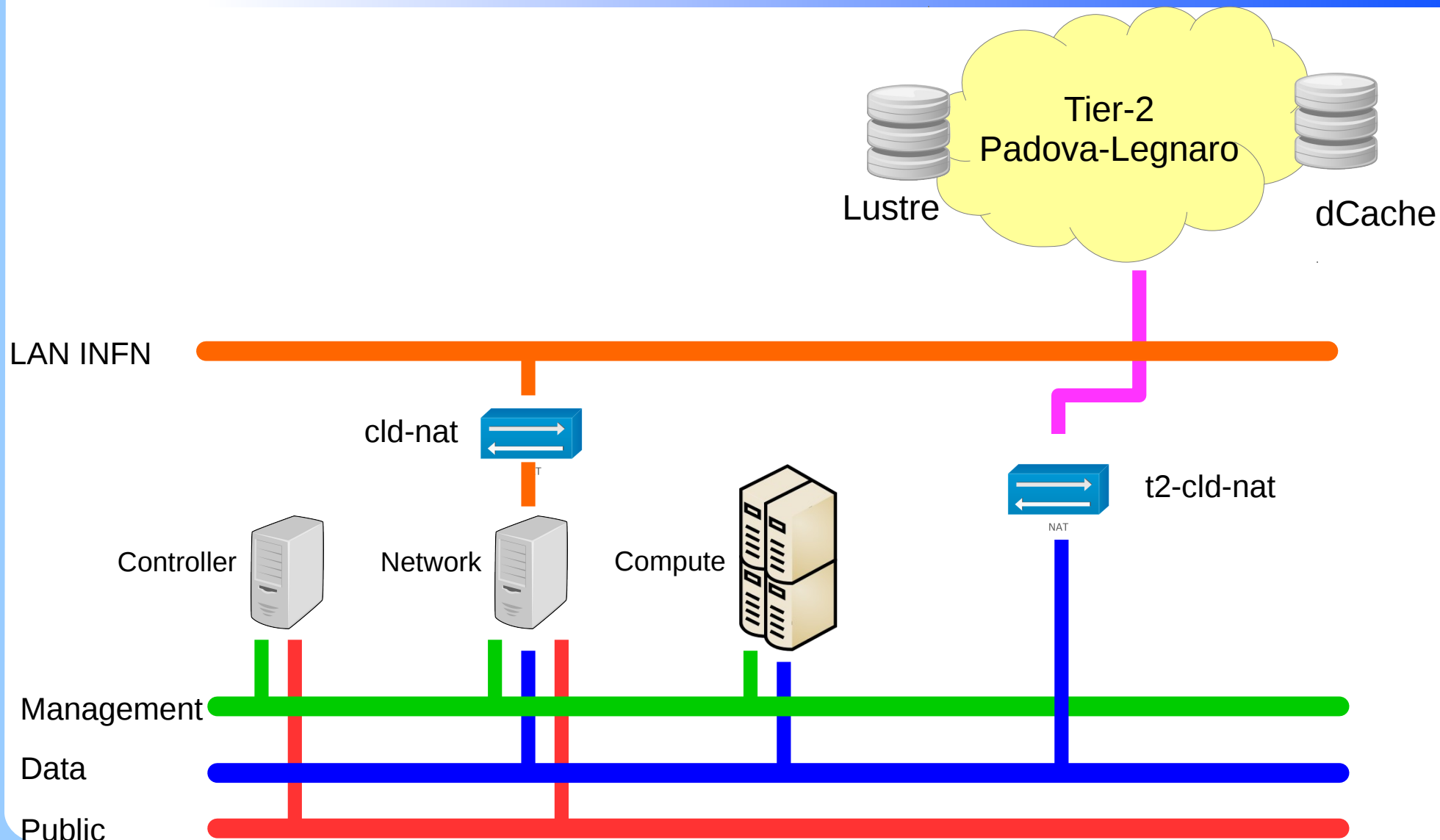
Padova

LNL



Hw per la Cloud di (pre)produzione

- A Padova
 - 4 lame DELL M620 ciascuna con 1 processore E52609, 32 GB RAM
 - 2 usate come controller node, 2 usate come network node in High Availability
 - 5 lame DELL M620 ciascuna con 2 processori E52670v2 (20 core fisici), 80 GB RAM
 - Usate come compute node
 - Altre 3 lame in configurazione compute in arrivo
 - 1 server iSCSI DELL MD3620i, con 23 dischi SAS da 900 GB
 - Usato per le immagini (Glance), le istanze (Nova), i volumi (Cinder)
 - Altro cassetto 12x4TB da installare
 - Altre risorse (fisiche o VM nel cluster Proxmox di Sezione) per altri servizi
 - Foreman/Puppet, mysql, HAProxy/KeepAlived, Ganglia, Nagios, NAT
- A Legnaro
 - 4 Fujitsu Primergy RX300S7 con 2 processori XEON E52650v2 (16 core fisici), 96 GB RAM
 - Usate come compute node
 - Server storage DELL PowerVault MD3600f (FC) con 12 dischi da 4 TB
 - Da usare come storage Posix



- Neutron: OpenvSwitch driver con GRE tunneling
- Configurazione “Provider router with private networks”
- Possibilita` 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) e 2 istanze L3
- Prima istanza L3 per interfaccia esterna “classica” con indirizzo pubblico e NAT
 - Gestisce le reti 10.63.x.0/24 (una per progetto)
 - Per le VM che devono avere floating ip sulla rete pubblica

- Seconda istanza per una seconda interfaccia esterna senza NAT
 - Esce sulla rete usando come gateway una macchina esterna (cld-nat)
 - Gestisce le reti 10.64.x.0/24 (una per progetto)
 - VM mantengono nelle LAN i loro indirizzi IP 10.64.x.y privati
 - Escono in internet con IP nattedo (quello di cld-nat)
 - VM sono raggiungibili direttamente dalle LAN di Padova e Legnaro
 - Route statica sul router di sezione per raggiungere la rete 10.64 attraverso cld-nat
 - L'accesso dalle VM alle LAN di Padova e Legnaro regolato da un firewall via iptables

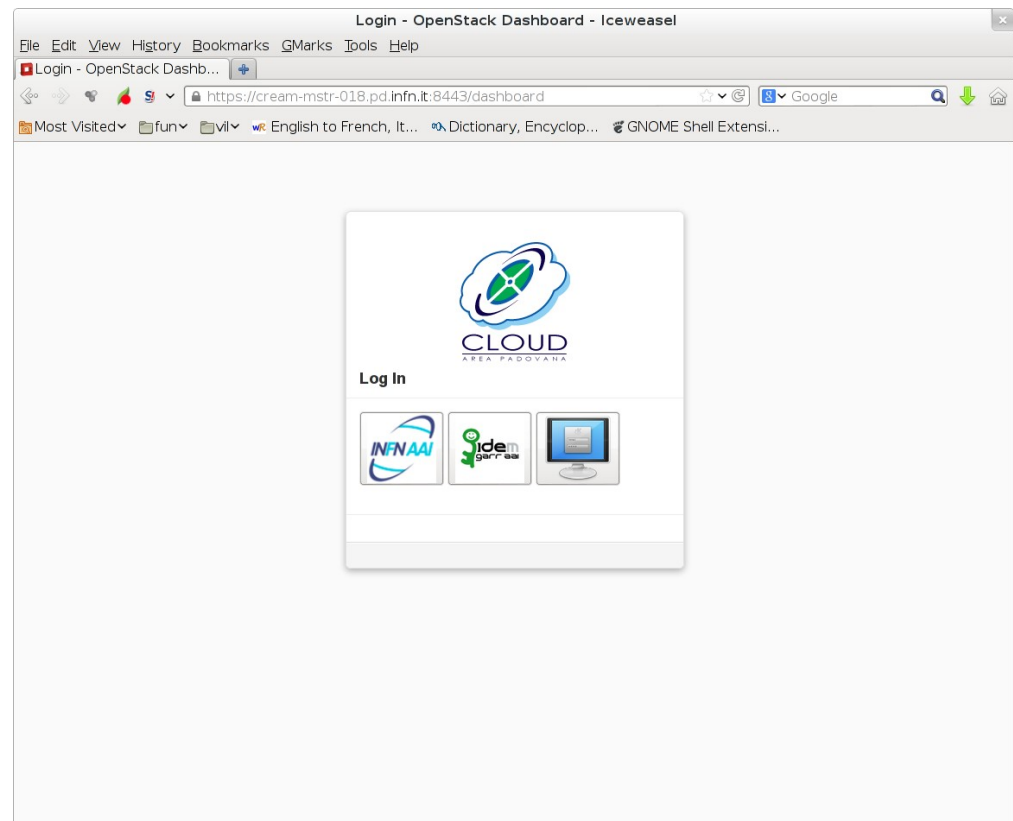
- Storage iSCSI a Padova usato per le immagini, per le istanze, per i volumi Cinder
 - Collegato ai 2 controller node (non e' stata una buona scelta !)
- Uso di glusterFS come backend
 - 2 brick, uno per controller node
- File system per le istanze condiviso tra i vari compute node
- Non usiamo Swift

- Oltre a Cinder prevediamo la possibilità` anche di accedere a storage “esterno” alla Cloud
- Definita una configurazione che evita il (possibile) bottleneck del Network Node
 - Neutron L2 installato e configurato sullo storage server, in modo che lo storage “entri” nella Cloud
 - http://wiki.infn.it/progetti/cloud-areapd/best_practices/storage_external
- Implementata per il momento per storage di Muon-Tomography (Gluster)
- Implementazione quasi ultimata per storage LHCb (Lustre)

High Availability (HA)

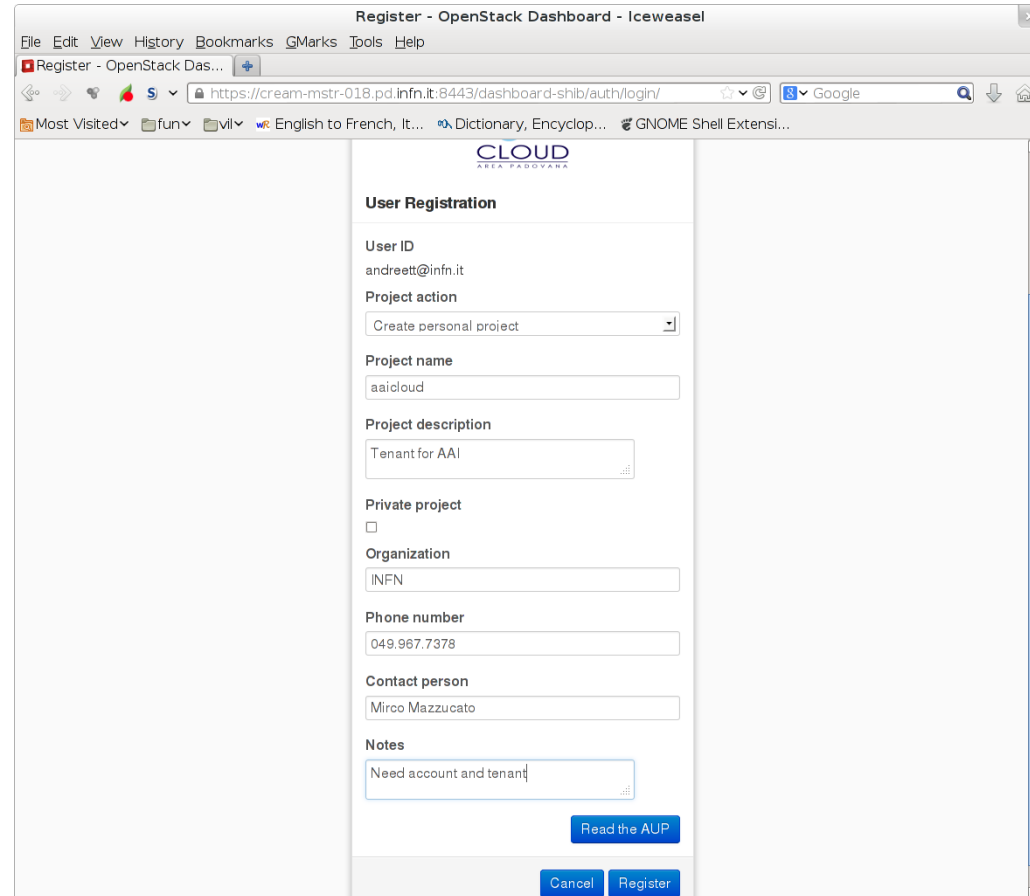
- High Availability in modalita` Active/Active
- Cluster HAProxy/Keepalived formato da 3 host per i servizi da configurare in HA (database, endpoint dei servizi OpenStack)
 - Uso di 3 VM (su 3 hypervisor diversi) del cluster Proxmox di Sezione
- Database
 - Cluster Mysql HA usando Percona XtraDB MultiMaster formato da 3 host
 - Uso di 3 VM (su 3 hypervisor diversi) del cluster Proxmox di Sezione
- 2 controller node in HA
 - Keystone, Glance, RabbitMQ, Nova API, Nova Schedd, Nova Console, Neutron Server, Cinder API, Cinder Schedd, Cinder Volume
 - RabbitMQ in cluster mode “nativo” (senza l'uso di HAProxy) con ha_queues
- 2 Network Node in HA
 - Neutron Agents (L2, L3, Metadata, DHCP), Cinder Volume
 - Migrazione router da un L3 Agent ad un altro gestita al momento manualmente

- Possibilita` di autenticarsi alla Cloud sia attraverso username e password, sia via INFN/AAI
- Integrazione con INFN-AAI implementata attraverso SAML2 (Shibboleth)
- Integrazione con IDEM in progress
- INFN-AAI/IDEM viene usato solo per l'autenticazione: autorizzazione gestita localmente



Supporto per registrazione

- Per ogni utente e' richiesta una registrazione, che si fa via Dashboard
- Nella registrazione (o in un secondo momento) si puo' richiedere la creazione di un nuovo progetto o l'inserimento in uno o piu' progetti esistenti
- Un progetto puo` essere "pubblico" (in genere un esperimento/gruppo di ricerca) o "personale" (concesso solo se ci sono validi motivi)
- Progetto "guest" con poche risorse per utenti che vogliono provare il servizio

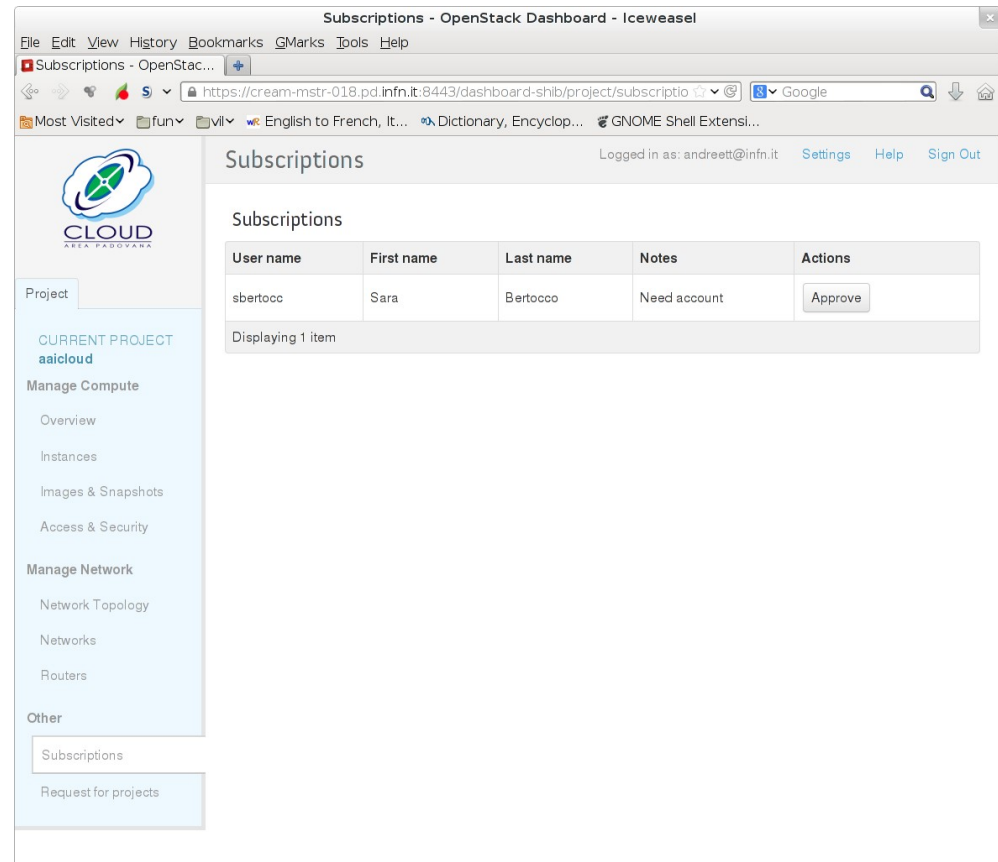


The screenshot shows a web browser window titled "Register - OpenStack Dashboard - Iceweasel". The address bar shows the URL: <https://cream-mstr-018.pd.infn.it:8443/dashboard-shib/auth/login/>. The page displays the "User Registration" form with the following fields and values:

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

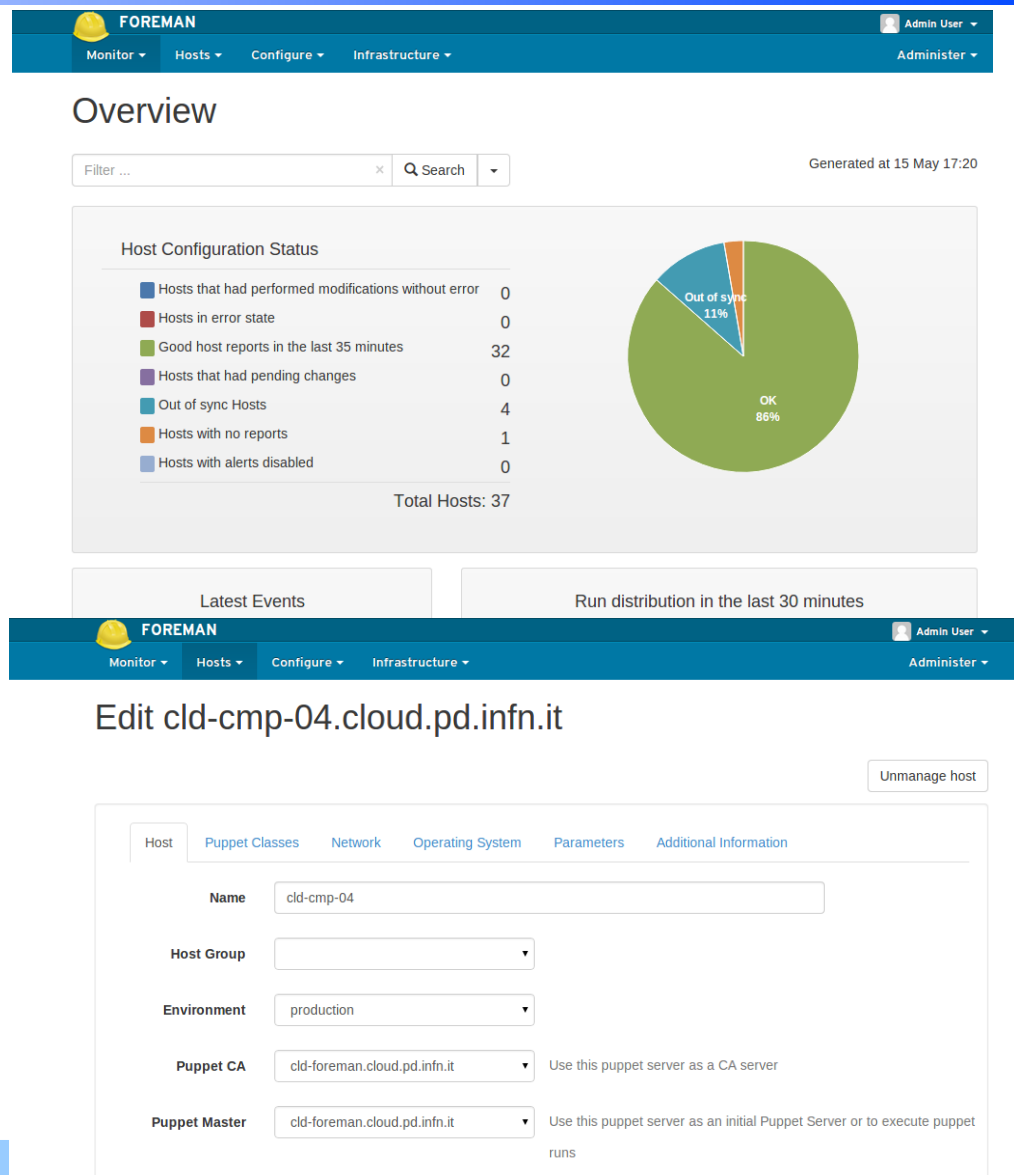
At the bottom of the form, there are three buttons: "Read the AUP", "Cancel", and "Register".

- 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
 - Approvazioni (sia per il cloud admin che per il project admin) effettuate via Dashboard
- Notifiche via mail per cloud admin/project admin/utente



- Sviluppo di uno scheduler avanzato di OpenStack per:
 - Gestire l'accodamento delle richieste che non possono essere immediatamente soddisfatte (perche` non ci sono risorse)
 - Utilizzo delle risorse in modalita` fair-share tra diversi progetti (esperimenti) senza necessita` di un partizionamento statico delle risorse
- Dettagli nella presentazione di Lisa Zangrando

- Gestione centralizzata e automatizzata dei nodi dell'infrastruttura Cloud via Foreman e Puppet
- Installazione e configurazione dei compute node implementata integralmente via Foreman/Puppet
- Per gli altri nodi Foreman/Puppet viene usato per l'installazione dell'OS e per la configurazione di alcuni servizi (ntp, ganglia gmond, nagios, etc.)
- Alcuni moduli presi in rete (in particolare Puppet Forge), altri sviluppati in casa



FOREMAN Overview

Filter ... Search

Generated at 15 May 17:20

Host Configuration Status

Hosts that had performed modifications without error	0
Hosts in error state	0
Good host reports in the last 35 minutes	32
Hosts that had pending changes	0
Out of sync Hosts	4
Hosts with no reports	1
Hosts with alerts disabled	0
Total Hosts:	37

Host Configuration Status Legend:

- OK: 86%
- Out of sync: 11%

Latest Events Run distribution in the last 30 minutes

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]

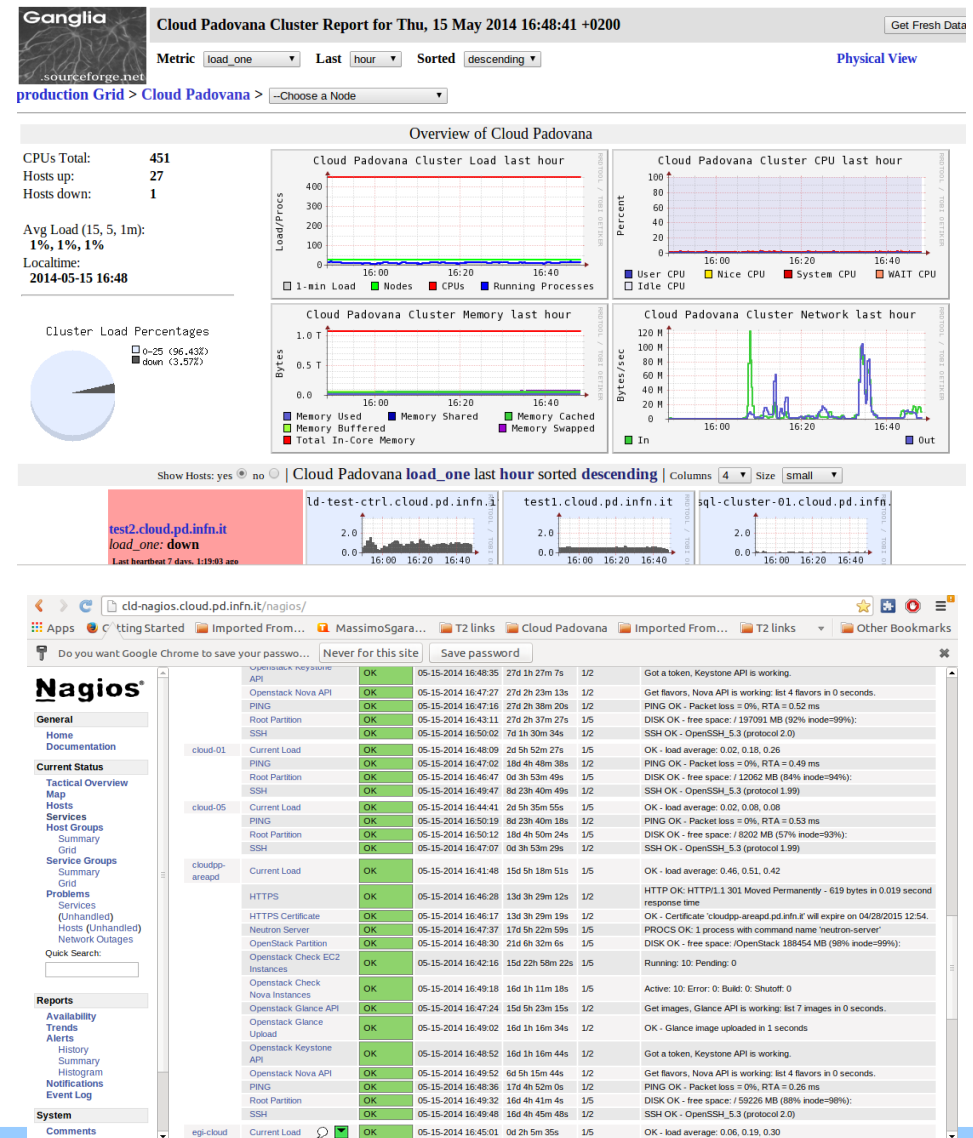
Environment: production

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

Monitoring dell'infrastruttura

- Implementata una infrastruttura di monitoring basata su Ganglia e Nagios
- Sensori Nagios per controllare la funzionalita` dei diversi servizi
- Alcuni presi in rete, altri sviluppati in casa



- Documentazione per admins nella wiki dell'INFN
<http://wiki.infn.it/progetti/cloud-area-padova/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 Tier-2 di Padova-LNL
- User guide per end-user disponibile in diversi formati
 - Scritta in DocBook (publican)
- Previsto seminario rivolto agli utenti in Sezione a inizio anno



Cloud - Area Padova

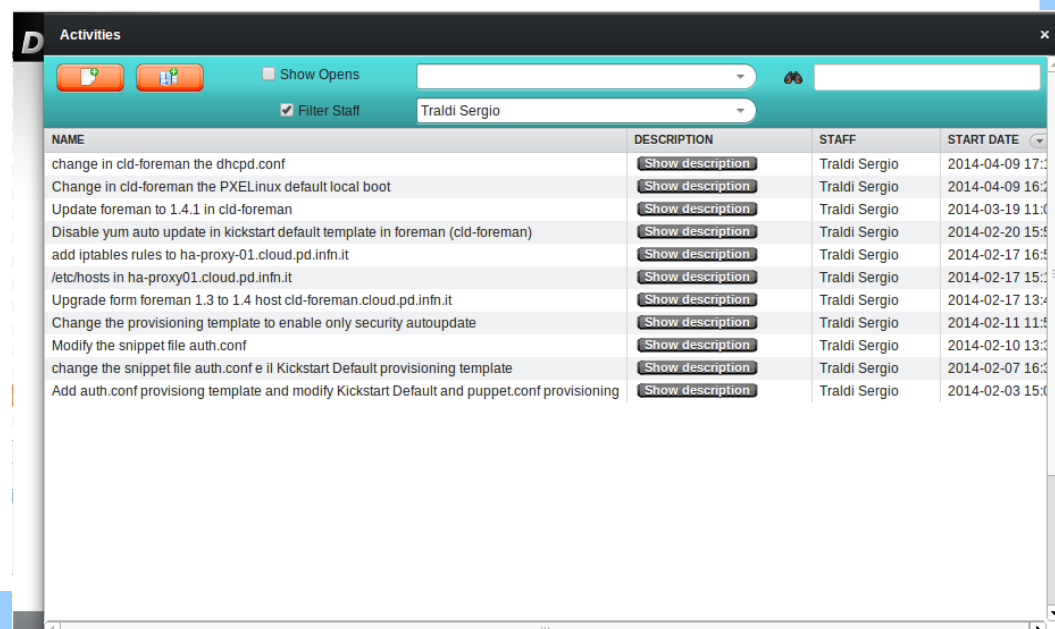
Questa è la wiki page per il progetto di Cloud dell'Area Padova, il cui obiettivo è dotare la sezione INFN di Padova e i Laboratori Nazionali di Legnaro di un servizio di Cloud computing e storage.

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 proseguirà 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

- Progetti attualmente attivi
 - ALICE
 - Virtual Analysis Facility (PRIN STOA-LHC)
 - Belle II
 - CMS
 - Analisi dati 8 e 13 TeV: ricerca di nuova fisica nel settore dell'Higgs, Z', Top e B-physics
 - CUORE
 - Analisi e confronto in frequenza ed energia del migliaio di canali bolometrici che realizzano il rivelatore Cuore
 - LHCb
 - GAMMA
 - Analisi dati e simulazioni per gli esperimenti AGATA e GALILEO
 - JUNO
 - Monte Carlo dell'esperimento: simulazione e ricostruzione
 - Muon Tomography
 - Uso di programmi per la generazione di immagini tomografiche tramite ricostruzione statistica degli eventi di diffusione dei mu all'interno di volumi estesi
 - OCP
 - Test di deployment di CloudFoundry

- Altre attività` Cloud related a Padova
 - EGI Federated Cloud
 - Cloud INFN multi-regione
 - OCP
 - Attività` Cloud CMS
- ... ma di queste non parlero` visto che sono state/saranno discusse in altre presentazioni

- Progetto promosso da A. Garfagnini che coinvolge 10 dipartimenti (tra cui Fisica e Astronomia)
 - Finanziamento dell'Ateneo + co-finanziamento da parte dei diversi dipartimenti
- Risorse hw acquistate:
 - 4 lame con due E5-2609 (2.4 GHz), 32 GB di RAM, 2 dischi SAS da 300GB
 - 12 lame con 2 E5-2670v2 (2.5 Ghz), 160 GB RAM, 2 dischi SAS da 300GB
 - 2 moduli DELL Force 10 MXL 10/40 GbE DCB Switch
 - Storage Equallogic con 17x1.2TB SAS 10000 giri (20 TB) + 7 x 800GB SSD (6 TB)
 - Storage Equallogic con 24 x 4TB SAS 7200 giri (96 TB)
- Queste risorse sono ospitate nella prima fase c/o il dipartimento di Fisica e Astronomia

- Lettera di intenti tra i 10 dipartimenti, INFN Padova e INFN Legnaro “per lo sviluppo, messa in opera e sperimentazione di un “Centro pilota di Elaborazione Dati Cloud a Padova – CED-C ” ad Alte Prestazioni a sostegno della ricerca di UNIPD e INFN- PDLNL”
 - Obiettivi
 - “unire gli sforzi per lo sviluppo e la messa in opera dell’attrezzatura avanzata di cloud computing tra UNIPD e INFN- PDLNL in modo da poter approfittare delle competenze già esistenti e delle esperienze già fatte, oltre che delle possibili sinergie, facilità e uniformità di accesso, gestione e crescita comune delle competenze nel campo delle tecnologie cloud open source,
 - dotarsi di appositi strumenti di coordinamento durante la fase di sviluppo e messa in funzione;
 - operare per il raggiungimento di un accordo di ampio respiro, affinché il CED-C possa svolgere un ruolo di riferimento tecnologico e di competenze a livello regionale non solo per il mondo della ricerca, ma anche per le imprese e per la Pubblica Amministrazione.”
- Lettera in questi giorni alla firma dei direttori
- Governance: comitato di gestione costituito da 5 membri
 - A. Garfagnini coordinatore scientifico, M. Sgaravatto responsabile tecnico

- Individuati alcuni tecnici Universitari che hanno iniziato a lavorare assieme a noi per l'implementazione di questa Cloud
- Installazione hardware completata
- Discussione in corso sull'architettura e sulle tecnologiche da adottare
 - Scelto OpenStack come middleware Cloud
 - Deciso di partire con IceHouse ma considerando CentOS 7 come OS (per facilitare successivi update)
 - Configurazione dei servizi in HA analoga a quanto fatto per la Cloud INFN
 - Discussione in corso sulla tecnologia da usare e sulla configurazione del backend di storage
 - Configurazione rete

- Molto difficile gestire un servizio di produzione usando un software che non e' production ready, ma noi ci proviamo lo stesso ...

THE END

Backup Slides

- Immagini pubbliche pubblicate dagli admins
 - Alcune prese in rete, altre create da noi
 - Immagini SL6.x in formato QCOW2, fully resizable, create con OZ
- Ogni utente puo` registrare la propria immagine (ma non puo` essere pubblica)
- Possibilita` di share di immagini tra diversi progetti
- Predisposta una macchina che puo` essere usata dagli utenti per costruirsi immagini (usando il software OZ)

- User guide
 - Scritta in DocBook (Publican)
 - Disponibile in diversi formati
 - “Ispirata” alla user guide del Cern
- Mailing list cloud-support@lists.pd.infn.it usata dagli utenti per riportare problemi
- Mailing list cloud-users@lists.pd.infn.it a cui vengono automaticamente iscritti tutti gli utenti usata per annunci di interventi programmati, ecc
- Valutando la possibilità di istituire turni di supporto

- Meeting settimanali in cui si discutono e si programmano le attività
 - Agenda e minute in Indico
(<https://agenda.infn.it/categoryDisplay.py?categId=658>)
- Attività tracciate come task in Jira
(<https://issues.infn.it/jira/browse/PDCL>)
- Software e documentazione in Github
(<https://github.com/CloudPadovana/>)