Model for a Azerbaijan Cloud Data Centre for Science

CERN April 24th 2015

Overview

- Why Clouds for Scientific Computing
- CERN's Cloud infrastructure
- Model for the evolution of a Cloud based data centre for science
- Discussions

Why Clouds?

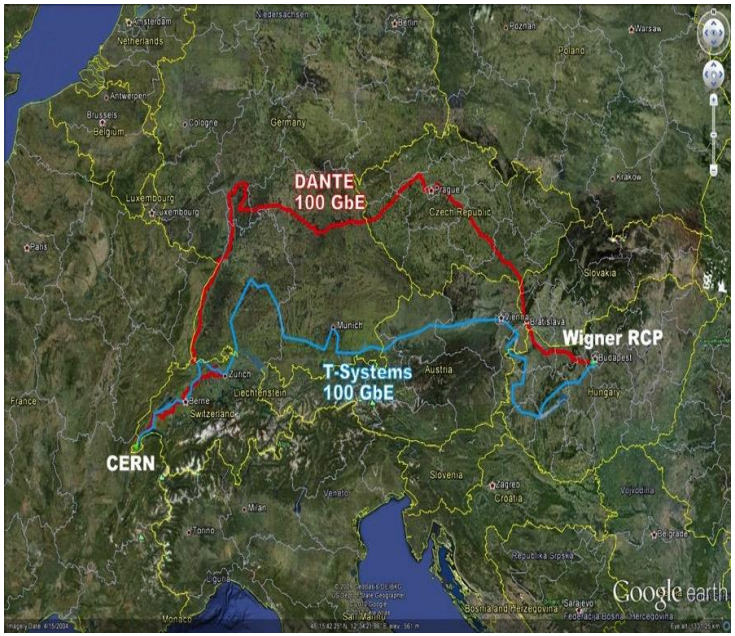
- Clouds offer flexibility
  - user workloads and system requirements are decoupled
  - dynamic allocation of resources
  - commercial and non-commercial providers
- Based on established, open technology and protocols
  - expertise is widely available
  - products and tools evolve rapidly
  - commercial and non-commercial users
- Proven scalability
  - small in-house systems to world wide distributed systems

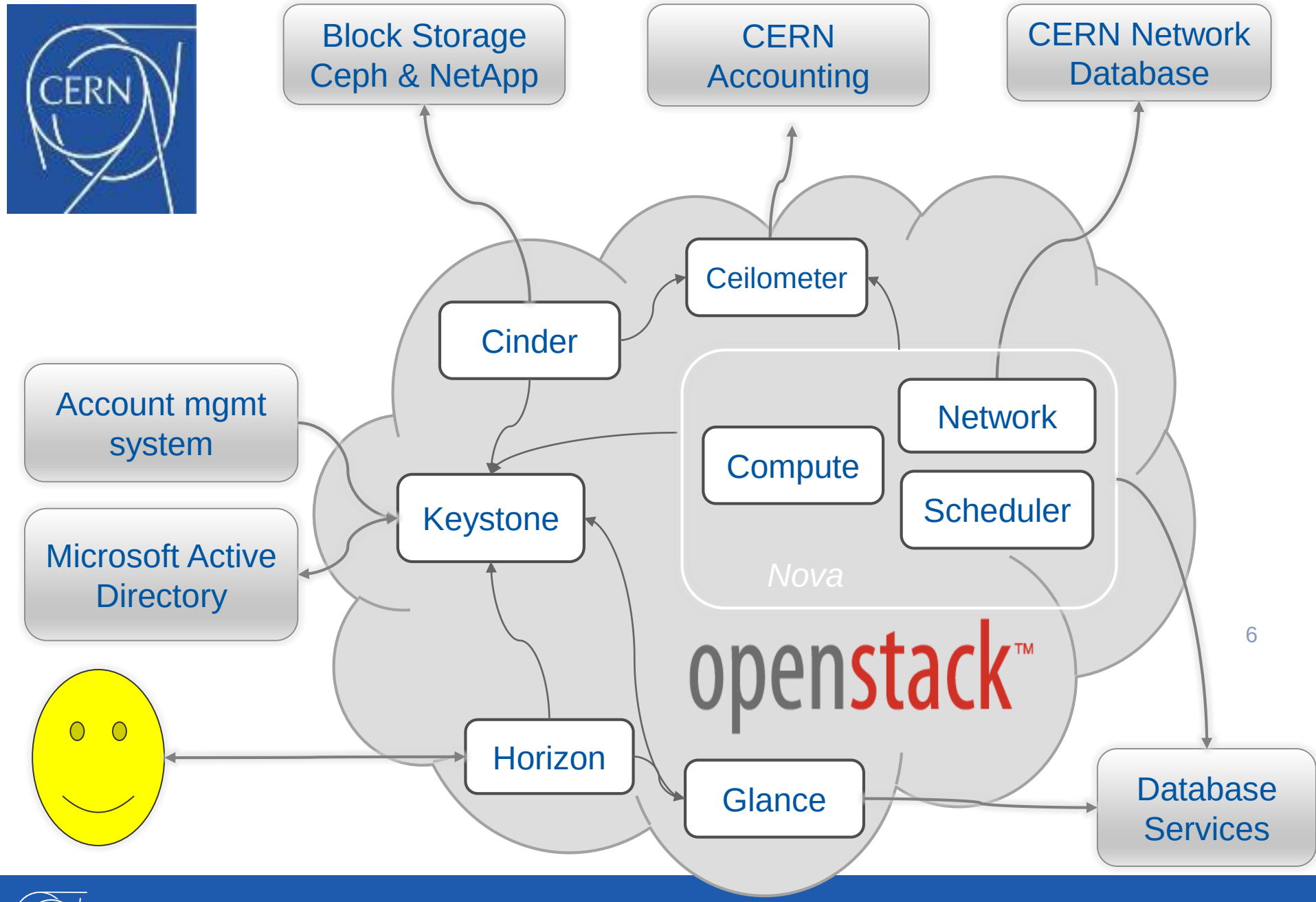
CERN Cloud use-cases

- Compute intensive physics computing
  - More than 90% of our resources
  - Linux VMs, managed LHC experiments
  - Many identical machines
    - “Cattle”, low SLA
- Service nodes, personal VMs
  - Webservers, development boxes, etc.
  - Includes Windows VMs
  - I/O intensive
  - Huge variety of machines
    - “Pets”, often requiring high uptime
- Both require high level of automation

CERN Private Cloud - Numbers

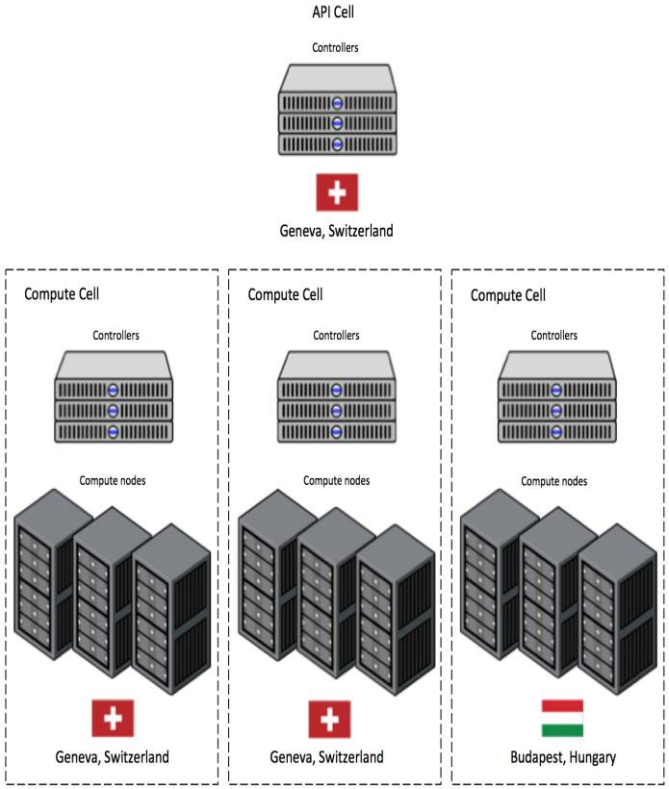
- Based on OpenStack Juno
- Spans between 2 datacentres
- 5000 hypervisors
  - 120000 cores
- 11000 VMs
- 1500 users
- 1800 projects



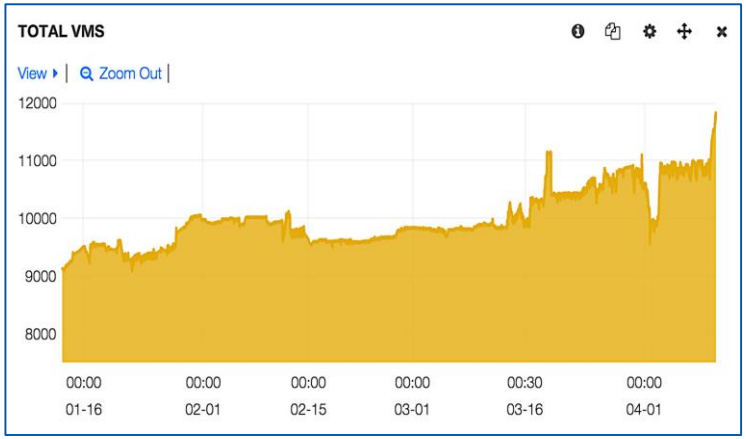

Block Storage
Ceph & NetApp
CERN
Accounting
CERN Network
Database
Ceilometer
Cinder
Account mgmt
system
Network
Compute
Scheduler
Keystone
Microsoft Active
Directory
Nova
openstack™
Horizon
Glance
Database
Services

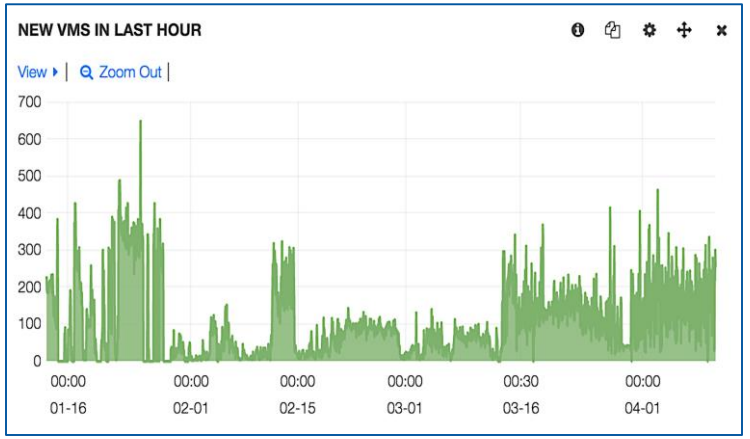
Scaling based on Nova Cells

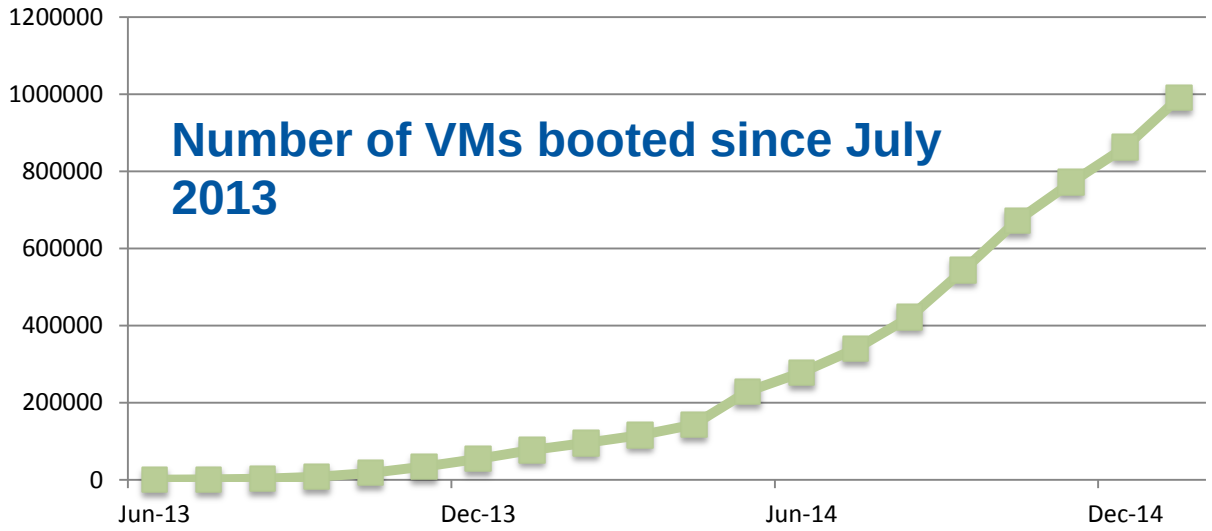
- Groups of 200 – 1000 hosts configured as a tree
- Add cells to grow the deployment
- CERN: 5000 servers in 15 cells



CERN Cloud usage figures







Storage for Science at CERN

- Cloud based storage for cloud services
  - OpenStack Cinder
    - Ceph reliable backend
- EOS for Physics research
  - XROOTD and HTTP
    - for high performance
  - Kerberos and X509
    - for Grid and local usage
  - scalable to >>100PByte
- CASTOR
  - tape archive system
  - >>100PByte
  - forever

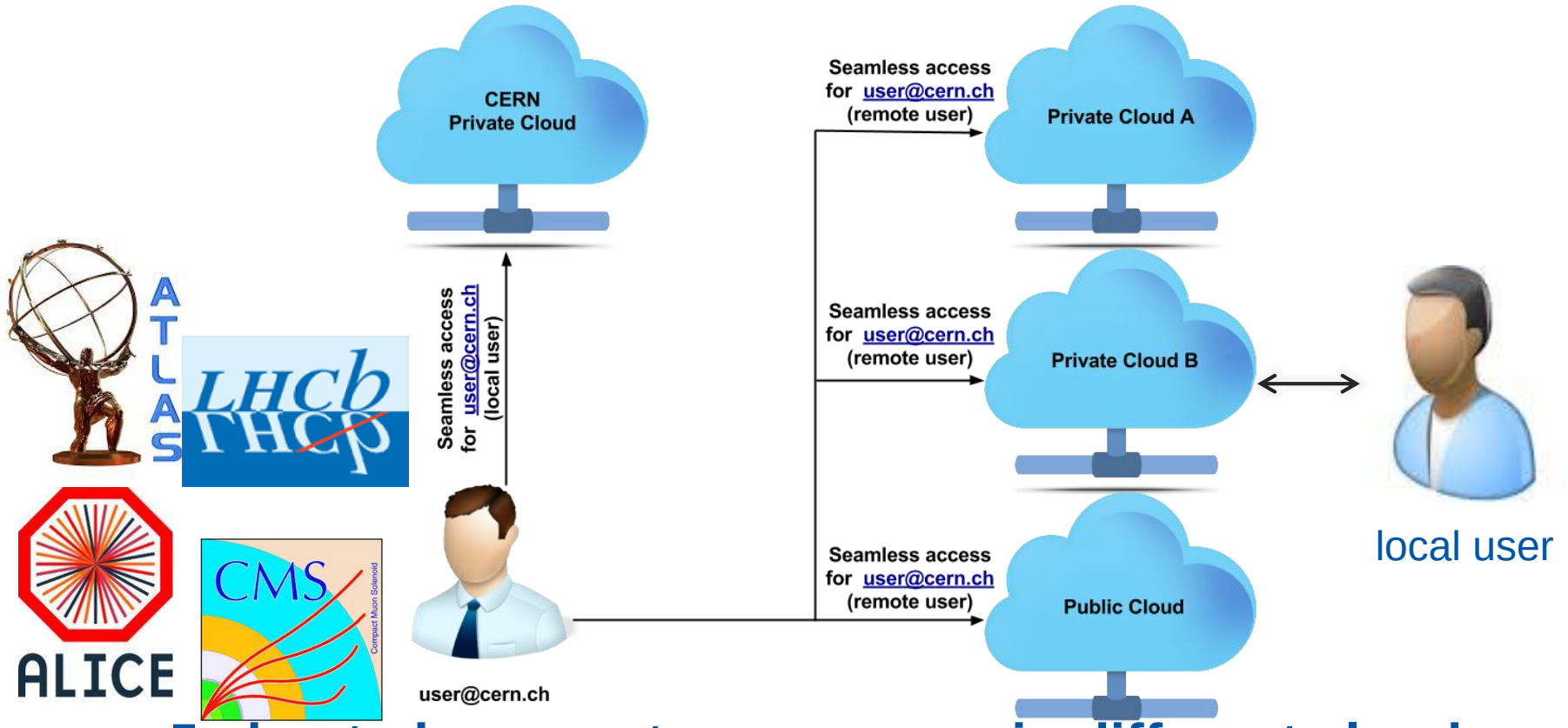
WLCG Experience

- Avoid:
  - "Everything has to work before anything works"
- Computational resources are manageable
  - everything is transient
  - partial systems are already useful
  - scaling/replacing is straight forward
- Storage is difficult
  - partial system is useless
  - reliability&availability a must
  - adding/replacing is complex
- → Multiphase deployment model works best

Possible deployment scenario: L1

- Start with subset of OpenStack components
  - Nova, Keystone, Glance, Neutron, Horizon, Ceilometer
- Use modern fabric/config management tools
  - kickstart, puppet, nagios etc.
- Deploy using off-the-shelf components
  - Commercial distributions exist
    - incl. training and support
  - CERN uses the RDO community distribution
    - with Puppet modules
- Integrate this into your operations infrastructure
  - Server deployment, monitoring, maintenance

Configure OpenStack Federation



Federated access to resources in different clouds

Azerbaijan Cloud Data Centre L1

- Can be deployed gradually
  - machines can be added as they get ready
- Can be used by LHC users like CERN
  - data and higher level services at CERN
  - including accounting/crediting
- Can provide computing resources to local users through cloud interfaces
- Can be evolved without disruption

Azerbaijan Cloud Data Centre L2

- Computing resources + local services
  - improved efficiency (latency hiding)
- Add OpenStack Cinder block storage
  - with CEPH as a backend
  - can be used by local (caching) services
    - deployed in VMs (scalable)
    - Squid, cvmFS.....
    - staging services
- Allows local users to deploy services

Azerbaijan Cloud Data Centre L3

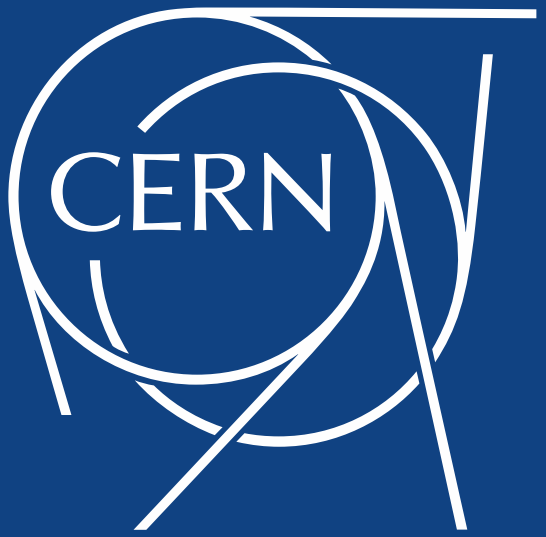
- Computing and Storage Cloud
- Adding local storage
  - CEPH backend can be scaled
  - Can be offered to users via standard interfaces
    - SWIFT, Cinder, S3, RADOS ...
- Local users can use this directly
  - Cloud object store or file system
- WLCG users
  - via VM based Grid Storage elements (EOS)
- This can grow into a WLCG T2/T1 centre

What is needed ?

- Expertise for setting up the system
  - unix/linux sysadmins (2-3 for 2k nodes)
    - training of OpenStack (commercially available)
  - fabric knowledge
    - racks, electrical system, cooling, cabling...
- Estimated time line for L1(difficult prediction)
  - “core servers” : few weeks + training
  - Configuration of Federation: week
  - Hypervisors and nodes can be added gradually

Discussion

- contacts for technical discussions:
- ian.bird@cern.ch
- tim.bell@cern.ch
- jan.van.eldik@cern.ch
- markus.schulz@cern.ch


CERN