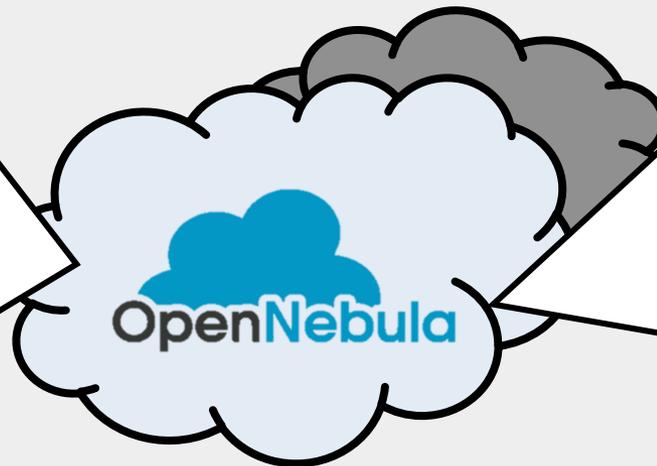


# CLOUD COMPUTING ACTIVITIES IN TORINO

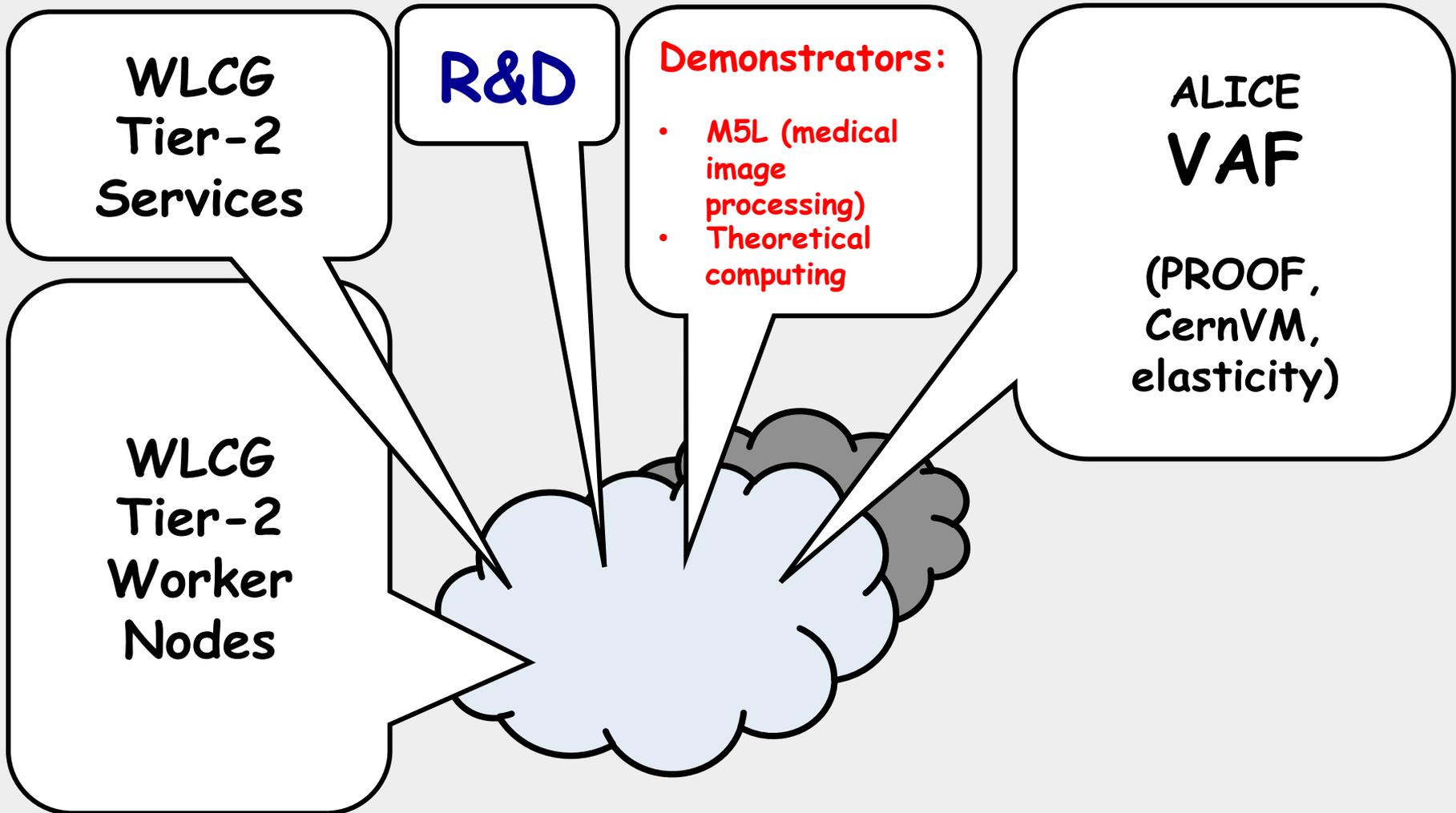


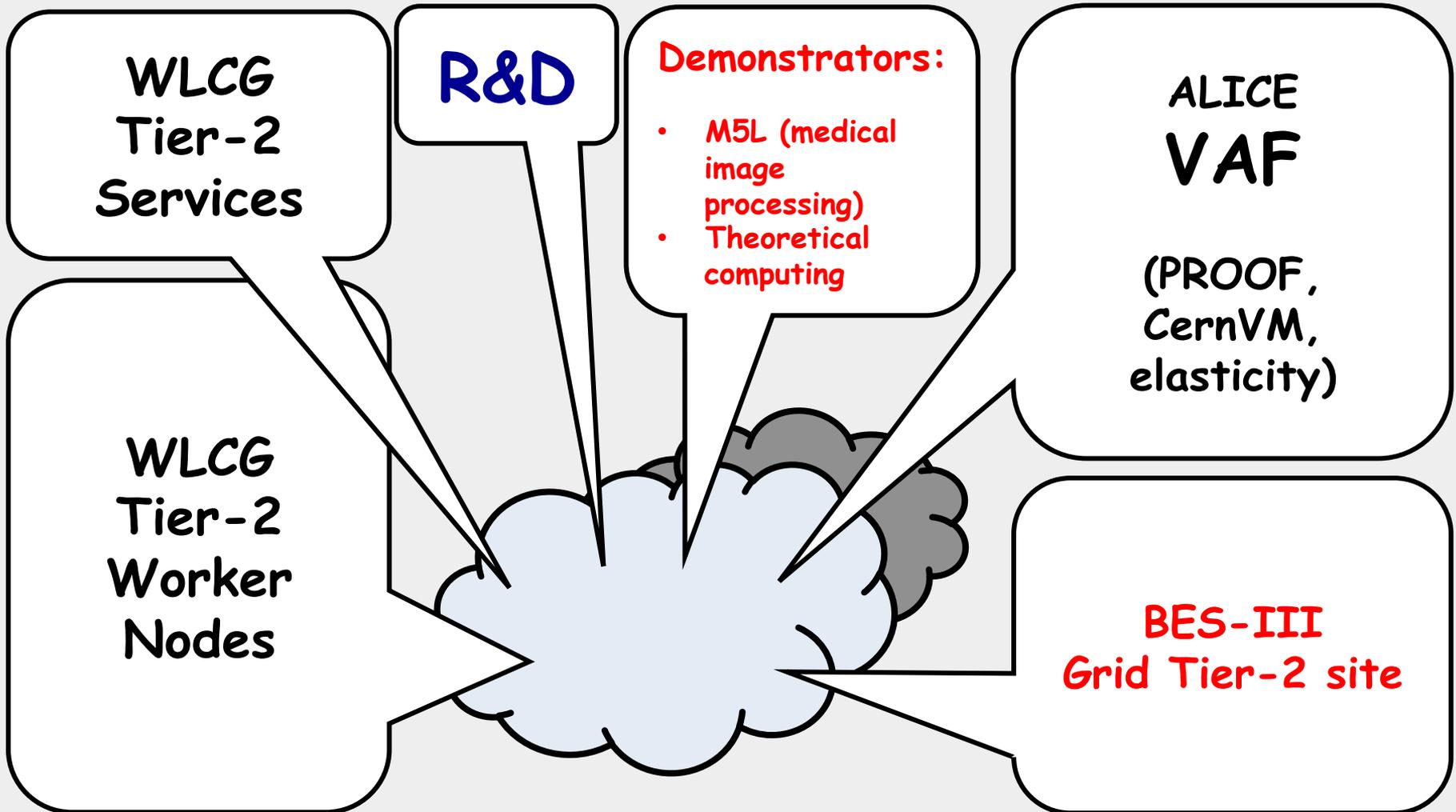
# IN THE BEGINNING...

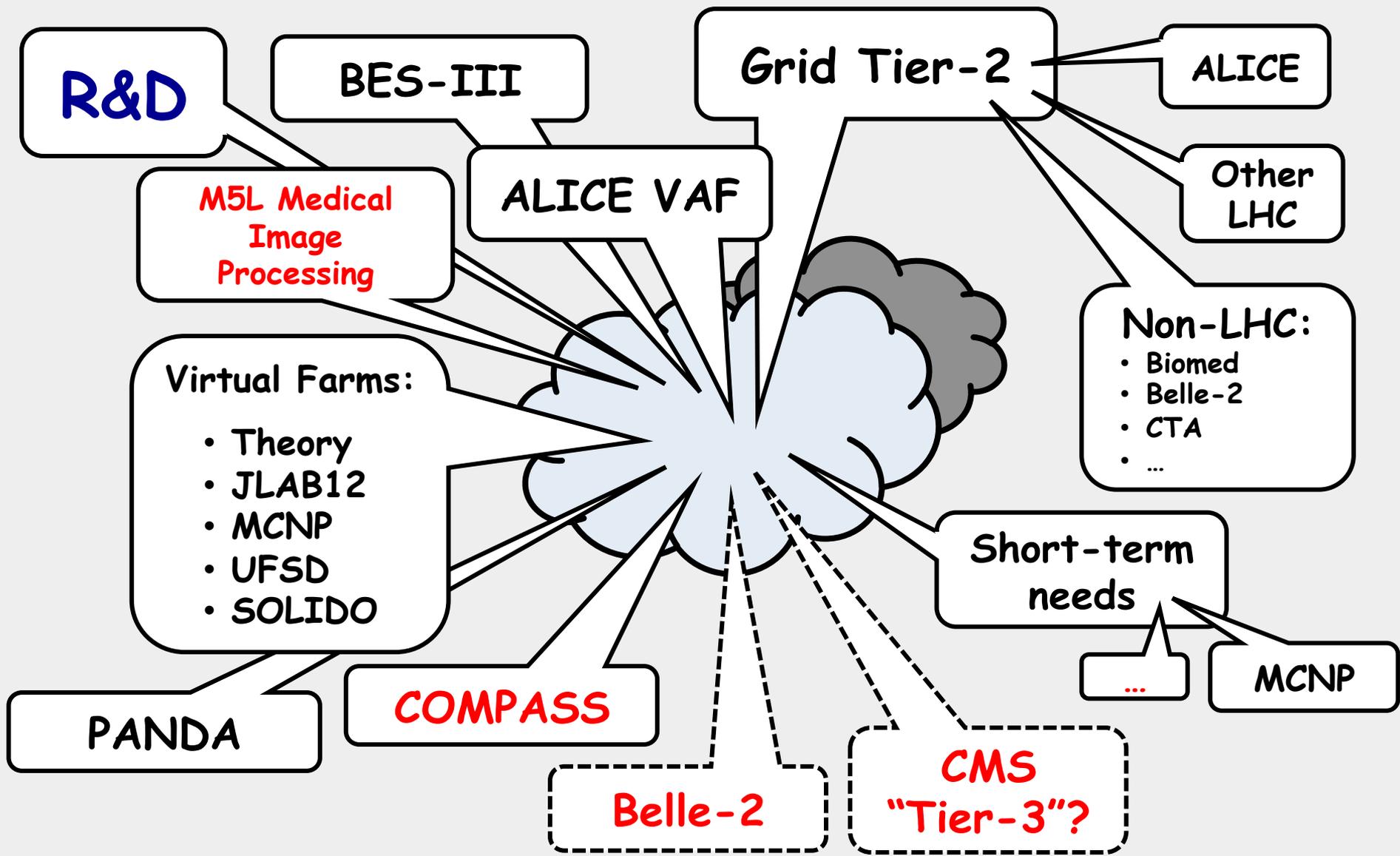
WLCG  
Tier-2  
Worker  
Nodes



ALICE  
Analysis  
Facility  
(PROOF)







# A TALE OF TWO CLUSTERS



VMs providing **critical services**:

- in- & out-bound connectivity
- public & private IP
- live migration
- no special I/O requirements



VMs providing **computing workforce**:

- example: Grid WNs
- private IP only
- high storage I/O performance

# A TALE OF TWO CLUSTERS

- Server-class hardware
- Shared image repository
- Resiliency-optimized FS for shared system disks
- Live migration possible

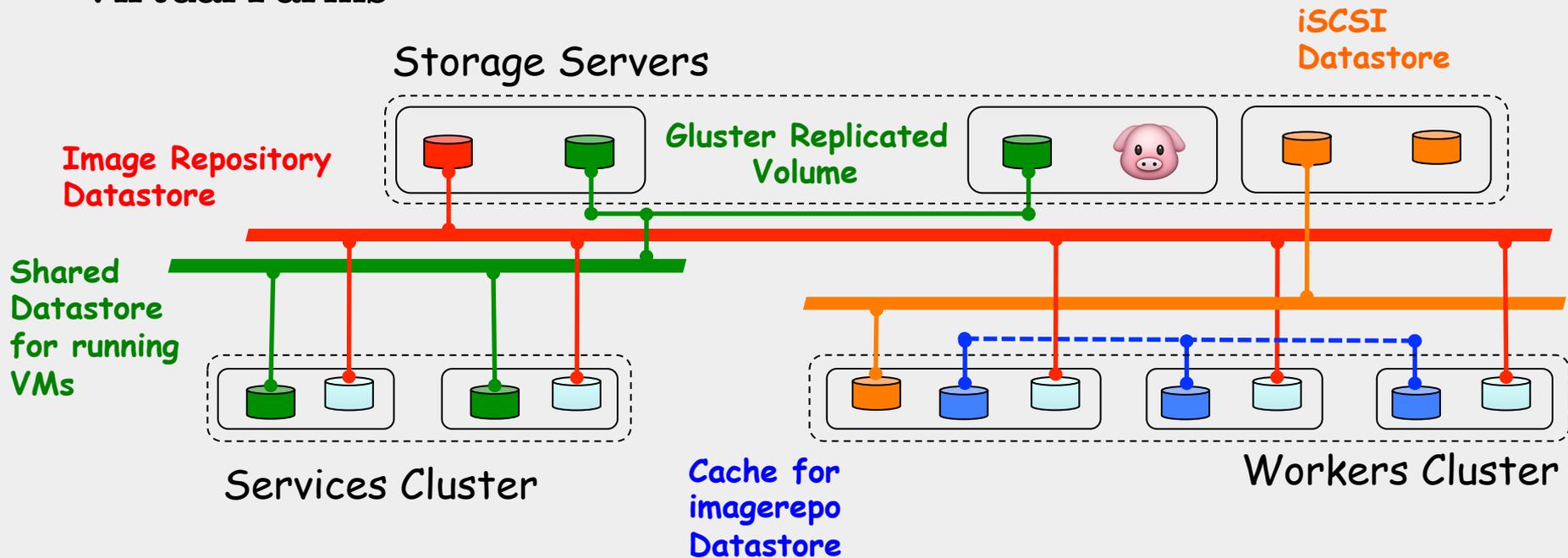


- Working-class hardware 😊
- Cached image repository
- Access to performance-optimized FS for data needs
- Runs about 200 VMs



# MULTIPURPOSE STORAGE: GLUSTERFS

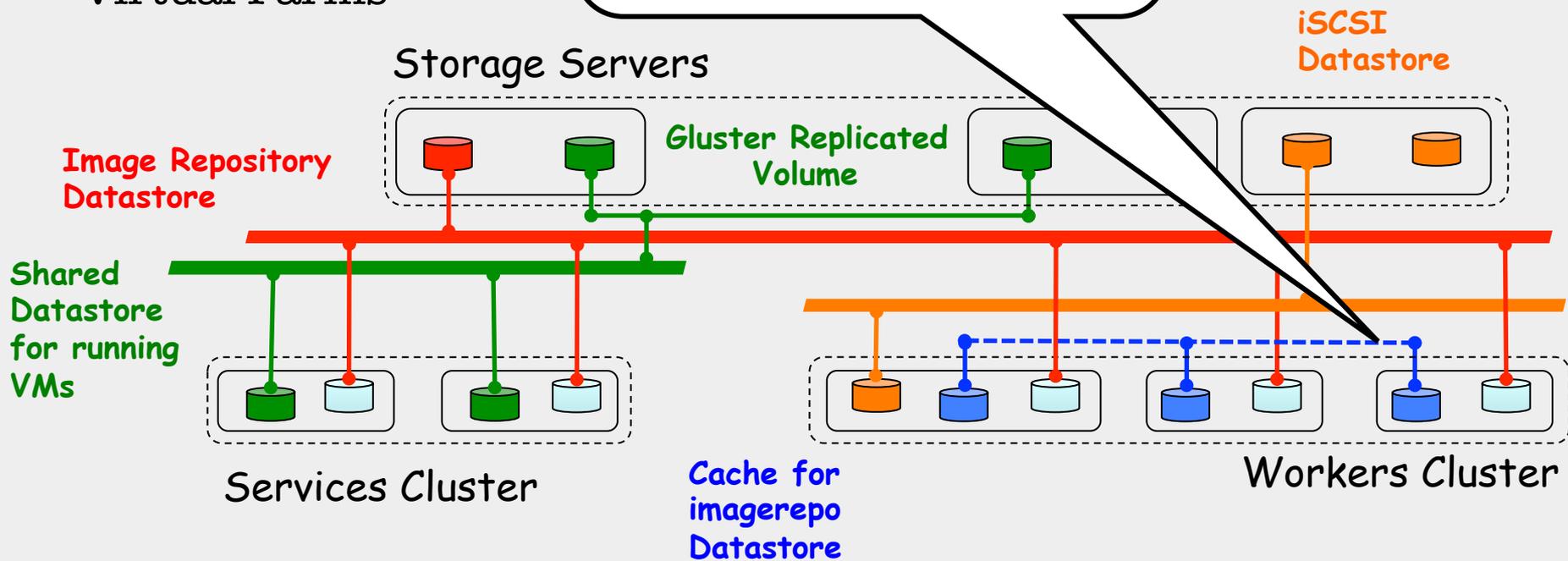
- **Services System Datastore** is **shared** to allow live migration
- **Workers System Datastore** is **local** to the hypervisors to increase I/O capacity. Images repository is locally **cached** on each hypervisor to reduce startup time
- **Persistent Space Datastore** is mounted on the relevant hypervisors using the **iSCSI** Transfer Manager to provide persistent storage to Virtual Farms



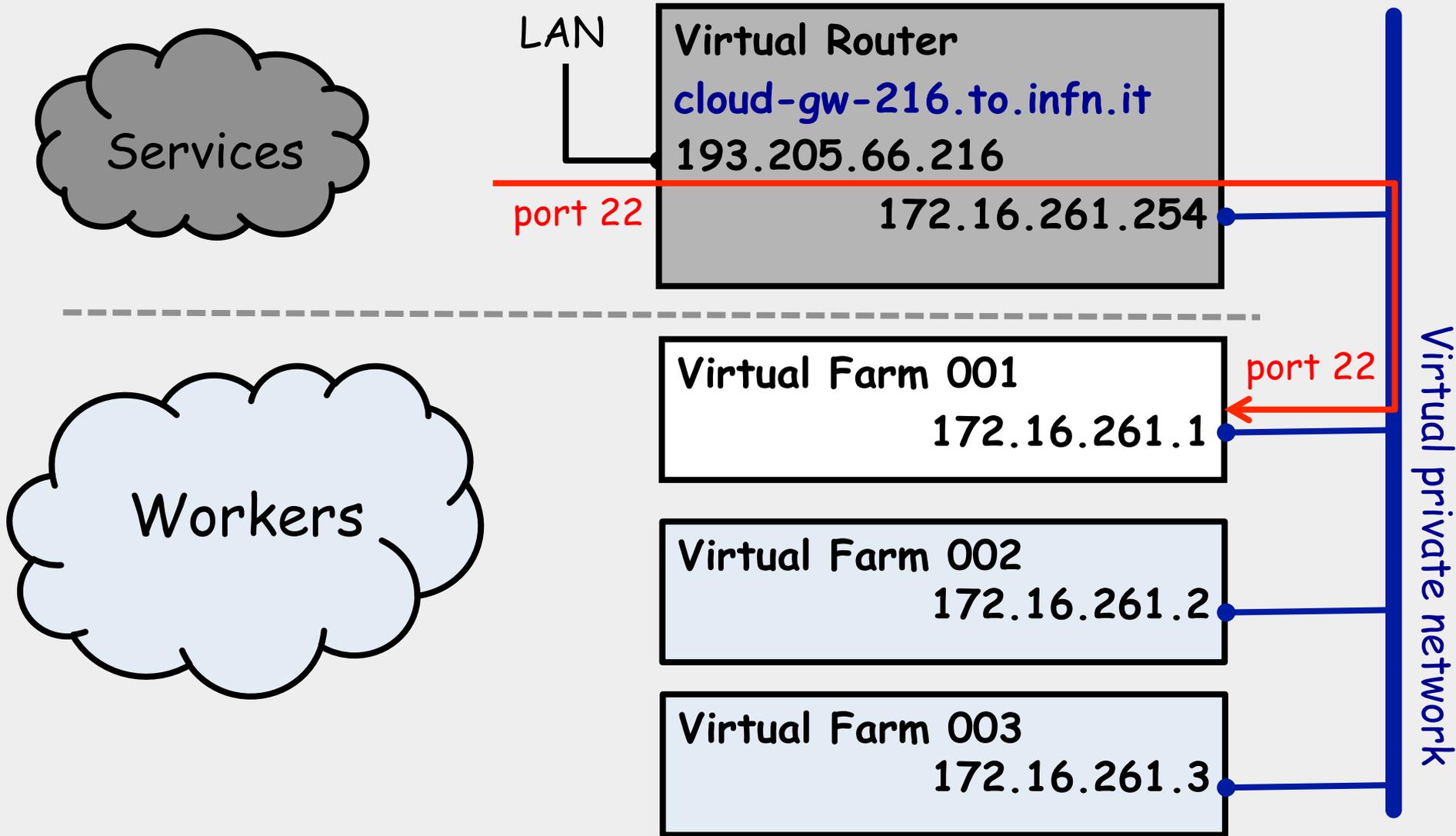
# MULTIPURPOSE STORAGE: GLUSTERFS

- **Services System Datastore** is **shared** to allow live migration
- **Workers System Datastore** is **local** to the hypervisors to increase I/O capacity. Images replicate to reduce startup time
- **Persistent Space Datastore** using the **iSCSI** Transfer Virtual Farms

An ad-hoc script synchronizes the local copies using a custom "torrent-like" tool (scpWave + rsync) when new versions of the images are saved

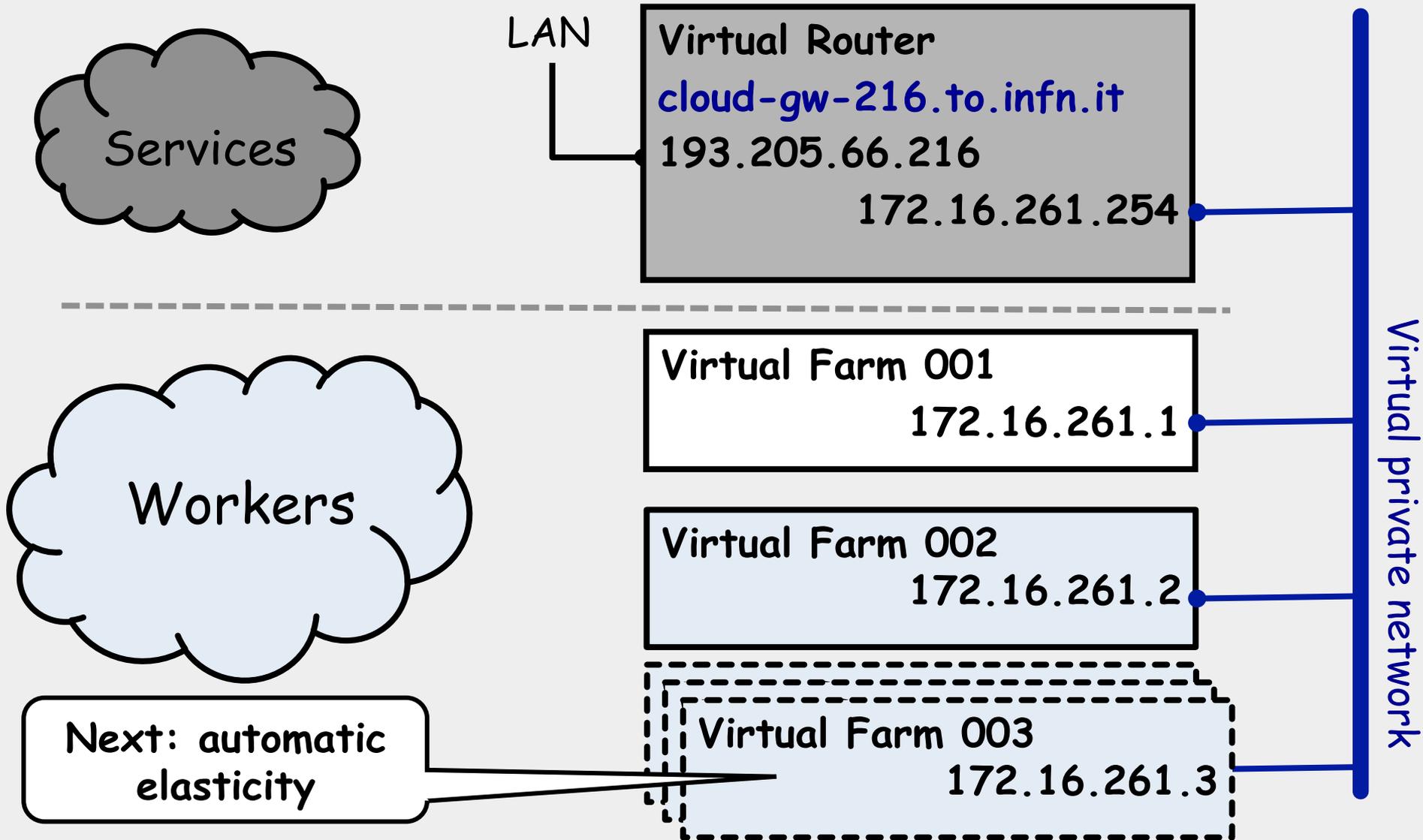


# BASIC VIRTUAL FARM PROVISIONING MODEL



- ONeUser & quotas (CPU, memory,...)
- VRouter (OpenWRT small VM)
- OS Images (CentOS, Ubuntu,...)
- Private network (e.g. 172.16.216.0/24)
  - With DHCP, DNS, NAT,...
- “Elastic” public IP (e.g. 193.205.66.216)
- Volatile & persistent storage
  - iSCSI datastore
- Context file templates (Cloud-config)
  - E.g. HTCondor head node & workers
- EC-2 CLI access and Sunstone dashboard

# VIRTUAL FARM PROVISIONING MODEL



# TENANT DASHBOARD



# SUNSTONE DASHBOARD

Recent - Google Drive | I ricercatori alla prova del | History | OpenNebula Sunstone: Clc x

https://one-master.to.infn.it

OpenNebula Sunstone | Documentation | Support | Community | Welcome oneadmin | Sign out

- Dashboard
- System
- Virtual Resources
  - Virtual Machines**
  - Templates
  - Images
- Infrastructure
  - Clusters
  - Hosts
  - Datastores
  - Virtual Networks
- Marketplace

<input type="checkbox"/>	6478	oneadmin	oneadmin	compass-VRouter	RUNNING	one-kvm-srv-05	172.16.219.254 193.205.66.219	
<input type="checkbox"/>	6482	oneadmin	oneadmin	SE-EC2	RUNNING	one-kvm-srv-02	192.168.0.231 193.205.66.192	
<input type="checkbox"/>	6512	oneadmin	oneadmin	WN-EC2-6slots	RUNNING	one-kvm-22	192.168.3.14	
<input checked="" type="checkbox"/>	6513	oneadmin	oneadmin	BDII-EMI3-CentOS6-v1	RUNNING	one-kvm-srv-05	192.168.0.252 193.206.184.19	
<input type="checkbox"/>	6522	oneadmin	oneadmin	WN-EC2-6slots	RUNNING	one-kvm-63	192.168.3.15	
<input type="checkbox"/>	6523	oneadmin	oneadmin	giunti-VRouter	RUNNING	one-kvm-srv-01	172.16.216.254 193.205.66.216	
<input type="checkbox"/>	6637	aguarise	INFN-TO	faust-proxy	RUNNING	one-kvm-33	192.168.5.91	
<input type="checkbox"/>	6640	prooftaf	ec2	ec2-m1-large	RUNNING	one-kvm-33	192.168.6.195	
<input type="checkbox"/>	6643	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-61	192.168.3.16	
<input type="checkbox"/>	6644	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-33	192.168.3.17	
<input type="checkbox"/>	6645	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-25	192.168.3.18	
<input type="checkbox"/>	6646	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-22	192.168.3.19	
<input type="checkbox"/>	6667	prooftaf	ec2	ec2-m1-large	RUNNING	one-kvm-66	192.168.6.207	
<input type="checkbox"/>	6687	bes	bes	WN-BESIII	RUNNING	one-kvm-38	192.168.4.40	
<input type="checkbox"/>	6690	bes	bes	WN-BESIII	RUNNING	one-kvm-15	192.168.4.41	
<input type="checkbox"/>	6691	bes	bes	WN-BESIII	RUNNING	one-kvm-13	192.168.4.42	
<input type="checkbox"/>	6720	oneadmin	oneadmin	amq-broker	RUNNING	one-kvm-srv-02	193.206.184.33 192.168.5.83	
<input type="checkbox"/>	6722	svallero	ec2	ec2-m1-small	RUNNING	one-kvm-40	172.16.213.79	
<input type="checkbox"/>	6726	aguarise	INFN-TO	amq-test	RUNNING	one-kvm-19	192.168.5.86	
<input type="checkbox"/>	6741	svallero	ec2	ec2-m1-small	RUNNING	one-kvm-36	172.16.213.81	
<input type="checkbox"/>	6788	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-01	172.16.216.5	
<input type="checkbox"/>	6789	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-18	172.16.216.6	
<input type="checkbox"/>	6790	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-19	172.16.216.7	
<input type="checkbox"/>	6791	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-20	172.16.216.8	
<input type="checkbox"/>	6798	prooftaf	ec2	ec2-m1-large	RUNNING	one-kvm-12	192.168.6.210	

Showing 51 to 100 of 106 entries

First Previous 1 2 3 Next Last

OpenNebula 3.8.4 by C12G Labs

# SUNSTONE DASHBOARD

OpenNebula Sunstone: Clc x

https://one-master.to.infn.it

OpenNebula Sunstone Virtual Machines

oneadmin OpenNebula

Dashboard System Virtual Resources **Virtual Machines** Templates Images Files & Kernels Infrastructure Marketplace OneFlow Support

<input type="checkbox"/>	ID	Owner	Group	Name	Status	Host	IPs	VNC
<input type="checkbox"/>	10914	prooftaf	ec2	ec2-m1-large	RUNNING	one-kvm-66	192.168.6.195	
<input type="checkbox"/>	10907	prooftaf	ec2	ec2-m1-large	RUNNING	one-kvm-65	192.168.6.192	
<input type="checkbox"/>	10906	testuser	ec2	ec2-m1-medium	RUNNING	one-kvm-67	172.16.225.3	
<input type="checkbox"/>	10886	jlab12	ec2	ec2-m1-small	RUNNING	one-kvm-22	172.16.220.5	
<input type="checkbox"/>	10885	jlab12	ec2	ec2-m1-medium	RUNNING	one-kvm-22	172.16.220.4	
<input type="checkbox"/>	10835	mcnp	ec2	ec2-m1-large	RUNNING	one-kvm-50	172.16.226.7	
<input type="checkbox"/>	10834	mcnp	ec2	ec2-m1-tiny	RUNNING	one-kvm-24	172.16.226.6	
<input type="checkbox"/>	10832	oneadmin	oneadmin	jlab12-VRouter	RUNNING	one-kvm-srv-01	172.16.220.254 193.205.66.220	
<input type="checkbox"/>	10760	rdh	ec2	ec2-m1-medium	RUNNING	one-kvm-67	172.16.222.1	
<input type="checkbox"/>	10676	solido	ec2	ec2-m1-medium	SUSPENDED	one-kvm-63	172.16.212.3	

Showing 1 to 10 of 194 entries

« 1 2 ... 19 20 » 10

194 TOTAL 191 ACTIVE 2 OFF 1 PENDING 0 FAILED

OpenNebula 4.8.0 by C12G Labs.

# SUNSTONE DASHBOARD

OpenNebula Sunstone: Clc x

https://one-master.to.infn.it

OpenNebula Sunstone Virtual Machines

oneadmin OpenNebula

Dashboard System Virtual Resources **Virtual Machines** Templates Images Files & Kernels Infrastructure Marketplace OneFlow Support

ID	Owner	Group	Name	Status	Host	IPs	VNC
8188	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-07	192.168.3.70	
8185	oneadmin	oneadmin	WN-EC2-8slots	RUNNING	one-kvm-24	192.168.3.69	
8131	oneadmin	oneadmin	APEL-EC2	RUNNING	one-kvm-srv-02	192.168.0.243 193.206.184.53	
7863	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-25	172.16.216.8	
7838	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-34	172.16.216.7	
7837	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-20	172.16.216.6	
7836	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-36	172.16.216.5	
7835	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-37	172.16.216.4	
7834	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-19	172.16.216.3	
7833	giunti	ec2	ec2-m1-large	RUNNING	one-kvm-18	172.16.216.2	

Showing 161 to 170 of 194 entries

10

« 1 2 ... 16 17 18 19 20 »

194 TOTAL 191 ACTIVE 2 OFF 1 PENDING 0 FAILED

OpenNebula 4.8.0 by C12G Labs.

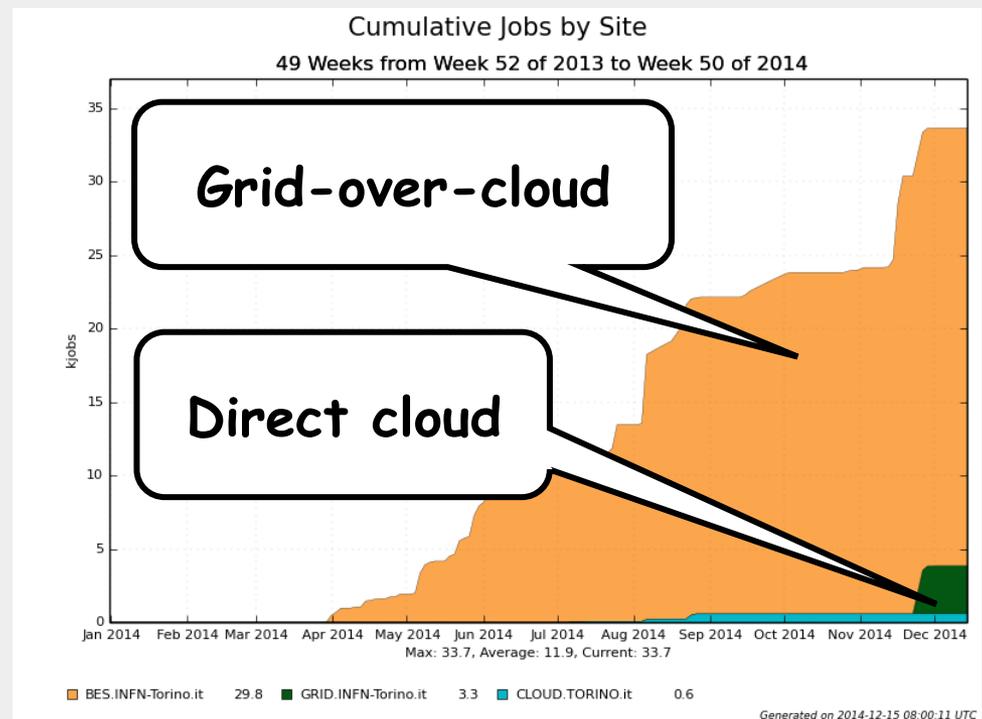
# RECENT AND CURRENT ACTIVITY 1

- Upgrade to OpenNebula 4.8 was not completely painless
  - Several things changed
  - We have a few hacks (to support the elastic IP model)
- Making plans for infrastructure 2.0
  - Big overhaul and rationalization planned for next year
  - iSCSI storage upgrade
  - fabric monitoring infrastructure
  - network redesign around 10Gb servers
  - database service consolidation?

# RECENT AND CURRENT ACTIVITY 2

- Coordination of the OCP Monitoring Working Group
  - See Sara's talk later
- Monitoring-as-a-service prototype in place
  - Also see Sara's talk later
- OpenStack test installation being built
  - Primarily for the OCP testbed
  - Will also serve as a test bench for interoperability
- Application integration and operations support
  - Lots thereof

- Direct interface of DIRAC to the OpenNebula Cloud Manager
  - Done using the rOCCI interface instead of EC2 because DIRAC's rOCCI modules are more mature
  - This is on a separate test instance of OpenNebula
  - DIRAC server at IHEP can directly start and stop VMs according to the number of running jobs
- Several production and stress tests
  - Another large one planned for the coming weeks
  - This is done using the Grid infrastructure, that runs on top of the production Cloud instance



- Monitoring-as-a-service and local accounting/billing
- Automatic elasticity where applicable
- Interoperability & integration
- Switch to new components, adopt new ONe tools
  - Use an HTCondor CE to simplify elasticity?
  - Replace homemade VRouters with OpenVSwitch?
  - Use OneFlow (or Cloudfify or whatever) to orchestrate complete virtual farms?
- Even more application support...

# NEXT BIG CHALLENGE: C3S

- The University of Torino applied to a call by Compagnia di San Paolo for a grant to build an inter-departmental scientific computing facility next year
- Funding approval is not yet final but deemed “highly likely”, total amount at least 500 kEUR
- INFN-Torino is a partner in the project and will host the facility

# NEXT BIG CHALLENGE: C3S

- The “Centro di Competenza per il Calcolo Scientifico” (**C<sup>3</sup>S**) will serve both as a production and an R&D facility
  - Diverse and manifold infrastructure: conventional nodes, “fat” nodes with 4-way servers and huge memory, GPU nodes
  - High-bandwidth, low-latency interconnection (InfiniBand?)
  - Two-tier storage
  - Will share the physical infrastructure with the existing cloud facility
- Will also help building a forum for scientific computing-related activities, and be a platform on which to host further projects

## Our ideas:

- Manage the “conventional” part of the facility with a cloud infrastructure like the existing one
  - But this is a true HPC cluster!
- Explore the opportunity of managing heterogeneous resources with a single tool
  - e.g.: GPU virtualization? Can containers help?
- Experiment with interoperability and cloudbursting

- Gaining a lot of operational experience, solving issues as they present themselves
  - “Agile” development cycle
  - Need to keep growth under control (in terms of number of tenants)
- We will stick to OpenNebula unless somebody provides extremely good reasons
  - Originally adopted for historical reasons
  - However, proved itself ideal for our environment
- Many future developments still vague
  - Will be guided by coordination with INFN-wide strategy
  - C3S will be a challenge but also an opportunity

- S. Bagnasco, D. Berzano, R. Brunetti, S. Lusso, S. Vallero, “Managing a Tier-2 Computer Centre with a Private Cloud Infrastructure”, Proceedings of ACAT2013, Beijing, *J. Phys.: Conf. Ser.* 1742-6596 **523** 012012 (2014)
- S. Bagnasco, D. Berzano, R. Brunetti, S. Lusso, S. Vallero, “Integrating multiple computing needs via a Private Cloud infrastructure”, proceedings of CHEP2013, Amsterdam, *J. Phys.: Conf. Ser.* 1742-6596 **513** 032100 (2014)
- D. Berzano, J. Blomer, P. Buncic, I. Charalampidis, G. Ganis, G. Lestaris and R. Meusel, “PROOF as a Service on the Cloud: a Virtual Analysis Facility based on the CernVM ecosystem”, proceedings of CHEP2013, Amsterdam, *J. Phys.: Conf. Ser.* 1742-6596 **513** 032007 (2014)
- S. Vallero, A. Guarise, S. Lusso, D. Berzano, S. Bagnasco, “Monitoring of IaaS and scientific applications on the Cloud using the Elasticsearch ecosystem”, poster presented at ACAT 2014, Prague.