CLOUD COMPUTING A TORINO

Stefano Bagnasco, Riccardo Brunetti, Stefano Lusso (INFN-Torino), Sara Vallero (Università di Torino), Dario Berzano (CERN)

Workshop Commissione Calcolo e Reti | Genova, 27-31 maggio 2013

PHILOSOPHY

- Ensure **interoperability**
- Favour **manageability** and **flexibility** over performance
- Provide a **production service** to applications

Keep it simple

Stay mainstream

Be user-oriented

PHILOSOPHY

- Don't use too many tools
- Develop as few pieces as possible
- Introduce features only when needed by applications
- Use few simple images plus contextualization

Keep it simple

Stay mainstream

Be user-oriented

PHILOSOPHY

Choose stable and widely used tools and components:

- OpenNebula cloud stack
  - Common interfaces: OCCI, EC2, OCA
- GlusterFS filesystem
- OpenWRT for network management

Keep it simple

Stay mainstream

Be user-oriented

PHILOSOPHY

- Adopt an **agile development** cycle
- Give resources to users as soon as possible
- Add functionalities as they become needed

Keep it simple

Stay mainstream

Be user-oriented

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

TWO CLUSTERS

- Server-class hardware
- Shared image repository
- Resiliency-optimized FS for shared system disks
- Currently 4 hosts

- Working-class hardware ☺
- Cached image repository
- Access to perfomance-optimized FS for data needs
- Currently 35 hosts

THE INGREDIENTS

- Cloud management Toolkit: **OpenNebula**
  - OpenStack, now widely adopted in new projects, was too embrionic when we started
  - ...and arguably* OpenNebula is better suited at Data Center Nebulization
  - Currently using version 3.6, will migrate to 3.8/4.0 soon
  - Templates based on few very simple images plus full contextualization via scripts and puppet (looking into CloudInit)
- Backend storage: **GlusterFS**
  - Flexible enough to cater to different needs with a single tool
- VM network management: **OpenWRT**
  - OpenVSwitch was not integrated in OpenNebula when we started

* See e.g. blog.opennebula.org/?p=4042

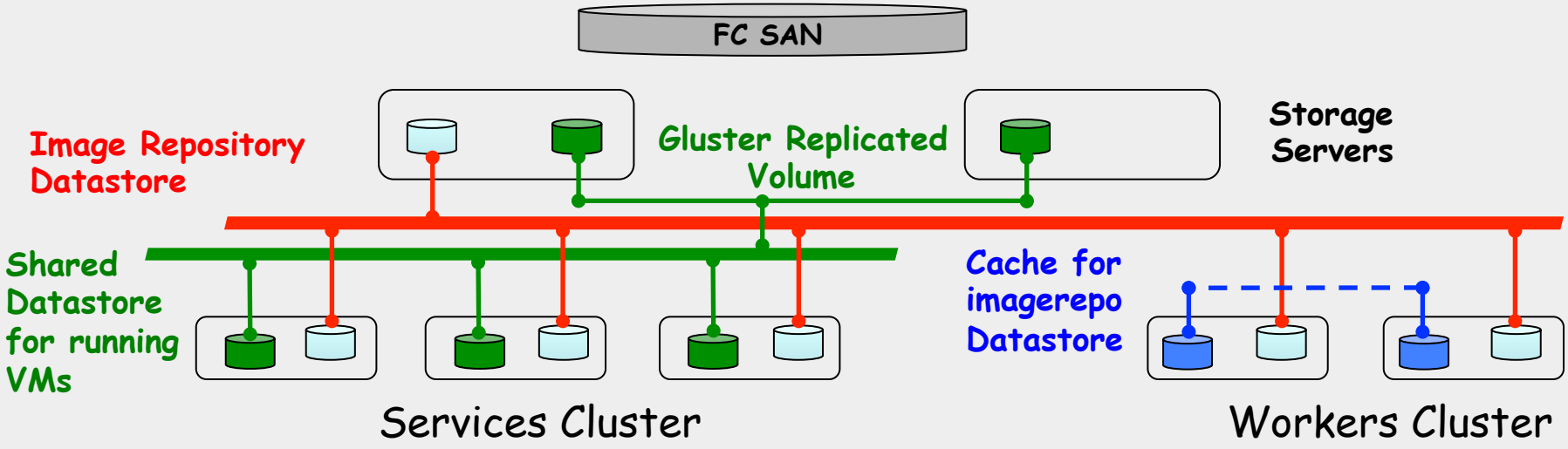
Our use cases:

- VM image repository:
  - one brick exported
- System datastore for service-class hosts:
  - replicated on two servers for redundancy.
  - Replica is synchronous, self-healing enabled.
  - Continuous r/w occurs
- Experiment data
  - pool of aggregated disks (currently ~50 TB).
  - Very high throughput towards many concurrent clients

MULTIPURPOSE STORAGE: GLUSTERFS

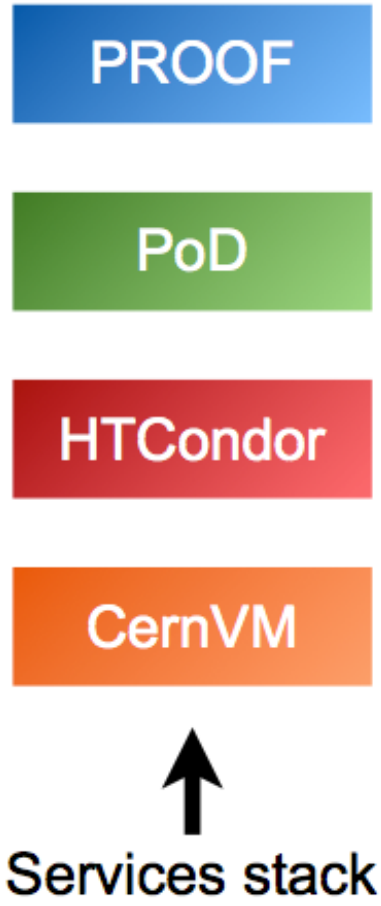
Two storage servers with 10Gbps interface provide some of the LUNs through GlusterFS

- All the virtual machines run on RAW or QCOW file images
- 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.
  - An ad-hoc script synchronizes the local copies using a custom "torrent-like" tool (scpWave + rsync) when new versions of the images are saved



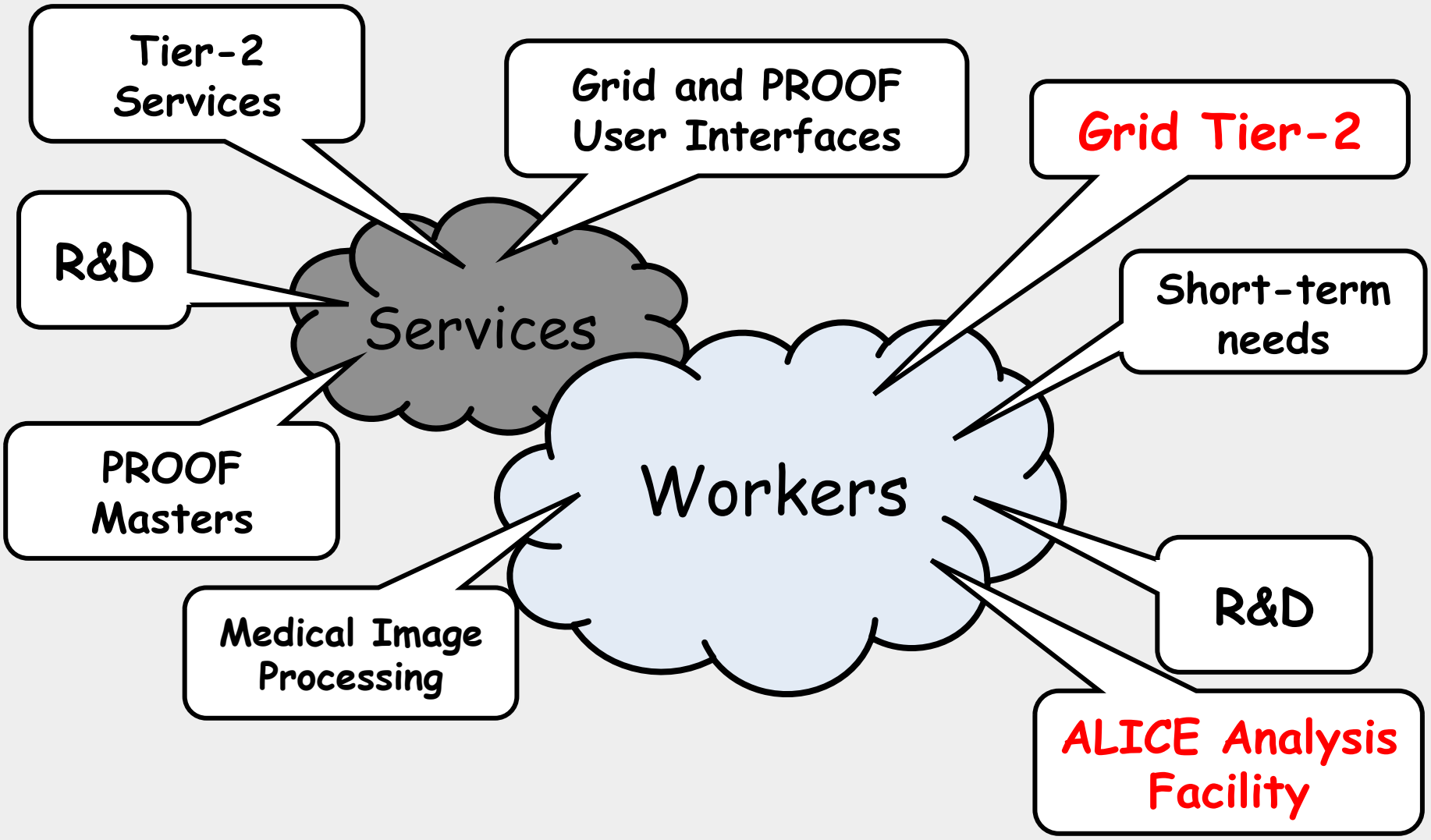
VAF components: overview

- User interacts with:
  - PoD to request and book workers
  - PROOF to execute jobs
- Under the hood:
  - worker requests are scheduled by HTCondor
  - CernVM virtual machines are part of the HTCondor cluster



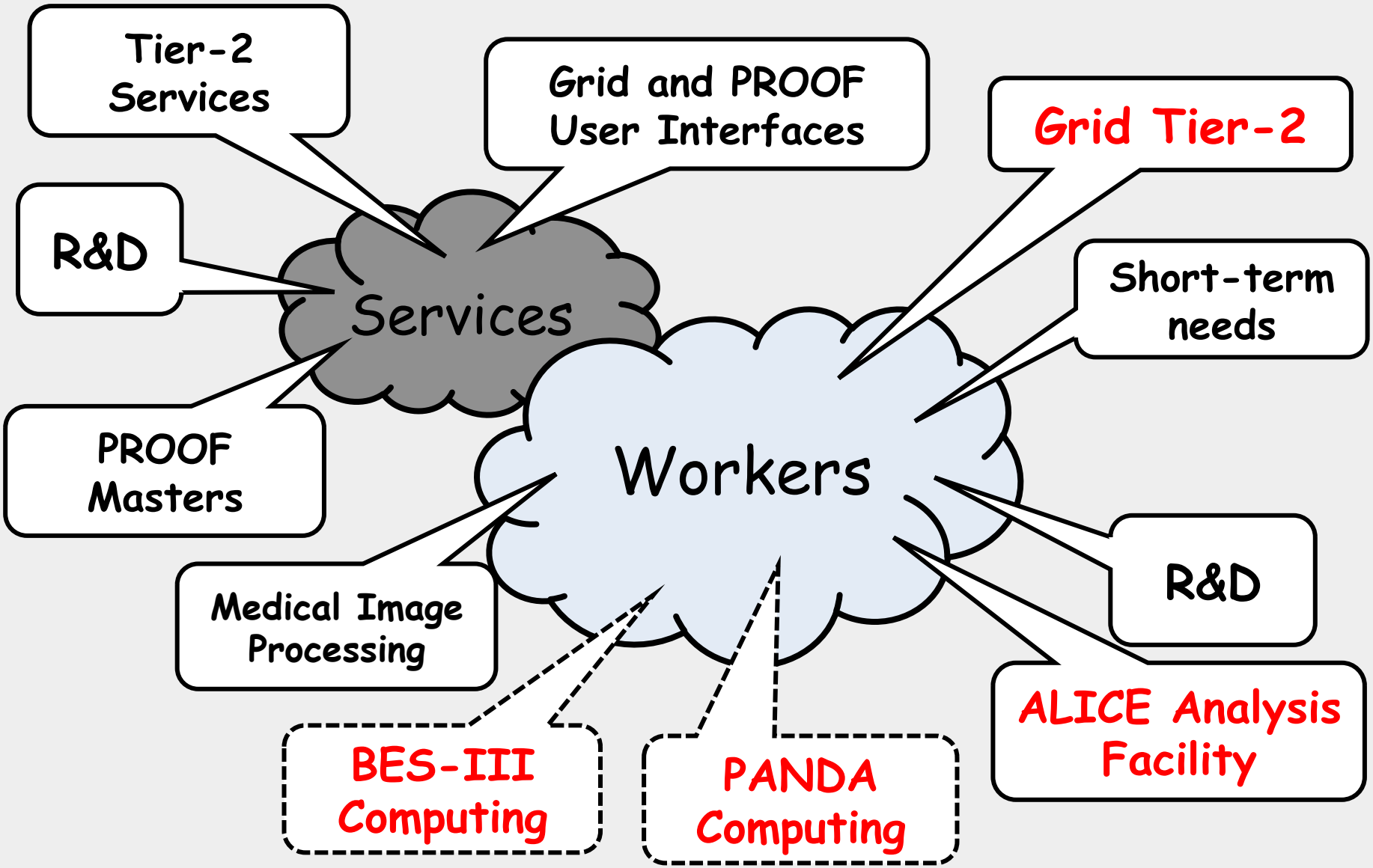
Dario.Berzano@cern.ch
http://goo.gl/CFnMM

STAKEHOLDERS



中國科學院高能物理研究所
Institute of High Energy Physics Chinese Academy of Sciences

STAKEHOLDERS



What is CernVM?

CernVM is a baseline Virtual Software Appliance for the participants of CERN LHC experiments. The Appliance represents complete, portable and easy to configure user environment for developing and running LHC data analysis locally and on the Grid, independently of Operating System software and hardware platform (Linux, Windows, MacOS).
The goal is to remove a need for the installation of the experiment software and to minimize the number of platforms (compiler-OS combinations) on which experiment software needs to be supported and tested.

CernVM Projects

FUTURE DEVELOPMENTS

- Understand the opportunities given by the CernVM "ecosystem"
- Study the integration of the OpenNebula AuthN/AuthZ system in a VO context or using federated authentication mechanisms.
- Try using CloudInit to consolidate and simplify contextualization
- Integrate (at least the AAF) into some accounting system
- Explore the GlusterFS UFO Object Storage to provide a "DropBox-like" storage to users.
- Participate in upcoming projects aimed to develop a higher-level federated cloud infrastructure