



cloudveneto

Massimo Sgaravatto
INFN Padova

Cloud computing per la ricerca: use case
dell'Istituto Nazionale di Fisica Nucleare

Calcolo scientifico nell'INFN

Perché Cloud computing ?

Esperienze Cloud @ INFN Padova
Cloud Area Padovana
CloudVeneto

Altre esperienze Cloud nell'INFN e nella ricerca

Istituto Nazionale di Fisica Nucleare (INFN): ente pubblico nazionale di ricerca, fondato nel 1951, dedicato allo studio dei costituenti fondamentali della materia e delle leggi che li governano

Attività di ricerca, teorica e sperimentale, nei campi della **fisica subnucleare, nucleare e astroparticellare.**

4 laboratori nazionali, e 20+ altre sedi

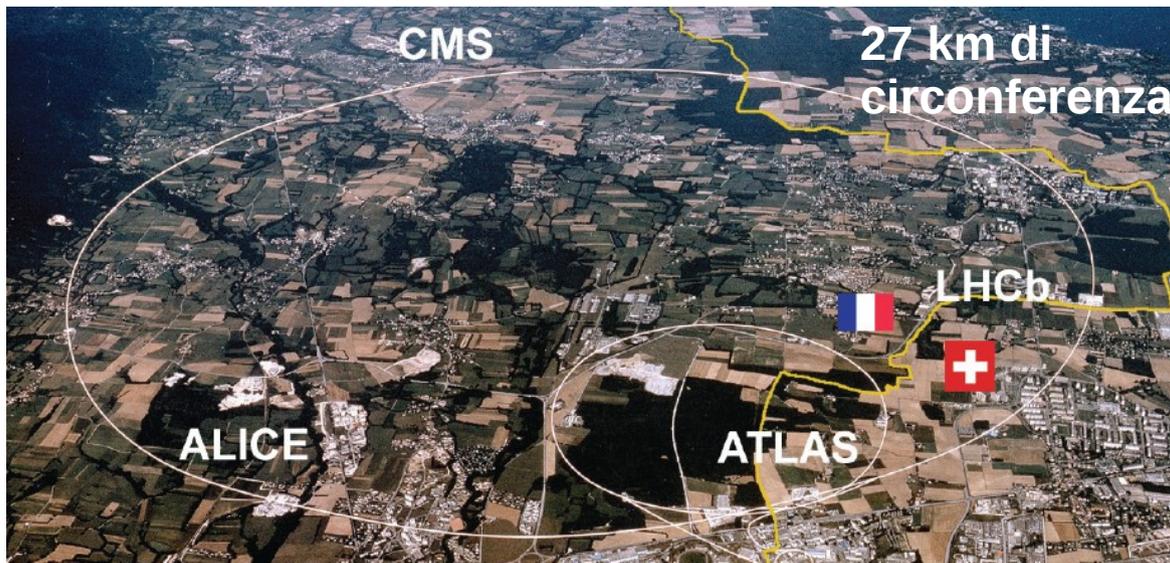
L'attività scientifica condotta dall'Istituto viene svolta nell'ambito di **collaborazioni internazionali**, sia presso proprie strutture in Italia, sia presso altri **laboratori internazionali**:

CERN (Ginevra), FERMILAB, SLAC, BNL e JLAB (Stati Uniti); PNPI, BINP e JINR (Federazione Russa); CIAE e IHEP (Cina); RIKEN e KEK (Giappone); BARC (India), DESY e GSI (Germania), ESRF (Francia), PSI (Svizzera) ecc.



Gli esperimenti più demanding dal punto di vista del calcolo sono gli esperimenti di **fisica delle alte energie** (HEP)

In particolare gli esperimenti **LHC** al **CERN** (ALICE, ATLAS, CMS, LHCb) sono di gran lunga i clienti maggiori



Qualche numero sul calcolo per gli esperimenti LHC

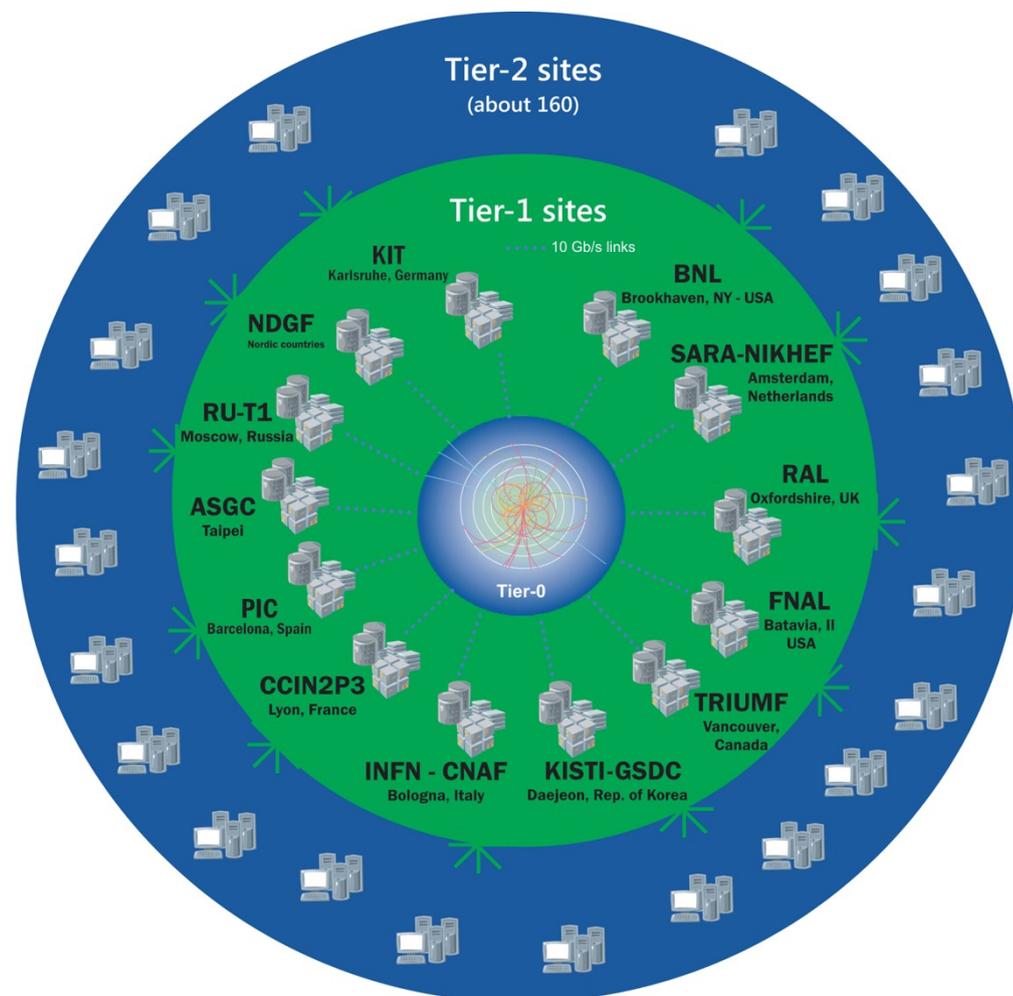
- ~ 900 Kcore
- ~ 470 PB disco
- ~ 750 PB nastro
- Risorse usate da **10K+ fisici** in tutto il mondo

Risorse distribuite

Gerarchia di centri calcolo:

- Tier-0 (Cern + Budapest)
- 13 Tier-1
- ~ 160 Tier-2

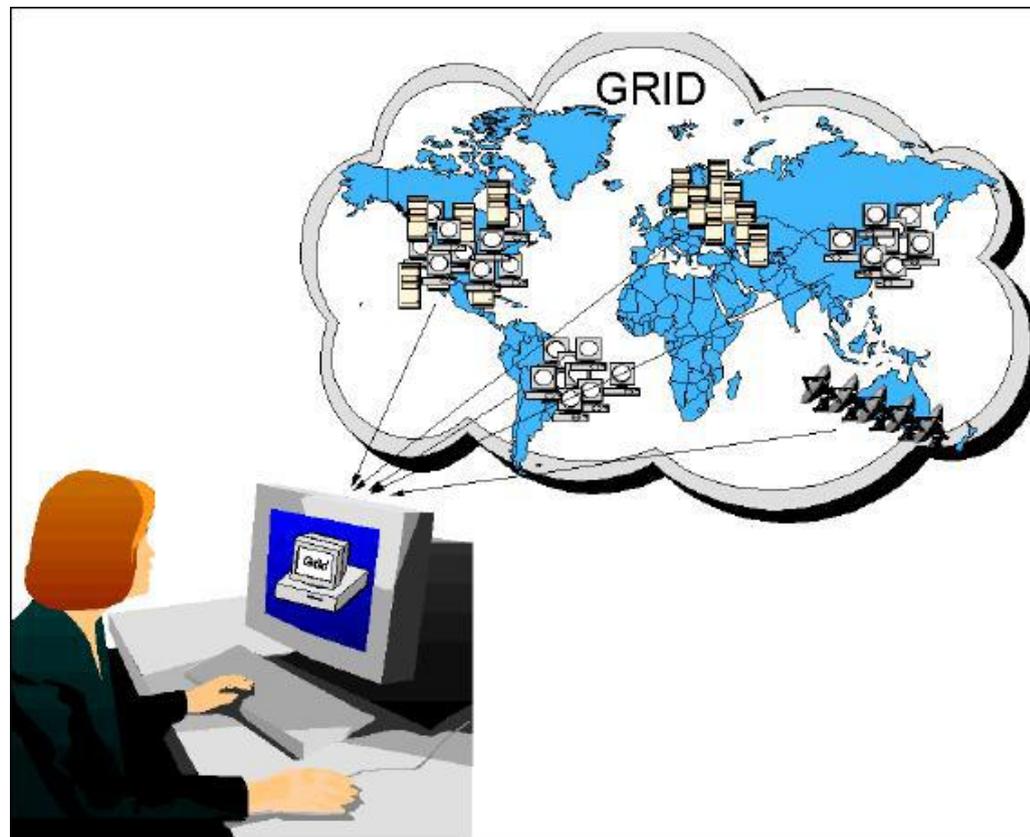
Grid computing per l'accesso a queste risorse



Grid computing

Il middleware **Grid** permette accesso trasparente a queste risorse

L'utente specifica caratteristiche e requirement delle sue applicazioni senza dover preoccuparsi di dove sono le risorse



Middleware Grid è **open-source**. Sviluppato nell'ambito di diversi progetti (molti EU-funded)

Grosso contributo INFN nello sviluppare, mantenere, operare questo middleware

In Italia

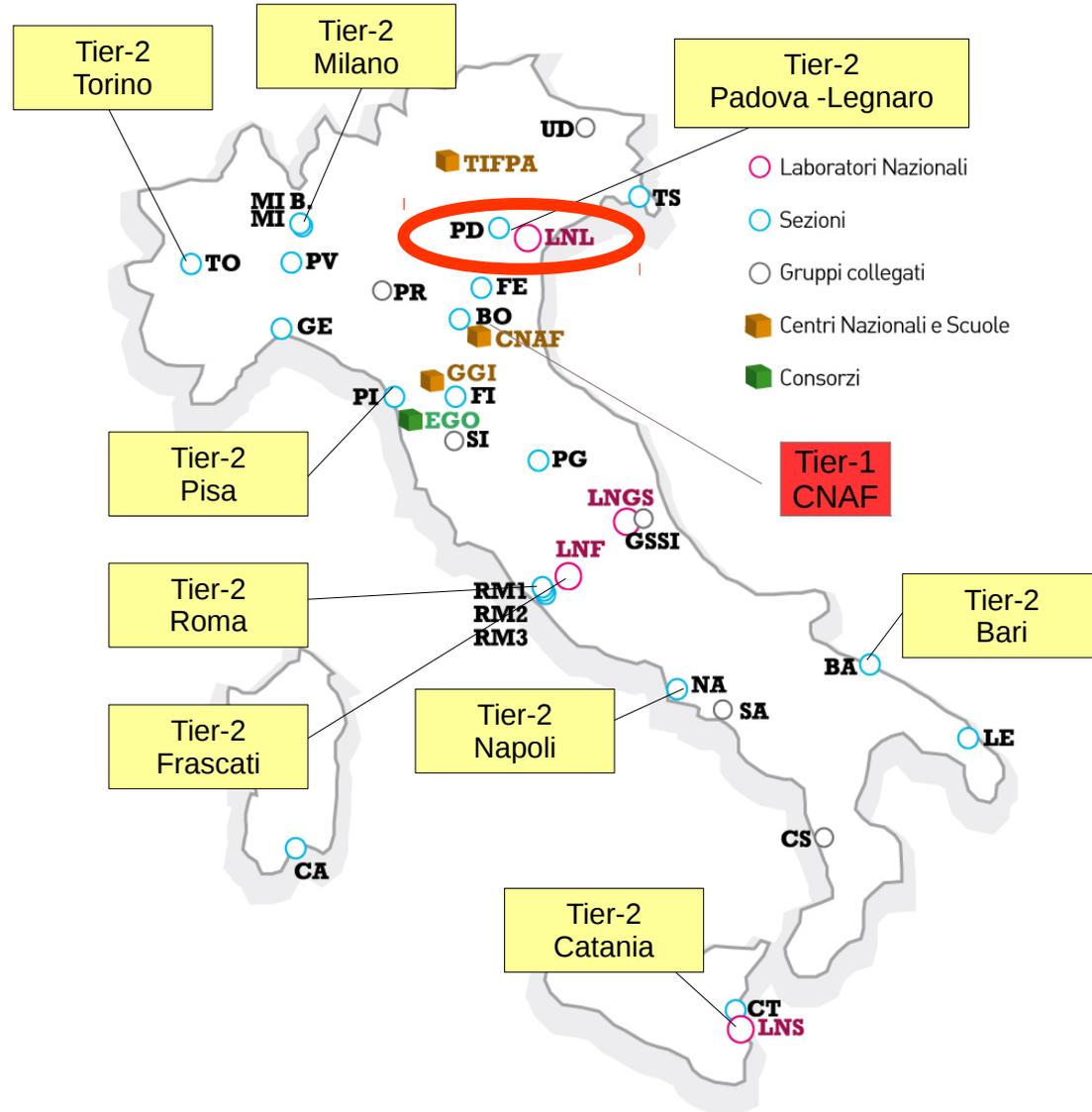
Tier-1 (INFN-CNAF) a Bologna

- ~ 26K core
- ~ 25 PB su disco
- ~ 70 PB su nastro

Centro di calcolo di riferimento per tutti gli esperimenti in cui INFN è coinvolto

9 Tier-2

- ~ 30 K core
- ~ 22 PB su disco



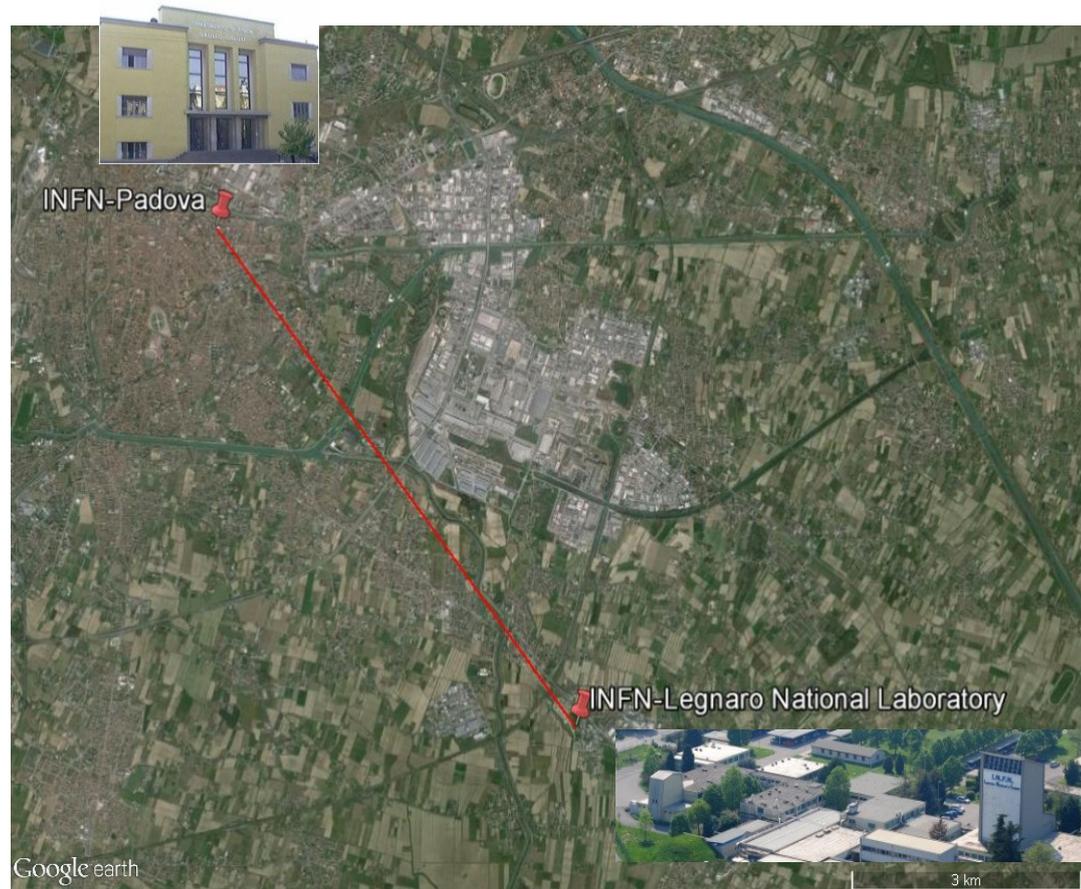
Tier-2 Padova-Legnaro

Singolo centro di calcolo ma con risorse distribuite in 2 siti

40 Gbps (dedicati) tra i 2 siti

Per il calcolo LHC:

- 240 worker node (6388 core)
- ~ 4 PB spazio disco per dati



Sale CED

15x7 mt

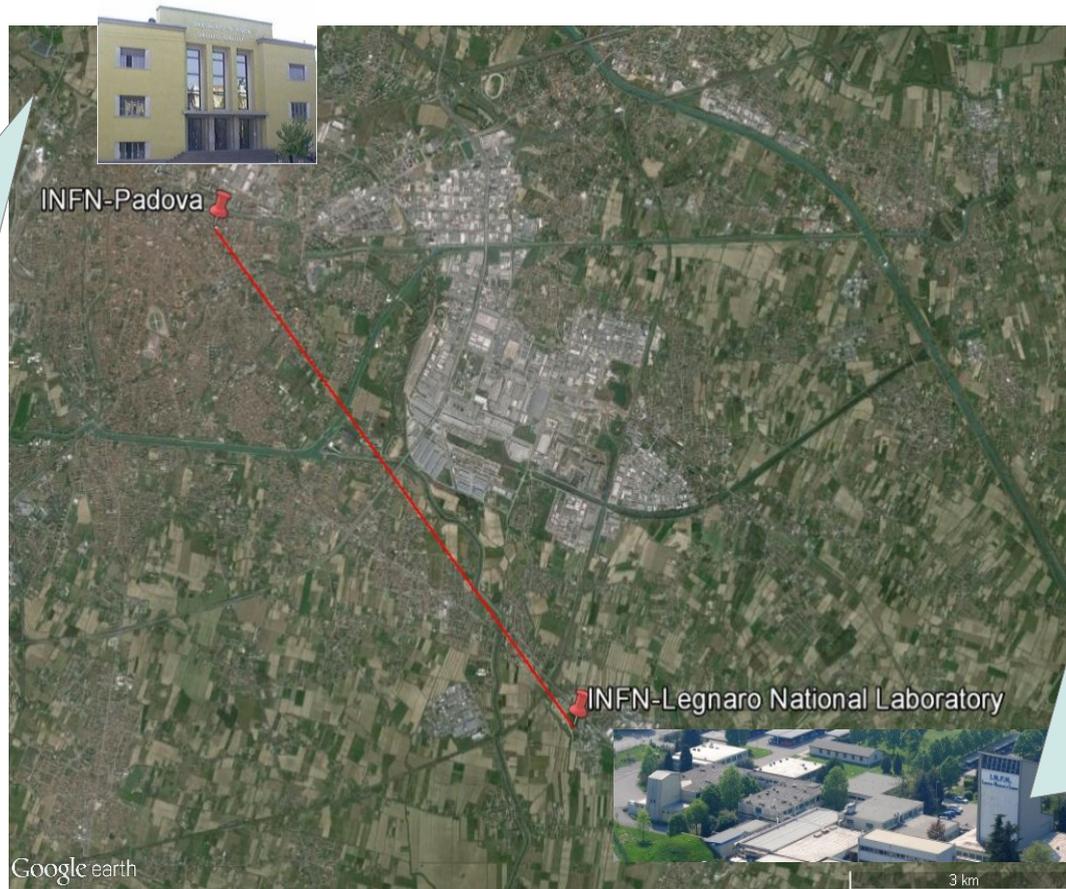
Rack di tipo standard

1 UPS da 200 KW

4 chiller da 40 KW

Impianto anticendio

Sala condivisa con
Il Dipartimento di
Fisica e Astronomia



10x7.4 mt

Rack APC con
cooler di
raffreddamento
in row

Gruppo elettrogeno
da 630 KW

5 UPS da 80 KW

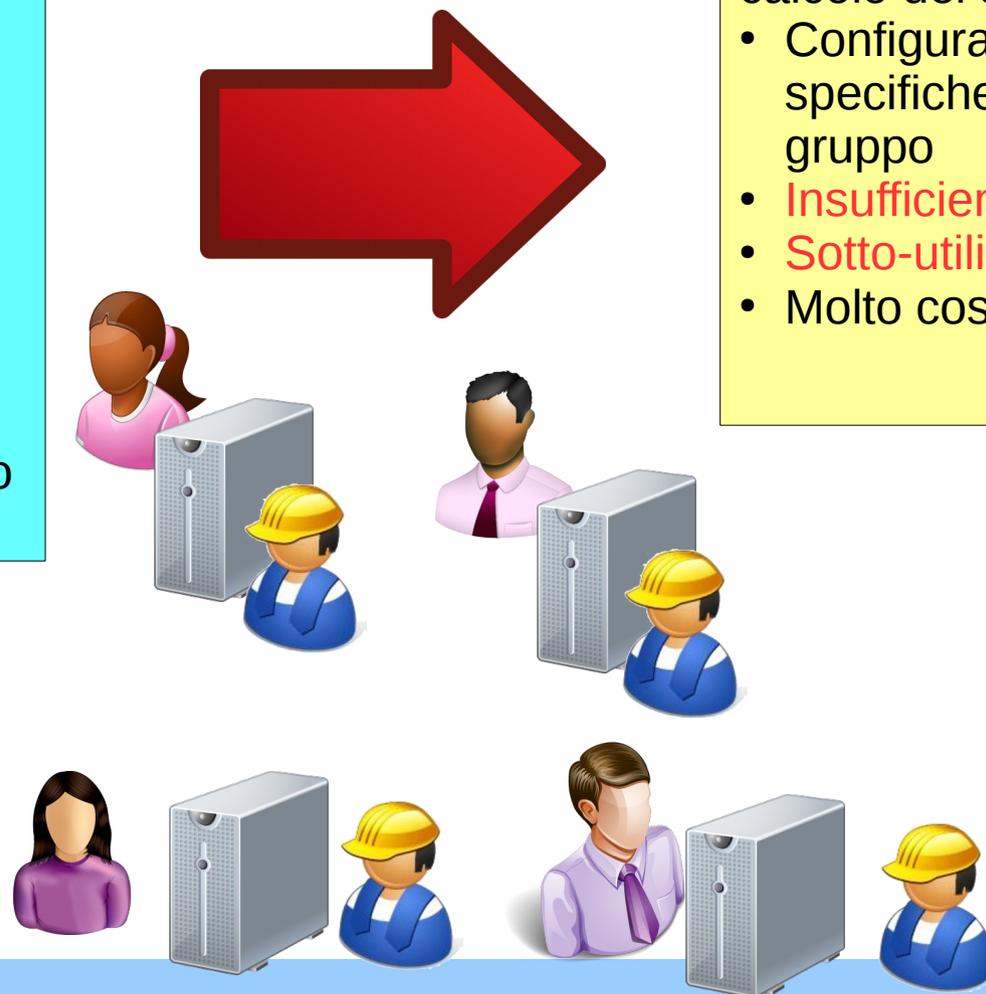
2 chiller da 105 KW

Impianto anticendio

Ma non tutto si fa in Grid ...

Non tutto il **calcolo scientifico** si riesce a fare in Grid

- Sostanzialmente usabile solo per attività di tipo batch
- Non semplicissima da usare (problema specialmente per gli esperimenti più piccoli e meno organizzati dal punto di vista del calcolo)



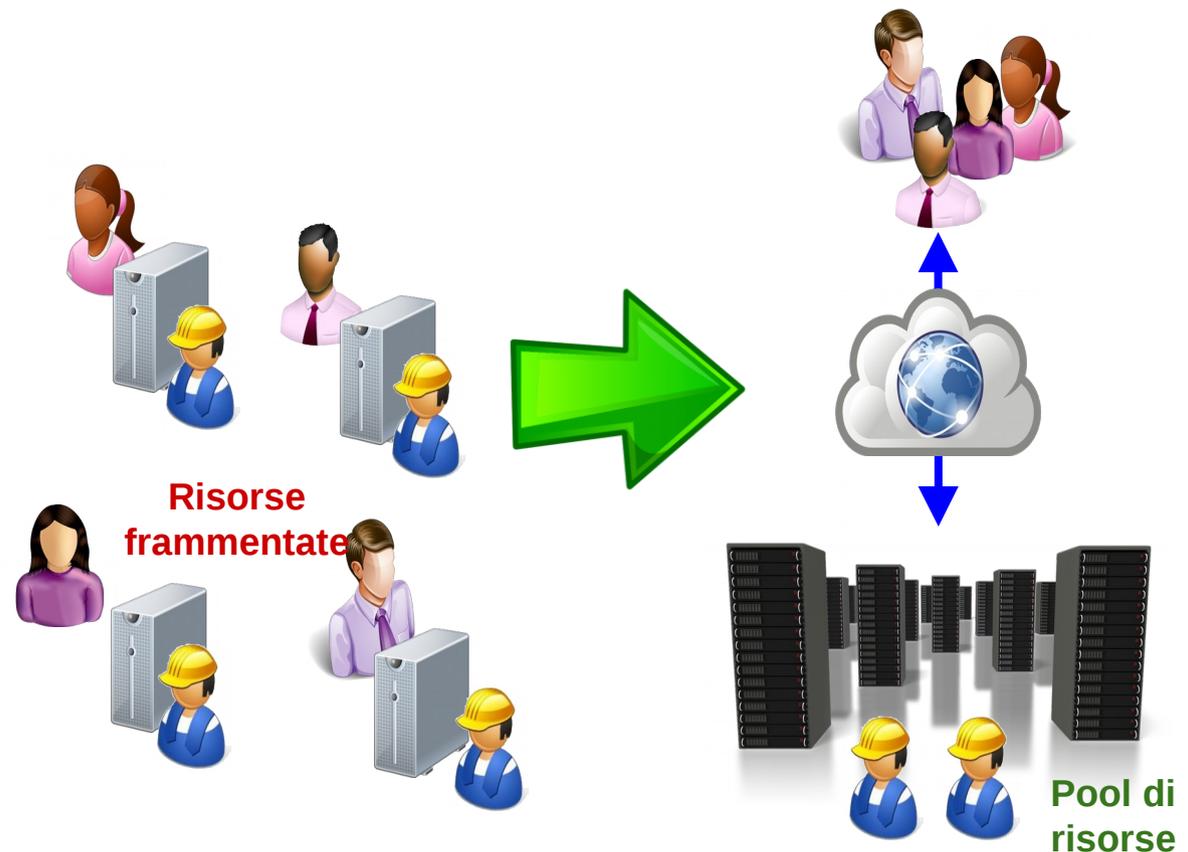
- Proliferazione di cluster di calcolo dei singoli gruppi
- Configurati per le specifiche esigenze di quel gruppo
 - **Insufficienti** in certi periodi
 - **Sotto-utilizzati** in altri
 - Molto costosi da gestire

Progetto proposto a fine 2013 per la **razionalizzazione dell'uso delle risorse**

Implementata una infrastruttura Cloud basata su **OpenStack**

In produzione da fine 2014

Unica infrastruttura di calcolo con **risorse distribuite** tra INFN Padova e INFN Legnaro (LNL)



Perché abbiamo scelto OpenStack ?

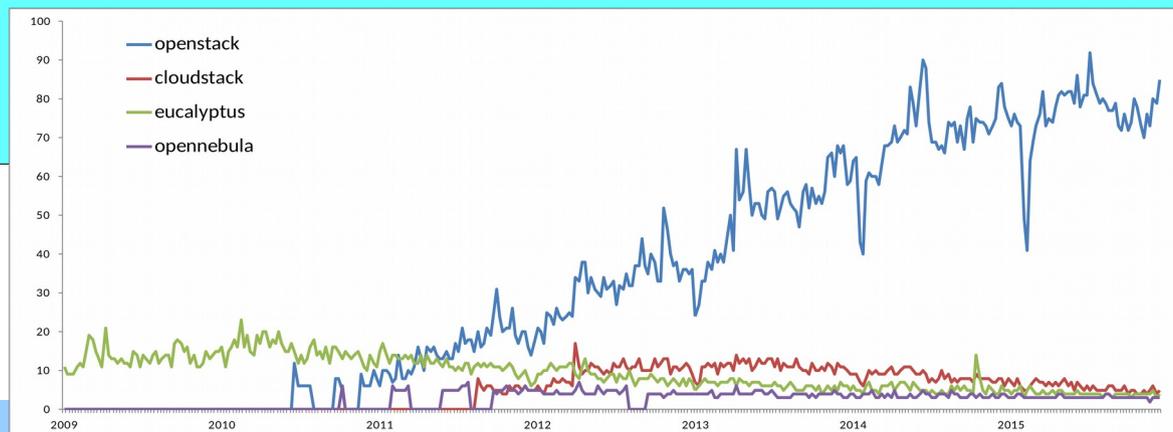
Perché è open-source e con dipendenze open-source

Per la sua architettura modulare (possibilità di utilizzare solo i moduli che interessano)

Per le sue prospettive di crescita (in termini di funzionalità, utenti, sviluppatori)

Per il forte supporto da parte dell'industria (ma non è in mano a nessun ente o impresa)

Perché i nostri partner (es. CERN) stavano convergendo verso OpenStack



Cloud dell'Università di Padova

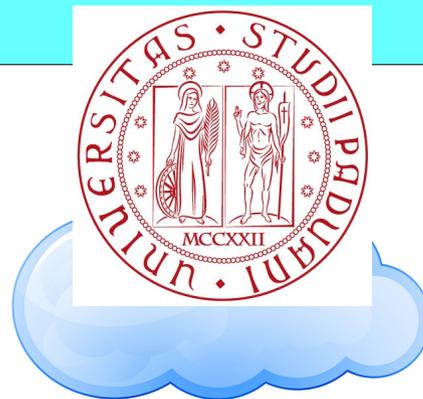


Circa un anno dopo (fine 2014) avviato un altro progetto Cloud dell'Università di Padova

Coinvolti 10 dipartimenti dell'Università di Padova, INFN Padova e INFN Legnaro

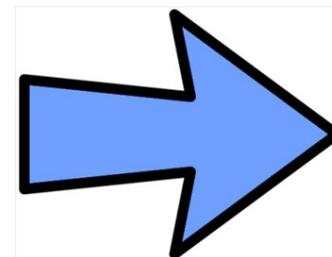
Con un finanziamento dell'Ateneo (più un co-finanziamento dei 10 dipartimenti coinvolti) implementata un'altra infrastruttura Cloud, sempre basata su OpenStack

Hardware di questa cloud ospitato c/o la sala CED del Dipartimento di Fisica e Astronomia (condivisa con INFN Padova)

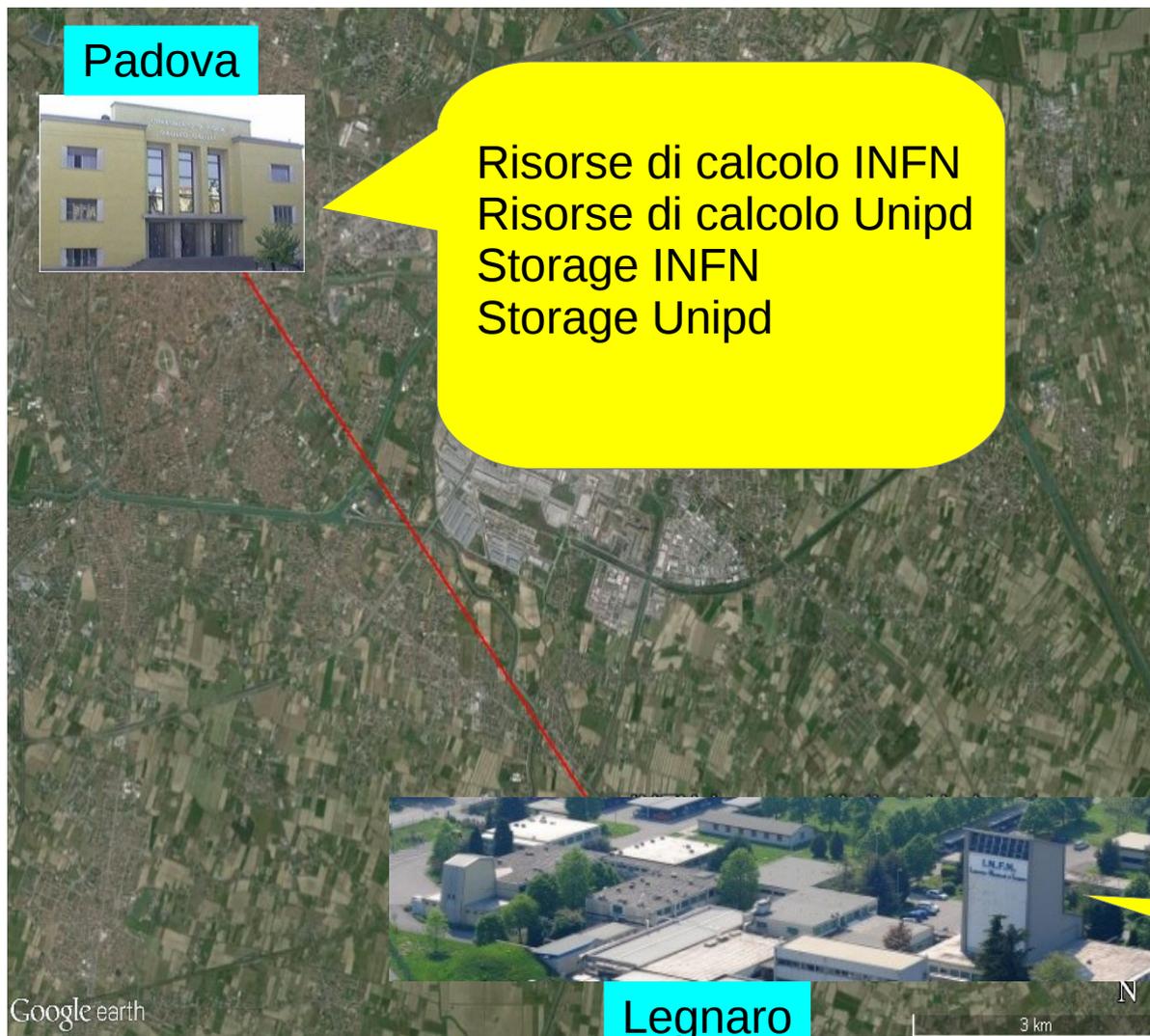




Effettuata
l'integrazione tra
Cloud Area Padovana
(INFN) e Cloud
Universitaria che si
sono unite in un'unica
infrastruttura chiamata
CloudVeneto



www.cloudveneto.it



Singola infrastruttura
Cloud con risorse
distribuite in 2 data
centre

Servizi configurati in
High Availability per
garantire la massima
disponibilità anche in
presenza di singoli
guasti

Risorse di calcolo INFN

~ 230 utenti
70+ progetti

~ 450 macchine
virtuali

~ 3500 VCPUs

~ 5 TB RAM

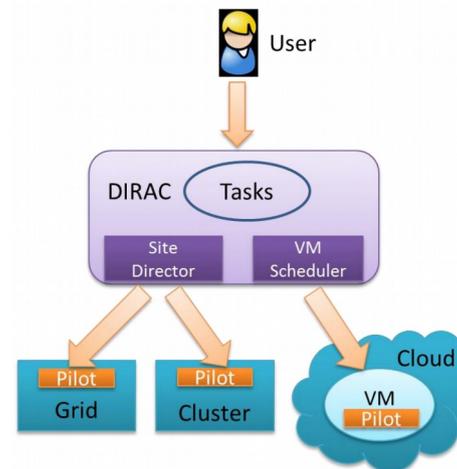
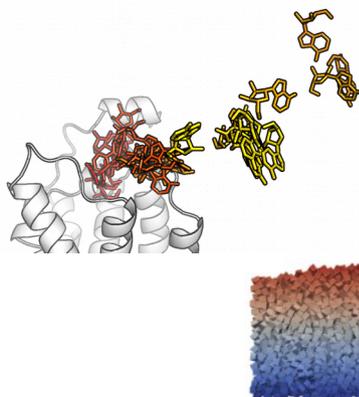
~ 150 TB disco

	Compute Nodes	Cores (in HT)	GPU	RAM (GB)	Storage for images and data (TB)
INFN-Padova	19	864	2	2912	240
INFN-Legnaro	13	416	0	1472	
Unipd	15	528	2	2080	90
Total	47	1808	4	6464	330

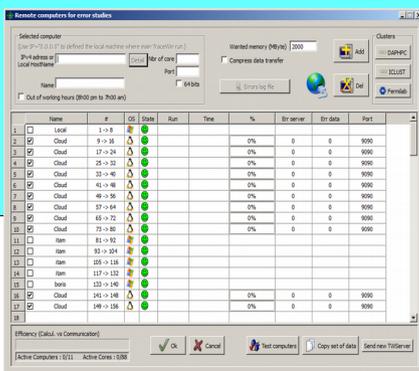
Utenti e risorse in **costante crescita**

Prevista a breve l'adesione di nuovi partner:
ARPAV, altri dipartimenti Unipd

- Simulazioni per il tuning di parametri dell'acceleratore
- Simulazioni di reti complesse
- Containerizzazione e architetture a MicroServizi
- Risoluzione numerica di integrali e equazioni differenziali
- Analisi dati HEP
- Big Data Analytics
- Simulazioni di terreni particellari
- Analisi di dati di RNAseq differenziali e mappatura su pathway
- Didattica
- ...



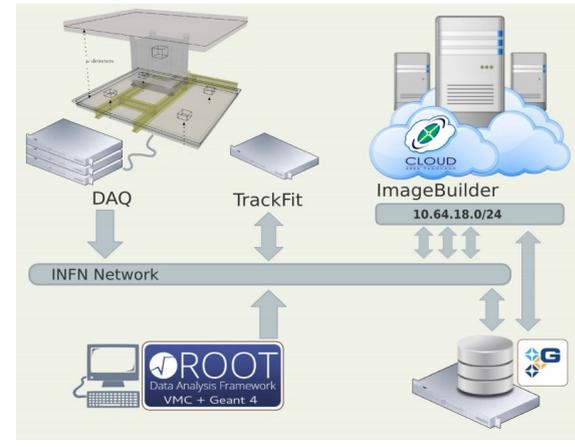
www.cloudveneto.it/testimonials.php



Name	#	OS	State	Run	Time	%	Err server	Err data	Port
1	Local	Linux	Running						9000
2	Cloud	Linux	Running			0%	0	0	9000
3	Cloud	Linux	Running			0%	0	0	9000
4	Cloud	Linux	Running			0%	0	0	9000
5	Cloud	Linux	Running			0%	0	0	9000
6	Cloud	Linux	Running			0%	0	0	9000
7	Cloud	Linux	Running			0%	0	0	9000
8	Cloud	Linux	Running			0%	0	0	9000
9	Cloud	Linux	Running			0%	0	0	9000
10	Cloud	Linux	Running			0%	0	0	9000
11	Cloud	Linux	Running			0%	0	0	9000
12	Cloud	Linux	Running			0%	0	0	9000
13	Cloud	Linux	Running			0%	0	0	9000
14	Cloud	Linux	Running			0%	0	0	9000
15	Cloud	Linux	Running			0%	0	0	9000
16	Cloud	Linux	Running			0%	0	0	9000
17	Cloud	Linux	Running			0%	0	0	9000
18	Cloud	Linux	Running			0%	0	0	9000
19	Cloud	Linux	Running			0%	0	0	9000
20	Cloud	Linux	Running			0%	0	0	9000

TraceWin Client with GUI on local desktop

- Remote Server
- ...



OpenStack

Creazione e gestione macchine virtuali

Block storage

Object Storage

Creazione e gestione immagini e snapshot

Dashboard grafica

Istanze private o accedibili in Internet

Orchestrazione

Integrate altri servizi e funzionalità di più alto livello (alcuni sono nostri sviluppi)

Cluster K8s as a service

Dynamic Batch Cluster as a service

Tool per deployment cluster
BigData

Autenticazione attraverso identity provider

Supporto per registrazione utenti e progetti

Servizi di accounting e monitoring



Foreman



kubernetes

Nagios®



PERCONA
Server for MySQL



HAPROXY
COMMUNITY EDITION



INDIGO - DataCloud
Better Software for Better Science



HTCondor
High Throughput Computing



mongoDB®

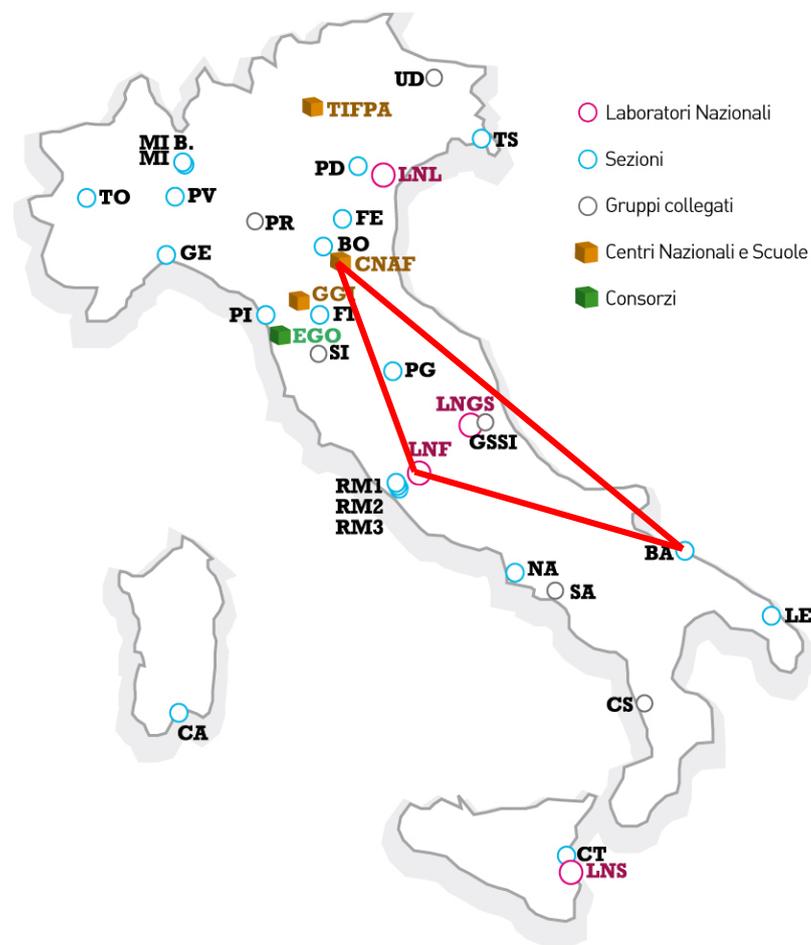
FOR GIANT IDEAS

Altre infrastrutture Cloud @ INFN

Molte altre sedi INFN si sono dotate di infrastrutture Cloud per razionalizzare l'uso delle risorse locali

Quasi tutte usano OpenStack
Unica eccezione è Torino dove si usa **OpenNebula**

INFN sta inoltre mettendo in produzione una infrastruttura Cloud distribuita multi-regione (CNAF Bologna, Bari, Laboratori Nazionali di Frascati) chiamata **INFN Corporate Cloud (INFN CC)**
Pensata in particolare per servizi e applicazioni geo-ridondati



Progetti Cloud

INFN ha partecipato e partecipa a molti progetti Cloud sia in ambito nazionale che internazionale



Sviluppo e sperimentare su scala nazionale di nuove soluzioni tecnologiche open, interoperabili e utilizzabili on-demand nell'ambito del Cloud Computing, insieme a nuovi modelli organizzativi sostenibili nel tempo per le Pubbliche Amministrazioni,



EU H2020 funded project aimed at developing scalable technologies for federating storage resources and managing data in highly distributed computing environments



Better Software for Better Science

Project funded under the Horizon2020 framework program of the European Unio. Development of a data and computing platform targeting scientific communities, deployable on multiple hardware and provisioned over hybrid (private or public) e-infrastructures.



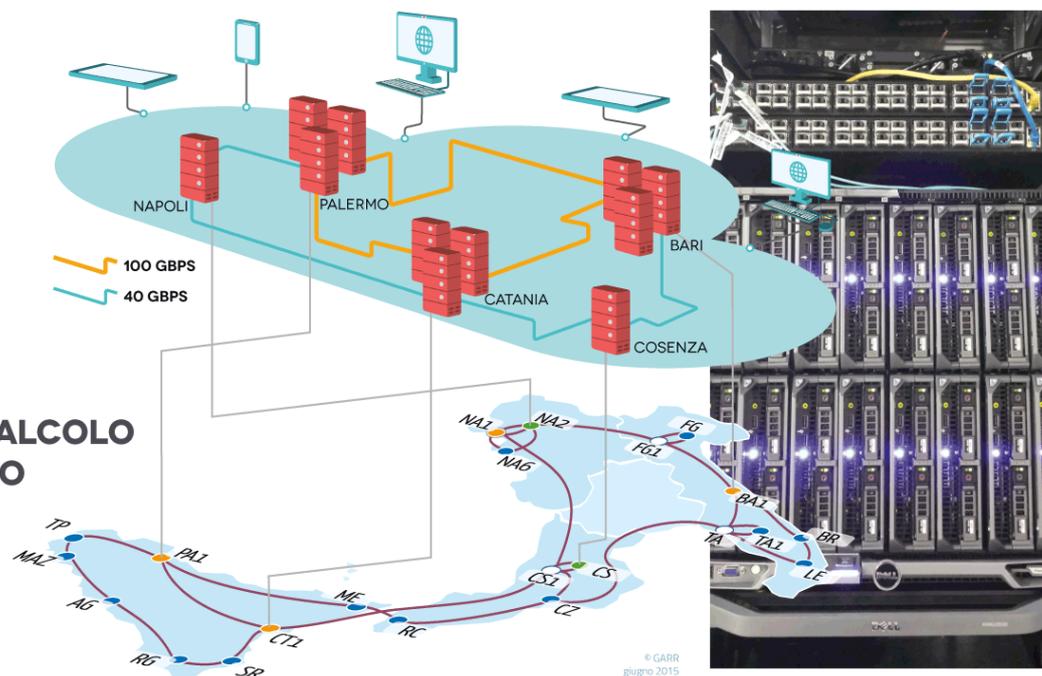
EOSC-hub brings together multiple service providers to create the Hub: a single contact point for European researchers and innovators to discover, access, use and reuse a broad spectrum of resources for advanced data-driven research.

Cloud nella ricerca in Italia

Anche molti altri enti di ricerca Italiani (CNR, etc.) si sono dotati di infrastrutture Cloud private

Anche il consorzio GARR (che si occupa di networking per Università e enti di Ricerca in Italia) offre Servizi Cloud per la comunità Italiana accademica e di ricerca

Basata su OpenStack+ Juju



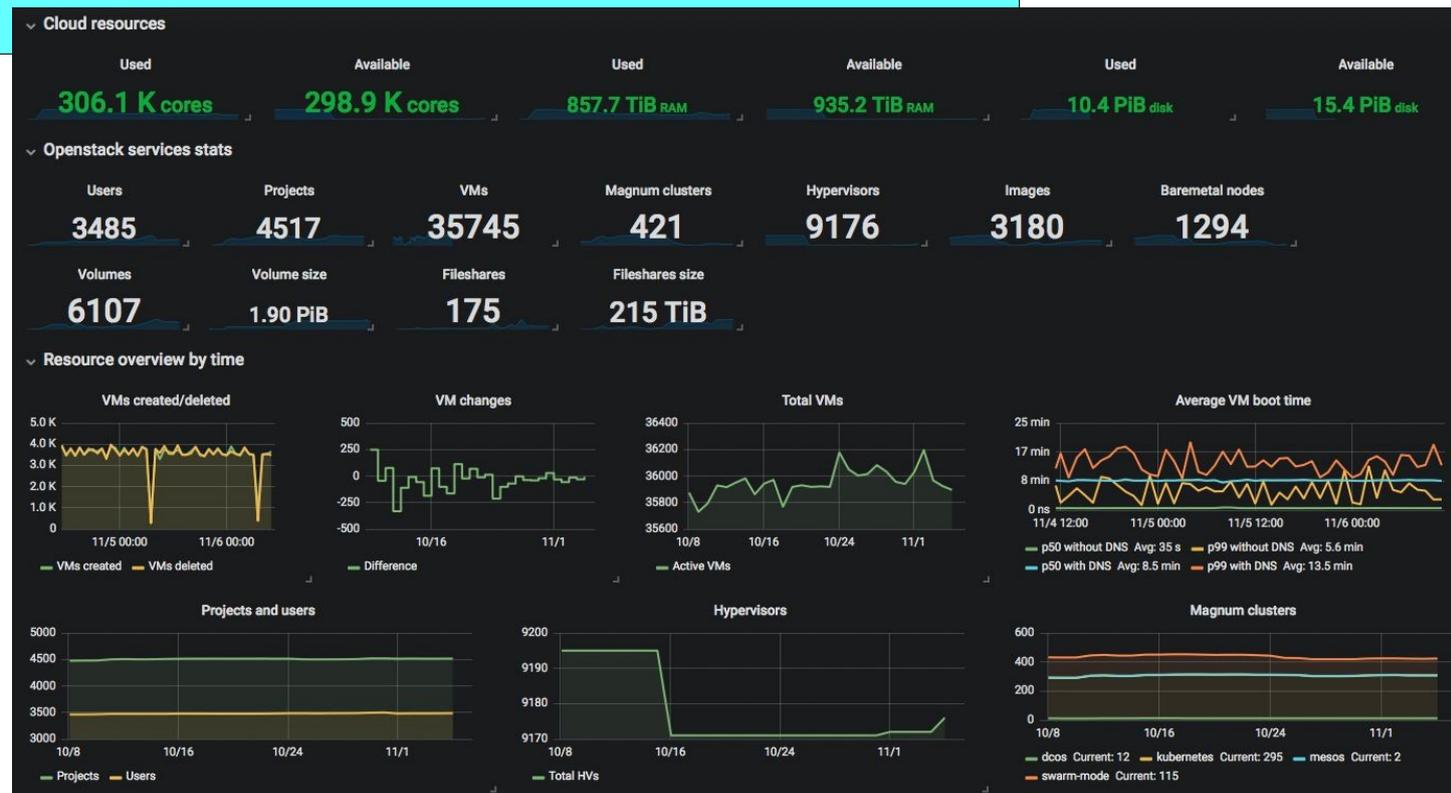
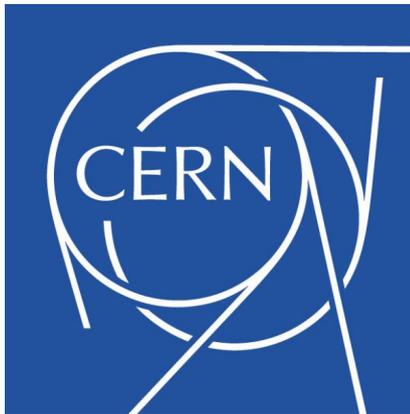
INFRASTRUTTURA DI CALCOLO E STORAGE DISTRIBUITO

- 📍 5 siti distribuiti
- 🖨️ 8.448 virtual CPU
- 💾 10 PB spazio storage

Cloud @ CERN

Cloud usata in molti altri istituti internazionali di fisica con cui collaboriamo

Es. il CERN opera dal 2013 una delle Cloud private OpenStack più grandi al mondo



L'infrastruttura CloudVeneto è un esempio di successo di collaborazione fruttuosa fra INFN e Dipartimenti dell'Università di Padova

È un servizio di produzione, ben supportato, usato con soddisfazione da parte degli utenti, in continua espansione

Ci permette un uso più efficiente delle risorse

Agevole modificare l'allocazione delle risorse ai diversi gruppi di ricerca in base alle esigenze

Il middleware Cloud OpenStack continua ad evolvere molto velocemente

Ora molto più maturo e stabile rispetto a qualche anno fa

È comunque un sistema estremamente complesso, che necessita di un certo sforzo per la sua gestione

THE END

Questions?



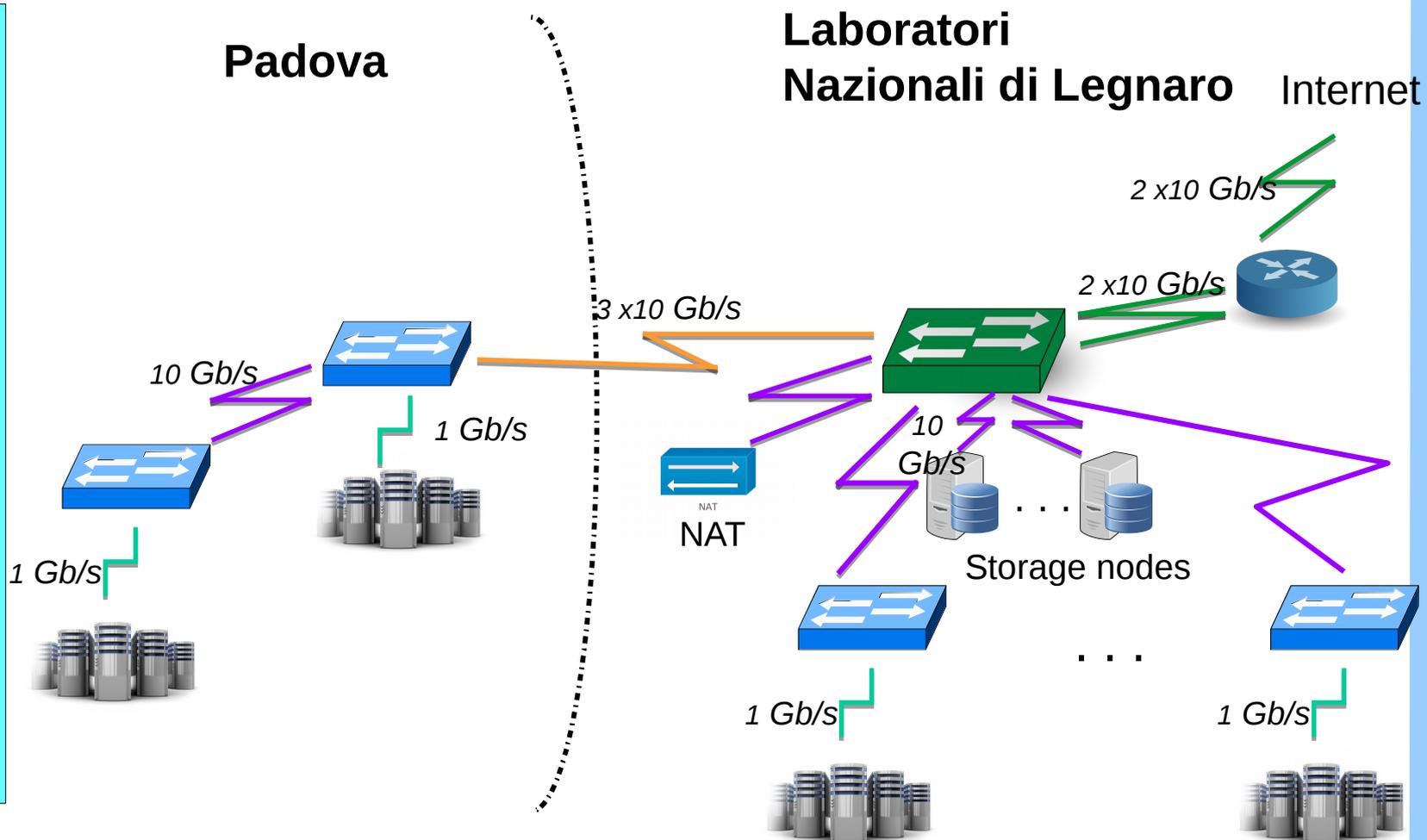
Backup slides

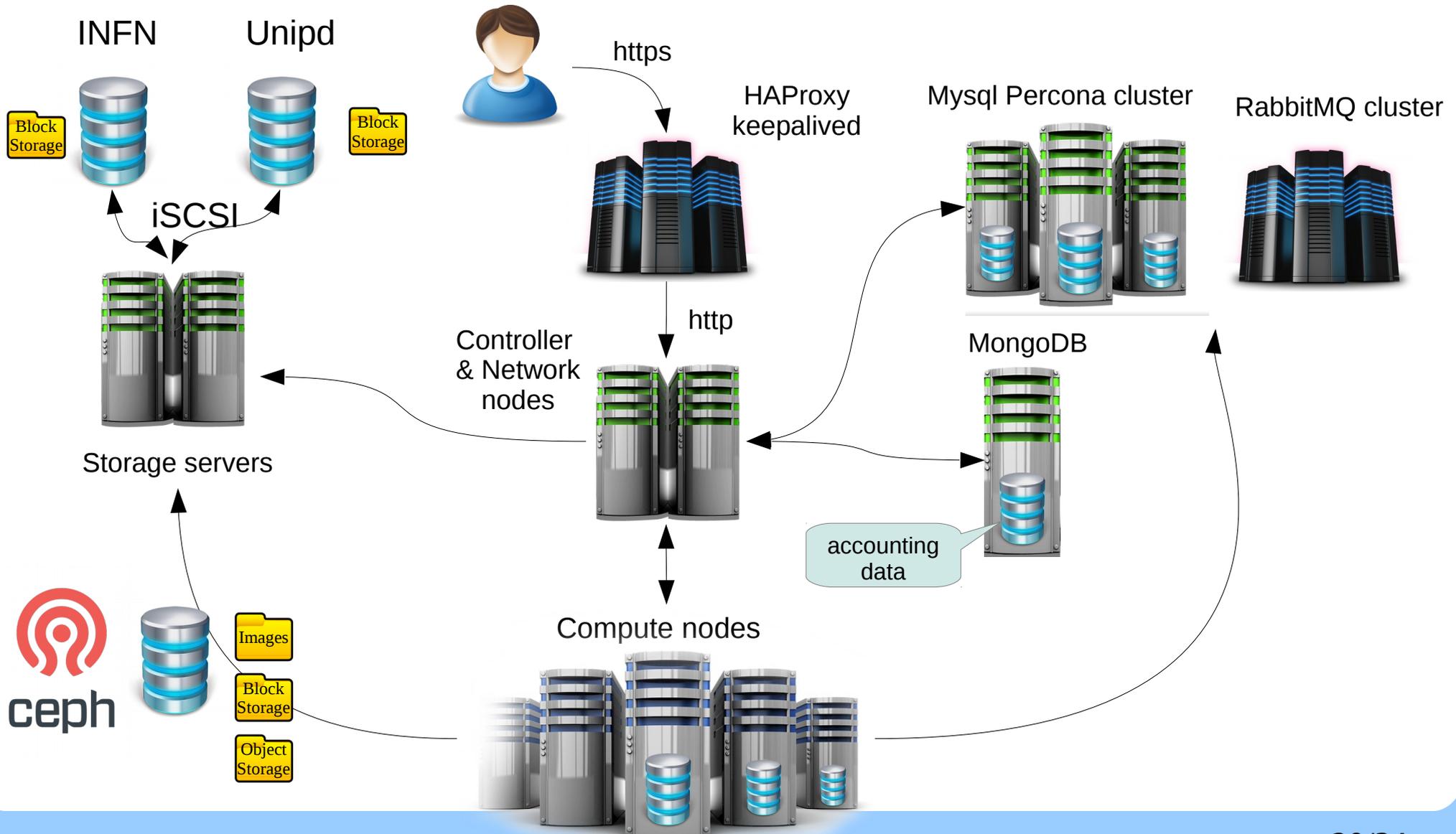
Tier-2 Padova Legnaro

Singolo centro di calcolo ma con risorse distribuite in 2 siti

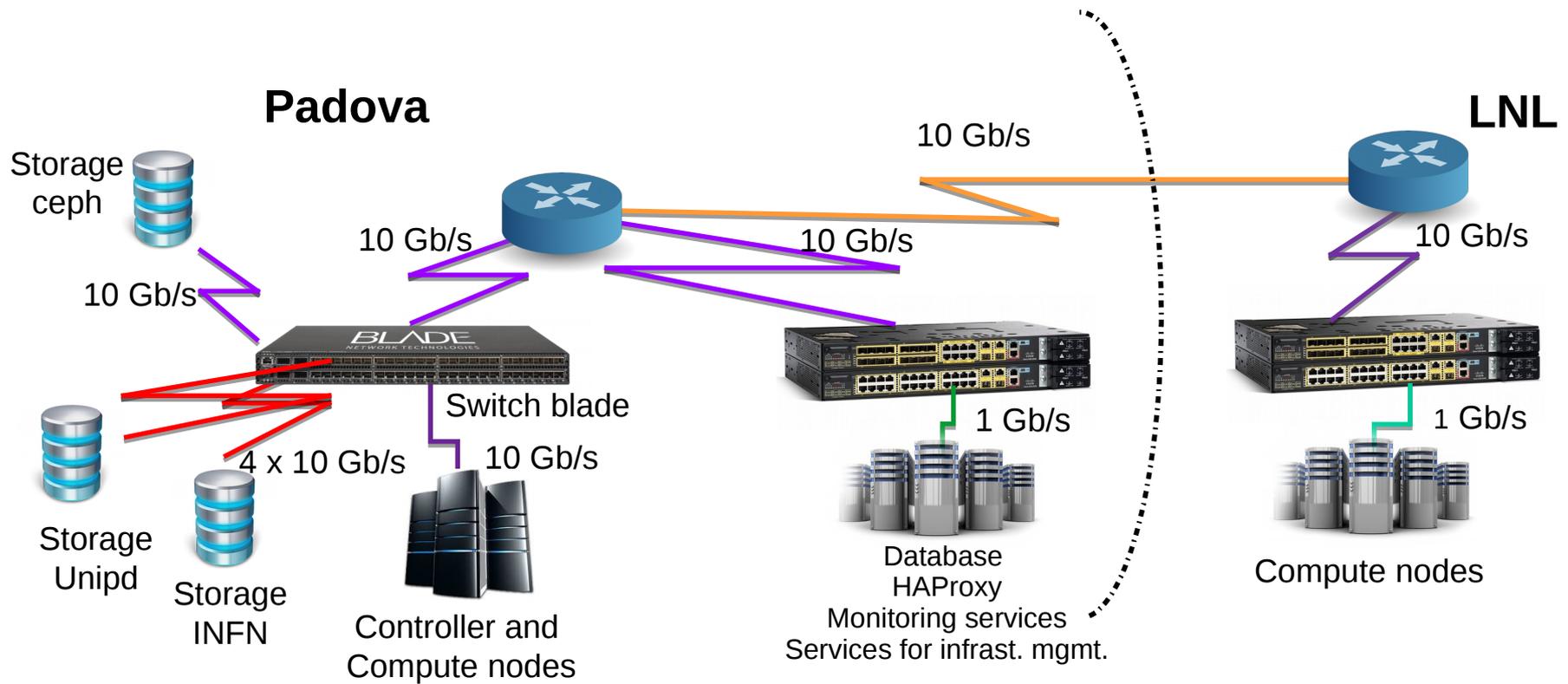
240 worker node (6388 core) organizzati in un singolo batch cluster (LSF)

~ 2.7 PB di dati





“High level” topology



Gestione delle informazioni di accounting

Gestione delle informazioni storiche di monitoring

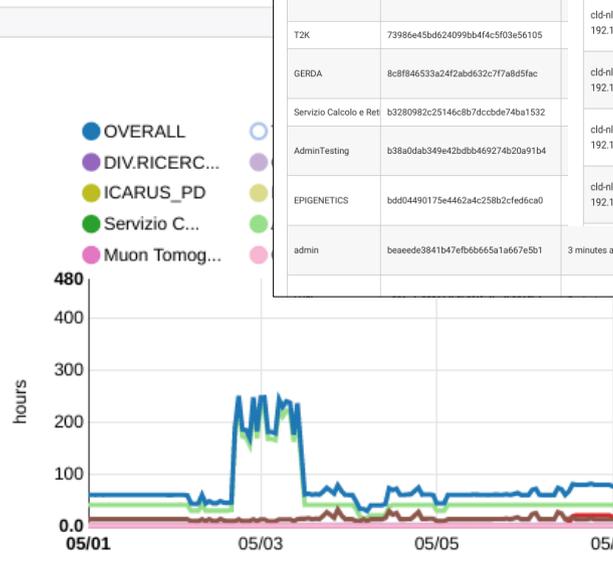
Active VMs	Allocated VRAM (TB)		Allocated VCPUs	
	Virtual	On bare	Virtual	On bare
259	Virtual: 3,15 82,34% Virtualizable: 5,05	Virtual: 3,15 80,96% Total Bare: 3,54	Virtual: 1926 84,43% Virtualizable: 4144	Virtual: 1926 176,66% Total Bare: 1072

Name	Last seen	State / Status	VMs	Allocated VRAM (MB)		Allocated VCPUs	
				Virtual	On bare	Virtual	On bare
				Virtual	On bare	Virtual	On bare
clid-ctrl-01.pd.inf.it (IP: 192.168.60.40)	4 minutes ago	down / disabled	0	Virtual: 512 1,53% Virtualizable: 48088.5	Virtual: 512 1,00% Total Bare: 32059	Virtual: 0 0,00% Virtualizable: 32	Virtual: 0 0,00% Total Bare: 8
clid-nl-01.cloud.pd.inf.it (IP: 192.168.60.56)	4 minutes ago	up / enabled	10	Virtual: 78848 54,25% Virtualizable: 144705	Virtual: 78848 17,65% Total Bare: 96470	Virtual: 59 1,00% Virtualizable: 128	Virtual: 59 184,08% Total Bare: 32
clid-nl-02.cloud.pd.inf.it (IP: 192.168.60.57)	4 minutes ago	up / enabled	11	Virtual: 113152 78,19% Virtualizable: 144705	Virtual: 113152 117,29% Total Bare: 96470	Virtual: 87 1,57% Virtualizable: 128	Virtual: 87 273,98% Total Bare: 32
clid-nl-03.cloud.pd.inf.it (IP: 192.168.60.58)	4 minutes ago	up / enabled	10	Virtual: 100864 69,76% Virtualizable: 144705	Virtual: 100864 104,55% Total Bare: 96470	Virtual: 81 1,50% Virtualizable: 128	Virtual: 81 253,13% Total Bare: 32
clid-nl-04.cloud.pd.inf.it (IP: 192.168.60.59)	4 minutes ago	up / enabled	10	Virtual: 107008 74,63% Virtualizable: 144703.5	Virtual: 107008 110,92% Total Bare: 96469	Virtual: 90 1,64% Virtualizable: 128	Virtual: 90 281,25% Total Bare: 32
clid-nl-05.cloud.pd.inf.it (IP: 192.168.60.60)	4 minutes ago	up / enabled	14	Virtual: 112640 77,84% Virtualizable: 144703.5	Virtual: 112640 116,76% Total Bare: 96469	Virtual: 69 1,27% Virtualizable: 128	Virtual: 69 215,63% Total Bare: 32
clid-nl-06.cloud.pd.inf.it (IP: 192.168.60.61)	4 minutes ago	up / enabled	14	Virtual: 82944 57,25% Virtualizable: 144703.5	Virtual: 82944 87,92% Total Bare: 96469	Virtual: 61 1,13% Virtualizable: 128	Virtual: 61 186,67% Total Bare: 32
clid-nl-07.cloud.pd.inf.it (IP: 192.168.60.62)	4 minutes ago	up / enabled	11	Virtual: 104960 72,64% Virtualizable: 192799.5	Virtual: 104960 111,64% Total Bare: 128533	Virtual: 83 1,53% Virtualizable: 128	Virtual: 83 259,38% Total Bare: 32
clid-nl-08.cloud.pd.inf.it (IP: 192.168.60.63)	4 minutes ago	up / enabled	16	Virtual: 121344 83,54% Virtualizable: 192799.5	Virtual: 121344 126,41% Total Bare: 128533	Virtual: 91 1,69% Virtualizable: 128	Virtual: 91 284,39% Total Bare: 32

Home Projects Hypervisors Accounting

Accounting

Metric: CPU Time



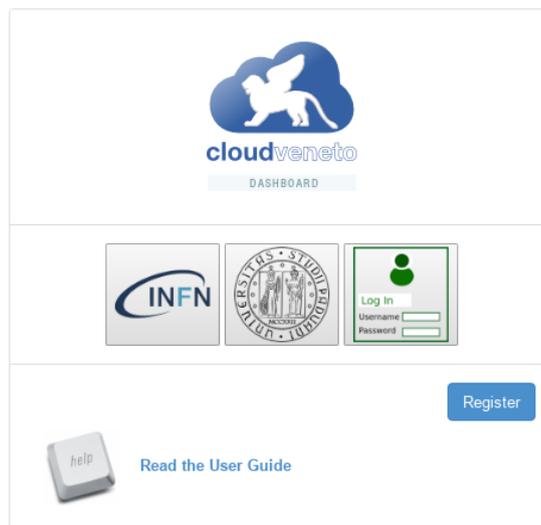
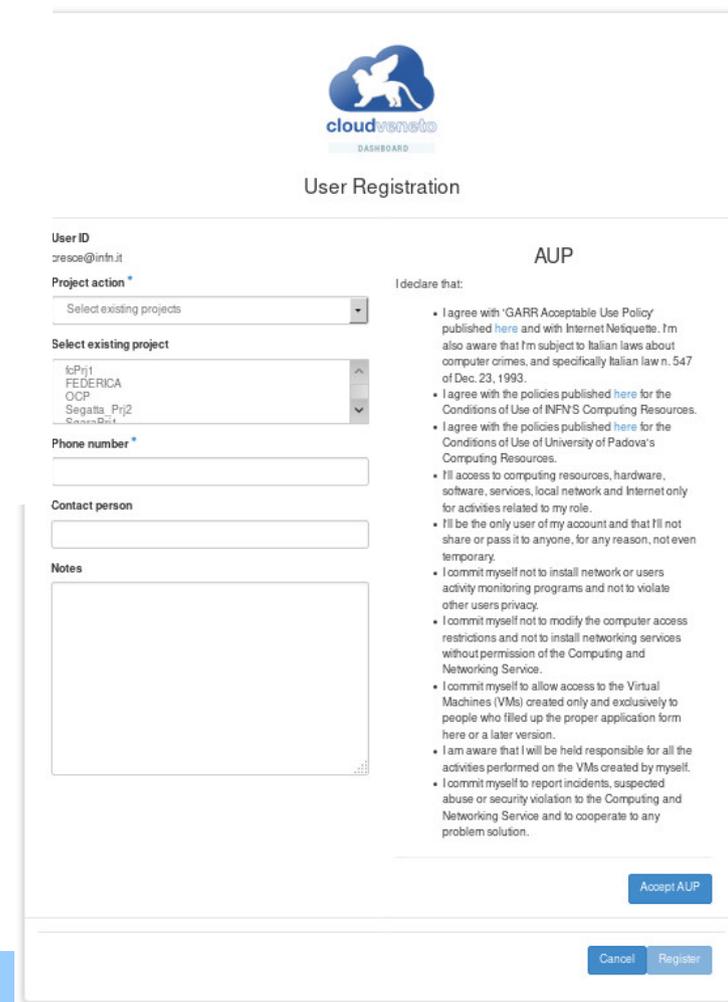
Project	Usage	Quota
Muon tomography	16.337,94	3,20%
EPIGENETICS	8.511,42	1,67%
OCF	8.032,67	1,57%

Mon May 08 07:00 2017 +0200
SPES 40.057 hours

Supporto per autenticazione attraverso Identity Provider

Procedura per la registrazione di nuovi utenti e progetti integrata nella dashboard

Ogni progetto Cloud ha un amministratore che gestisce (via dashboard) le richieste di affiliazione al suo progetto

User Registration

User ID
cresce@infn.it

Project action *
Select existing projects

Select existing project
fcPrj1
FEDERICA
OCP
Segatta_Prj2
Sgaravati

Phone number *
[input field]

Contact person
[input field]

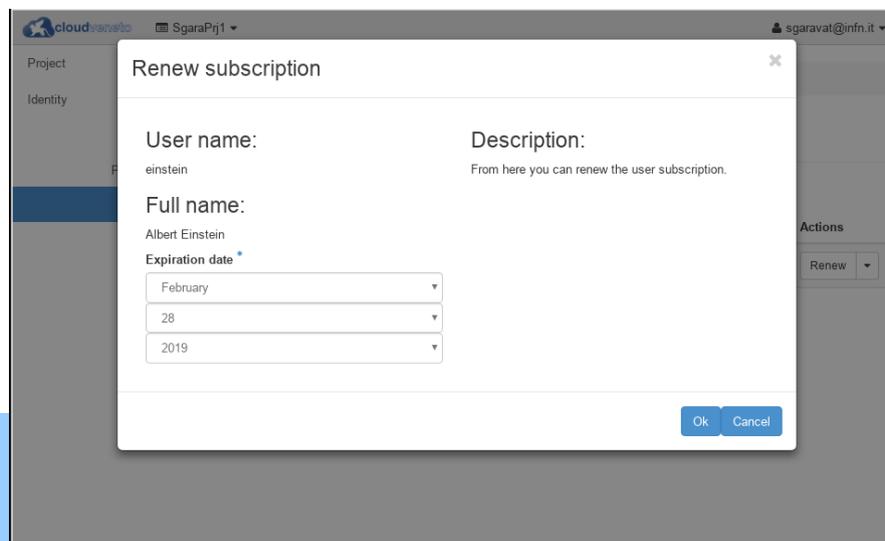
Notes
[text area]

I declare that:

- I agree with 'GARR Acceptable Use Policy' published [here](#) and with Internet Netiquette. I'm also aware that I'm subject to Italian laws about computer crimes, and specifically Italian law n. 547 of Dec. 23, 1993.
- I agree with the policies published [here](#) for the Conditions of Use of INFN'S Computing Resources.
- I agree with the policies published [here](#) for the Conditions of Use of University of Padova's Computing Resources.
- I'll access to computing resources, hardware, software, services, local network and Internet only for activities related to my role.
- I'll be the only user of my account and that I'll not share or pass it to anyone, for any reason, not even temporary.
- I commit myself not to install network or users activity monitoring programs and not to violate other users privacy.
- I commit myself not to modify the computer access restrictions and not to install networking services without permission of the Computing and Networking Service.
- I commit myself to allow access to the Virtual Machines (VMs) created only and exclusively to people who filled up the proper application form here or a later version.
- I am aware that I will be held responsible for all the activities performed on the VMs created by myself.
- I commit myself to report incidents, suspected abuse or security violation to the Computing and Networking Service and to cooperate to any problem solution.

Accept AUP

Cancel Register



cloudveneto SgaraPrj1 sgaravat@infn.it

Project
Identity

Renew subscription

User name:
einstein

Description:
From here you can renew the user subscription.

Full name:
Albert Einstein

Expiration date *
February
28
2019

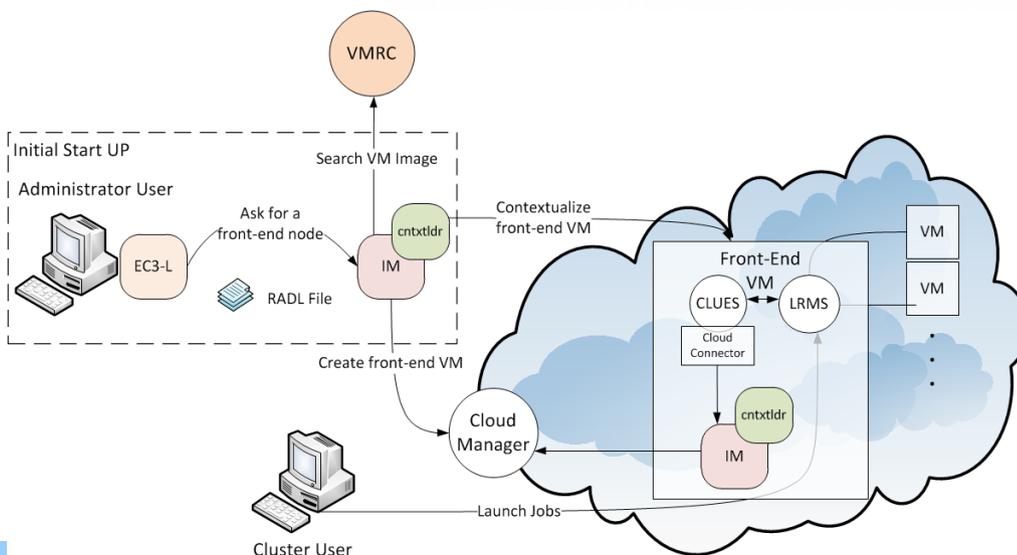
Actions
Renew

Ok Cancel

Gli utenti hanno la possibilità di creare cluster batch elastici

Nuovi slave vengono automaticamente istanziati se ci sono job in coda; vengono automaticamente distrutti quando non servono più

Supporto per HTCondor e SLURM



Run 2 Data Flow

RAW:
'raw' detector
Information (detector
element hits, etc)

AOD:
Analysis Object Data
'Distilled' version
of the RECO
data

Mini-AOD:
Used for physics
analysis

**RAW
and
RECO**

- Twice per year
- From prompt reco
- Online-reco

3B-12B events
AOD
480kB/ev

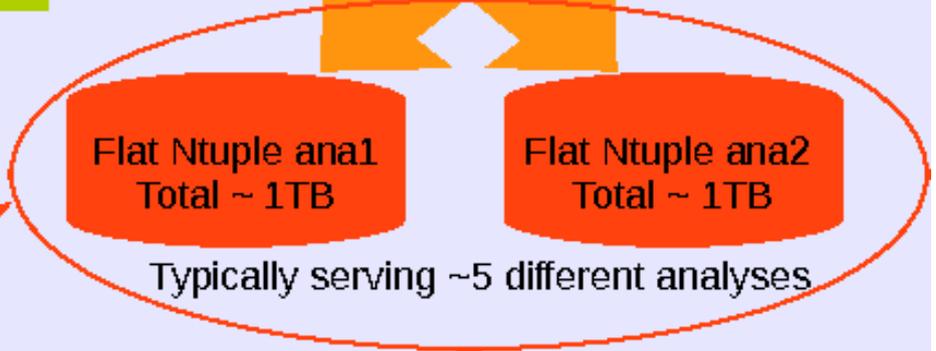
- Two days run time
- Monthly
- No skimming

Mini-AOD
35-60kB/ev

Takes 24h to run

Official, centrally produced

RECO:
RAW data digitized
and collision events
reconstructed



Private formats