

Production Large Scale Cloud Infrastructure Experiences at CERN

CHEP – Amsterdam, 2013

Belmiro Moreira

belmiro.moreira@cern.ch @belmiromoreira



CERN IT Infrastructure challenges in 2011

- Need to manage twice the servers as today
- No increase in staff numbers
- Increasing number of users / computing requirements
- Legacy tools - high maintenance and brittle

Why Build CERN Cloud Infrastructure?

Improve operational efficiency

- Hardware interventions with long running programs
- Multiple operating system demand

Improve resource efficiency

- Exploit idle resources
- Highly variable load such as interactive or build machines

Improve responsiveness

- Self-service

Identify a new Tool Chain

- Identify the tools needed to build our Cloud Infrastructure
 - Configuration Manager tool
 - Cloud Manager tool
 - Monitoring tools
- Storage Solution

OpenStack

“OpenStack is a global collaboration of developers and cloud computing technologists producing the ubiquitous open source cloud computing platform for public and private clouds.”

“The project aims to deliver solutions for all types of clouds by being simple to implement, massively scalable, and feature rich.

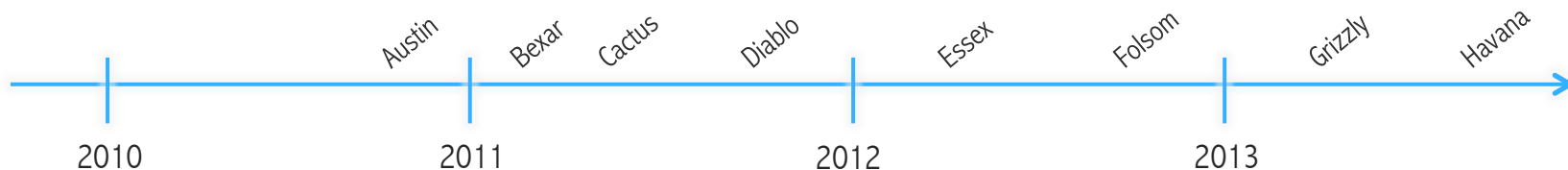
The technology consists of a series of interrelated projects delivering various components for a cloud infrastructure solution.”

OpenStack

- Started in 2010 by NASA and Rackspace

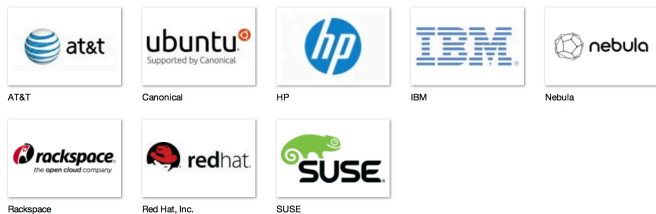


- Openstack releases:

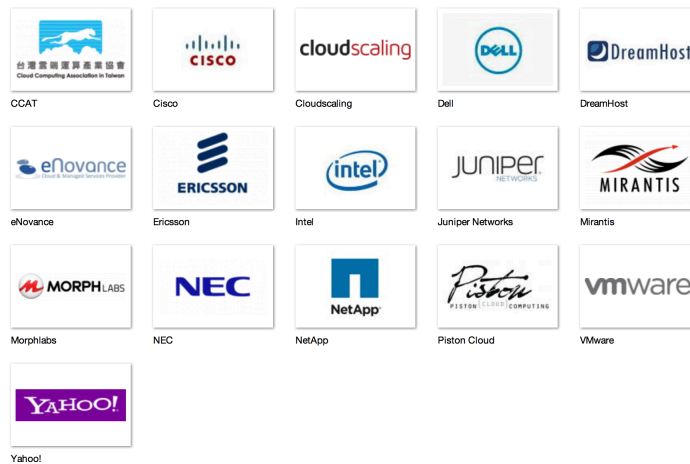


OpenStack Community

- Run by an independent foundation
- + 10000 members
- ~ 200 sponsors

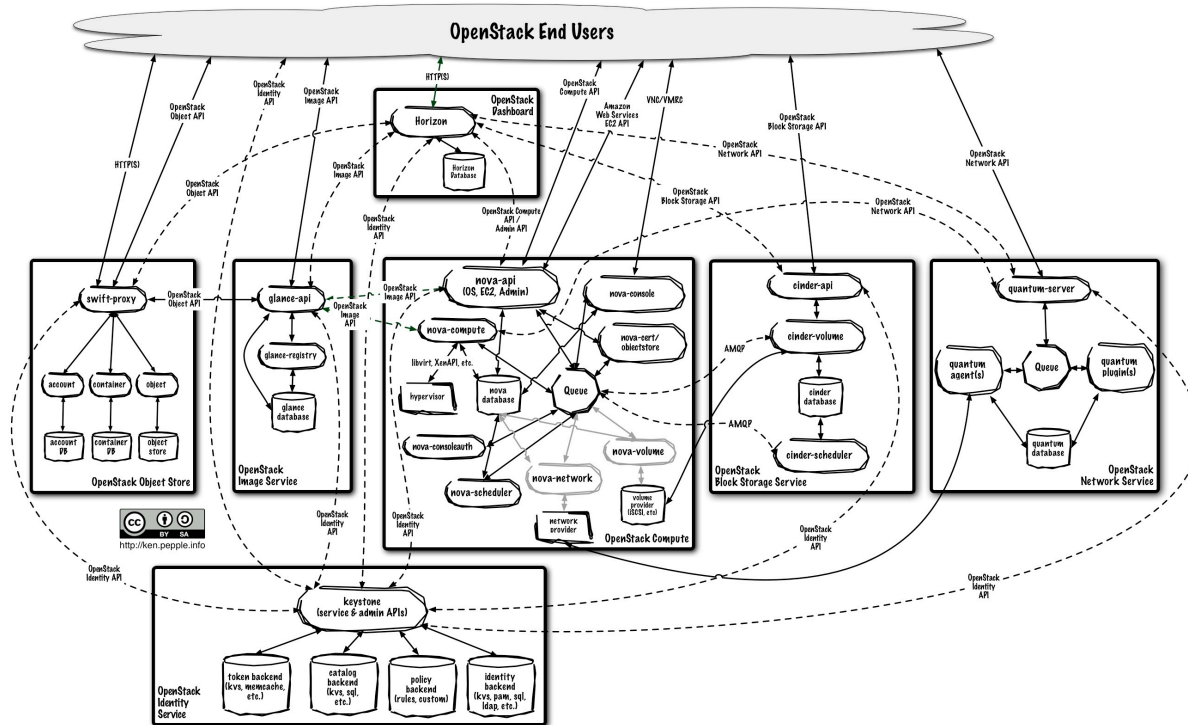


Platinum members



Gold members

OpenStack Components



Strategy to deploy OpenStack

- Configuration infrastructure based on Puppet
- Community Puppet modules for OpenStack
- SLC6 Operating System
- EPEL/RDO - RPM Packages

Strategy to deploy OpenStack

- Deliver a production IaaS service through a series of time-based pre-production services of increasing functionality and Quality-of-Service
- Wigner Computer Center hardware deployed as OpenStack compute nodes
- Have an OpenStack production service in the Q2 of 2013

Pre-Production Infrastructure

Essex



"Guppy"

1st, June of 2012

Folsom



"Hamster"

7th, September of 2012



"Ibex"

1st, March of 2013

CERN Cloud Infrastructure - "Guppy"

- Target audience - AI project members
- Deployed on Fedora 16 (KVM hypervisor only)
- RPMs provided by the Fedora Cloud SIG team
- Community OpenStack puppet modules
- Used for functionality tests
- Limited integration with CERN infrastructure
 - single IP service
 - no CERN accounts

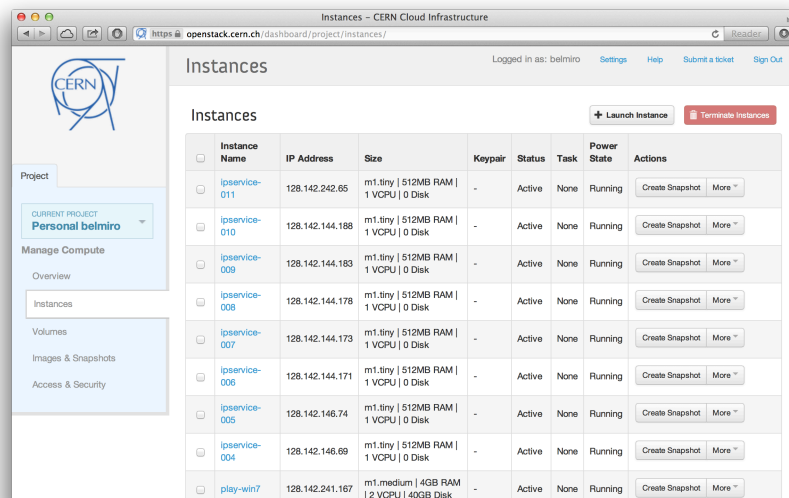
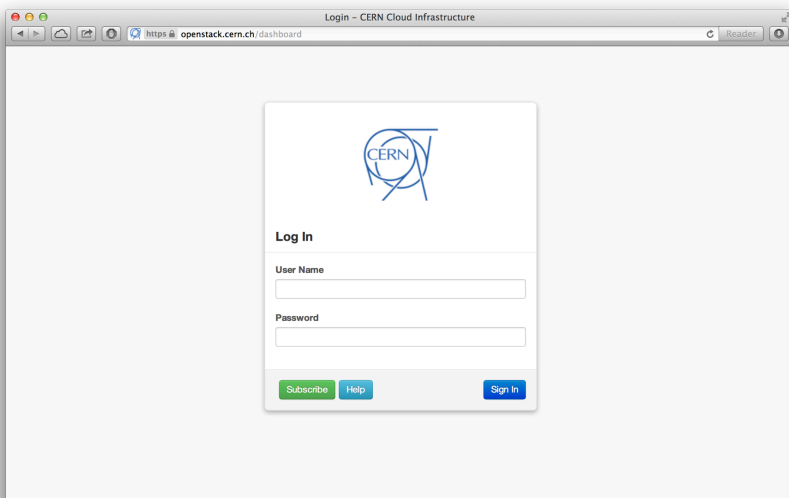
CERN Cloud Infrastructure - "Hamster"

- Open to early adopters
- Deployed on SLC6 and Hyper-V
- CERN Network DB integration
 - Support for multiple IP services
- Keystone LDAP integration

CERN Cloud Infrastructure - "Ibex"

- Open to a wider community
 - Atlas, CMS, LHCb, ...
- Some OpenStack services in HA
- Using CERN dbod service for MySQL
- +600 compute nodes

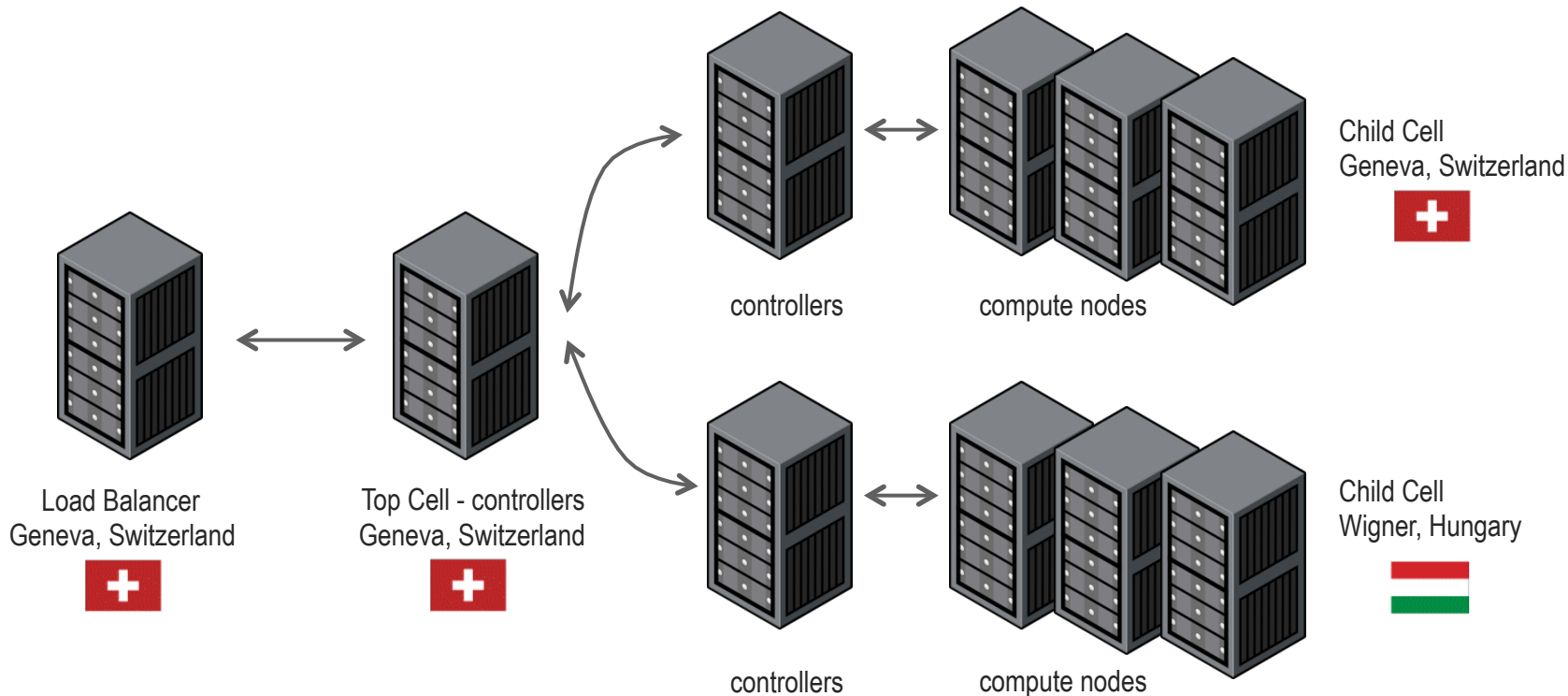
OpenStack at CERN - grizzly release



OpenStack at CERN - grizzly release

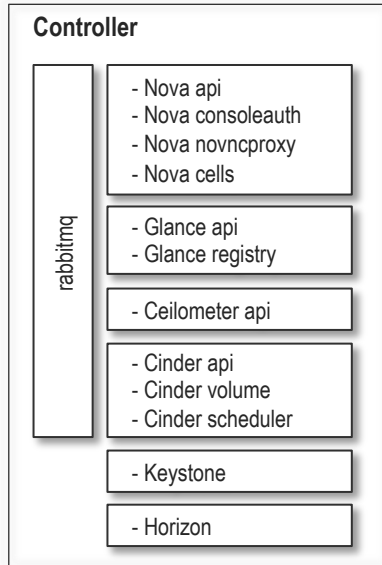
- +2 Cells – Geneva and Wigner Computer Centers
- HA+1 architecture
- Ceilometer deployed
- Integrated with CERN accounts and network
- Monitoring OpenStack components status
- Glance - Ceph backend
- Cinder tests - Ceph backend

Architecture Overview

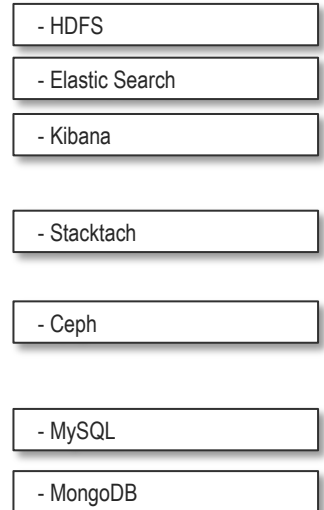
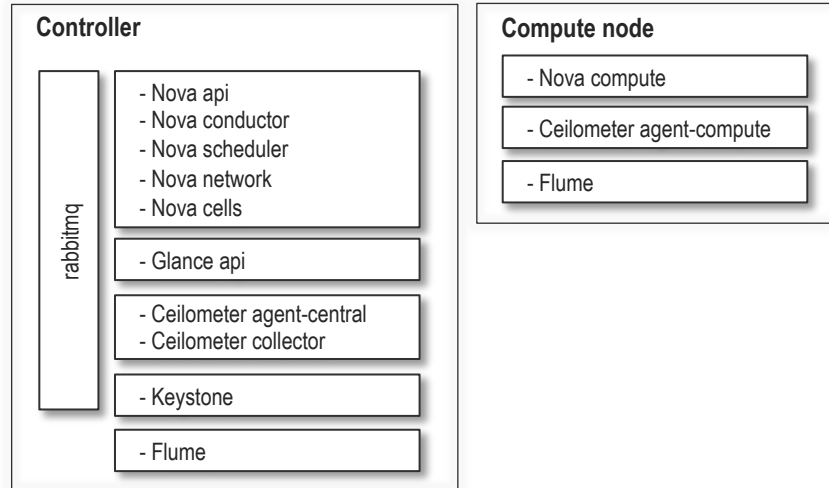


Architecture Components

Top Cell



Children Cells



Infrastructure Overview

- HAProxy as load balancer
- Master and Compute nodes
 - 3+ Master nodes per Cell
 - $O(1000)$ Compute nodes per Cell (KVM and HyperV)
 - 3 availability zones per Cell
- Rabbitmq
 - At least 3 brokers per Cell
 - Rabbitmq cluster with mirrored queues

Infrastructure Experiences

- With cells we lost functionality
 - Security groups
 - Live migration
- Cell Scheduling
 - Random cell selection on Grizzly
 - Implemented simple scheduler based on project
 - CERN Geneva only, CERN Wigner only, “both”

Infrastructure Experiences

- Challenge to integrate OpenStack with network model at CERN
 - Implemented Nova Network CERN driver
 - MAC address selected from pre-registered addresses of “host” IPService
 - Network constraints in some nova operations
 - Resize, Live Migration

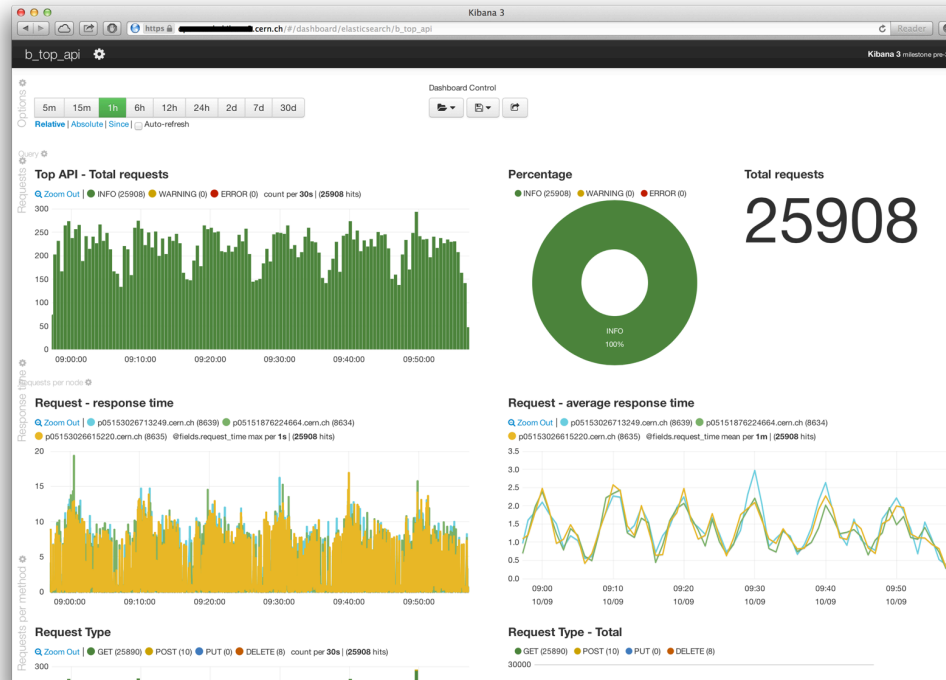
Infrastructure Experiences

- Keystone integrated with CERN identity
 - LDAP backend
 - CERN user subscribes the "cloud service"
 - Created "Personal Tenant" with limited quota
 - Shared Projects created by request
- Glance API
 - Using glance api v1
 - python-glanceclient doesn't support completely v2

Monitoring - Flume, Elastic Search, Kibana

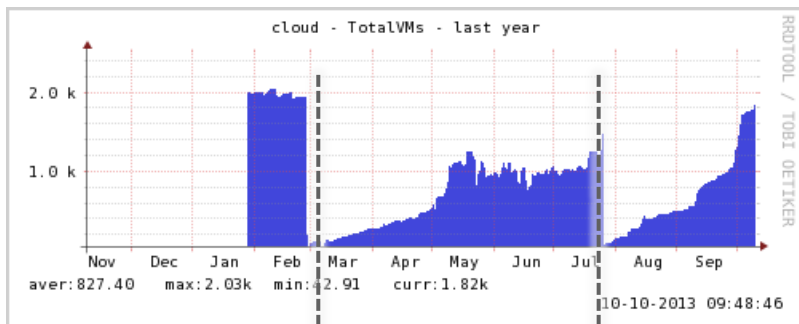
- How to monitor OpenStack status in all nodes?
 - ERRORS, WARNINGS – log visualization
 - identify in “real time” possible problems
 - preserve all logs for analytics
 - visualization of cloud infrastructure status
 - service managers
 - resource managers
 - users

Monitoring - Kibana



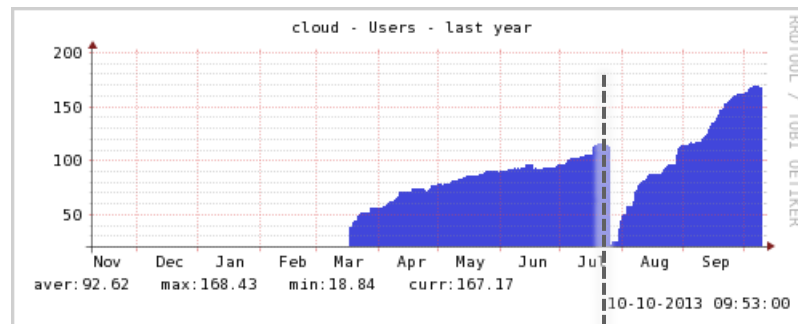
CERN Cloud Infrastructure adoption

Number of VMs



Grizzly

Number of Users



Grizzly

| Cell | Nodes | Cores | RAM (GB) | Disk (TB) | VMs |
|--------|-------|-------|----------|-----------|------|
| Geneva | 375 | 10976 | 21662 | 632 | 1438 |
| Wigner | 291 | 9312 | 18296 | 491 | 674 |
| Total | 666 | 20288 | 39958 | 1123 | 2112 |

(13/10/2013)

Next Challenges

- Growing the infrastructure
 - +100 compute nodes per week
 - 15000 servers – more than 300000 cores
- Migration from Grizzly to Havana
- nova-network deprecation
- Kerberos, X.509 user certificate authentication
- Keystone Domains

belmiro.moreira@cern.ch
@belmiromoreira



www.cern.ch