CERN Cloud status

CREATED BY THE CLOUD INFRASTRUCTURE TEAM.

Presented by Thomas Oulevey / @thomasnomason the behalf of the team.

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Agenda

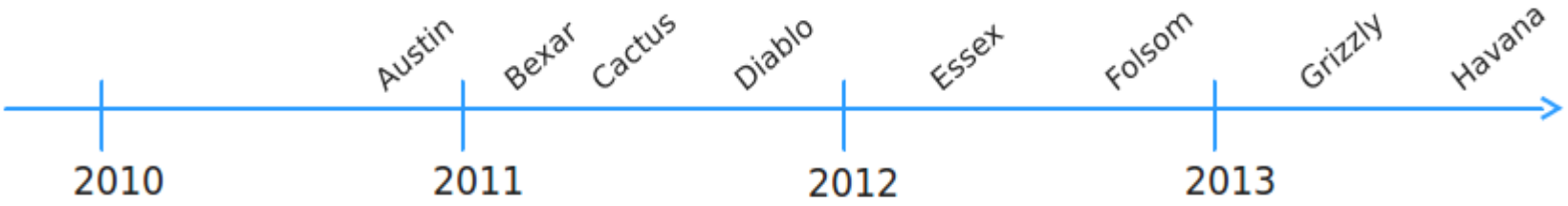
CERN IT Department

- Openstack
- Architecture
- What's new since Spring?
- Next challenges
- Questions

CERN IT Department
CH-1211 Genève 23
Switzerland
www.cern.ch/it

Started in 2010 by NASA and Rackspace

- Current release : Havana (October 2013)
- Next release : Icehouse (2014)

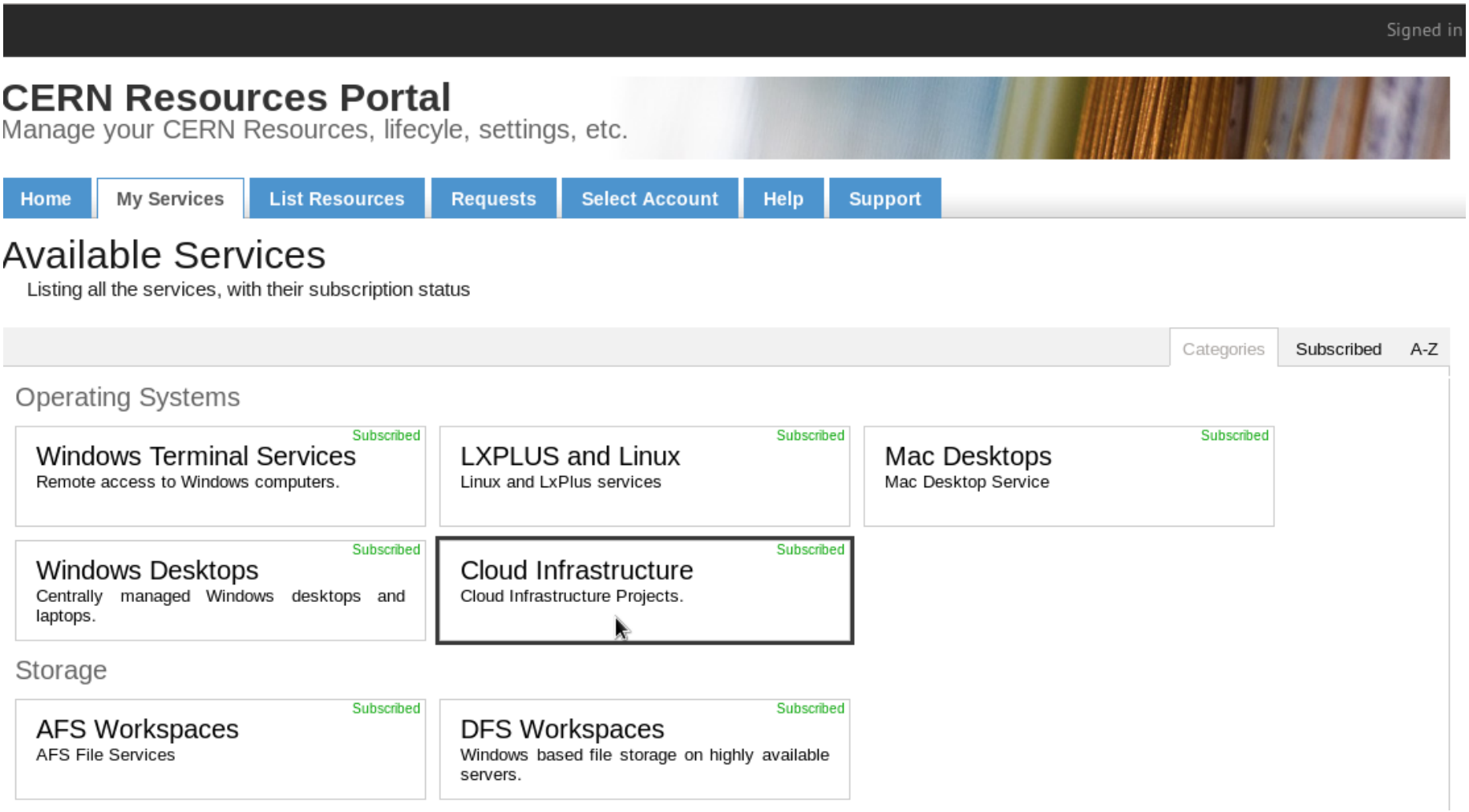


- Grizzly is our current production,
- based on RDO(upstream Redhat RHOS) packages,
- on Scientific Linux CERN 6,
- deployed with community puppet modules from stackforge.

openstack.cern.ch

- Production service,
- Around 700 hypervisors,
- Adding 100 per week to reach 15000 by the end of 2015,
- Self-service from web portal,
  - Personal project with 10 VMs quota
  - Request projects
- Support line in place and integrated with our tools.

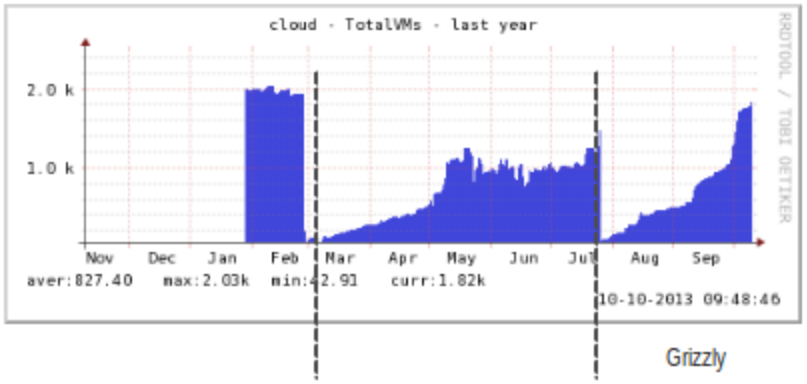
- code: https://github.com/cernops
- documentation: http://cern.ch/go/lWn8


Signed in
CERN Resources Portal
Manage your CERN Resources, lifecyle, settings, etc.
Home
My Services
List Resources
Requests
Select Account
Help
Support
Available Services
Listing all the services, with their subscription status
Categories
Subscribed
A-Z
Operating Systems
Subscribed
Windows Terminal Services
Remote access to Windows computers.
Subscribed
LXPLUS and Linux
Linux and LxPlus services
Subscribed
Mac Desktops
Mac Desktop Service
Subscribed
Windows Desktops
Centrally managed Windows desktops and laptops.
Subscribed
Cloud Infrastructure
Cloud Infrastructure Projects.
Storage
Subscribed
AFS Workspaces
AFS File Services
Subscribed
DFS Workspaces
Windows based file storage on highly available servers.

in numbers

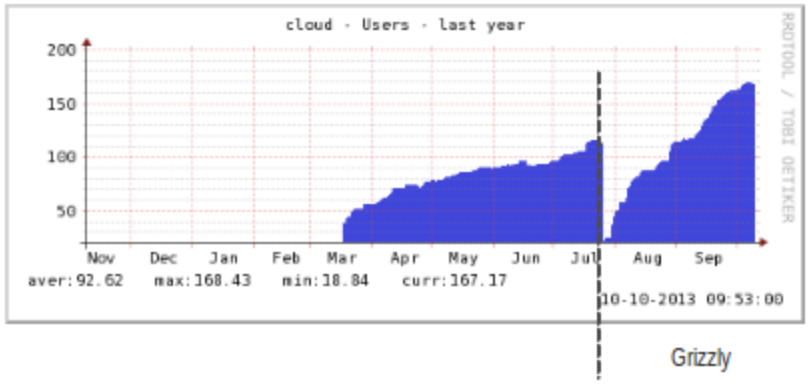
CERN IT Department

Number of VMs



Grizzly

Number of Users

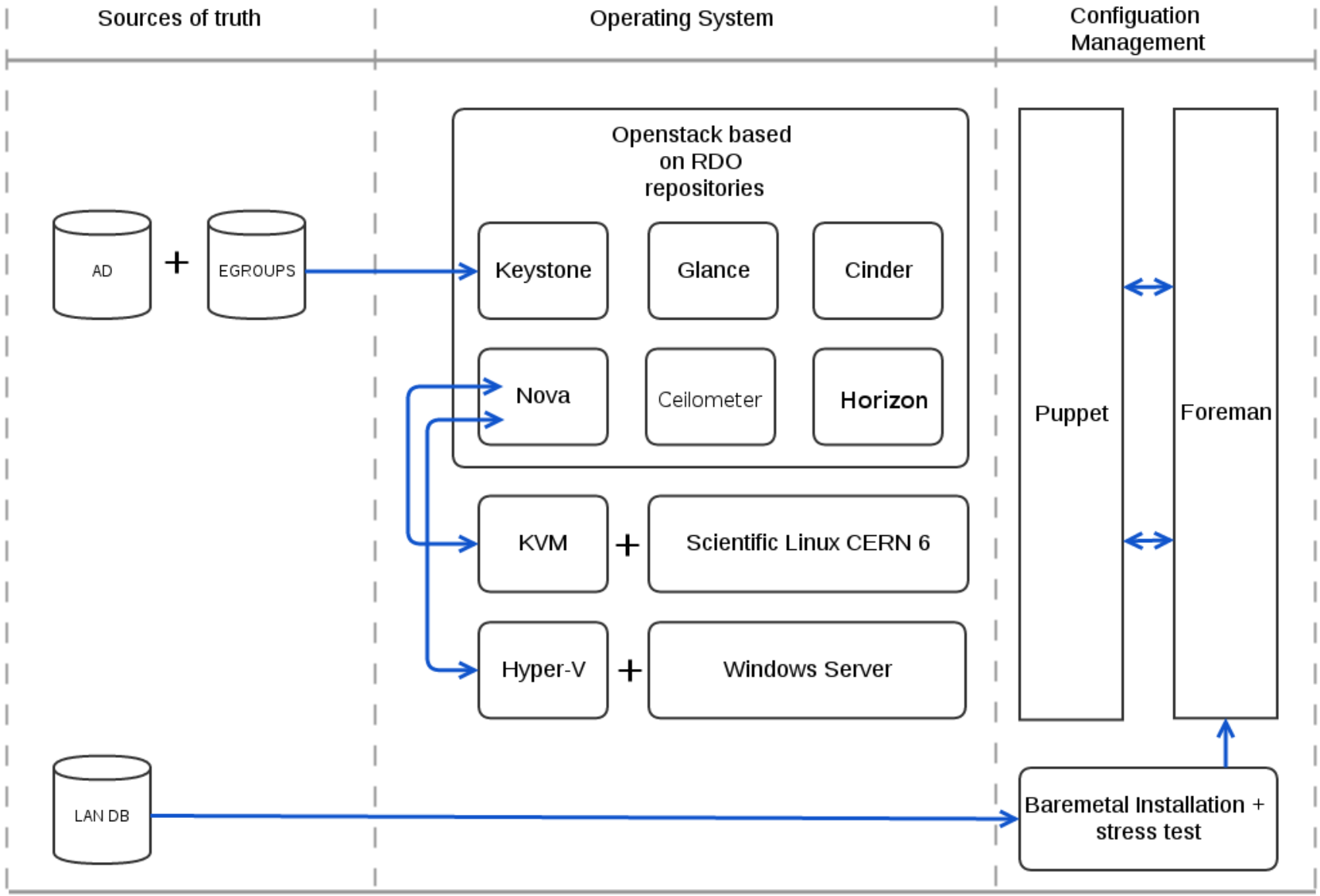


Grizzly

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)

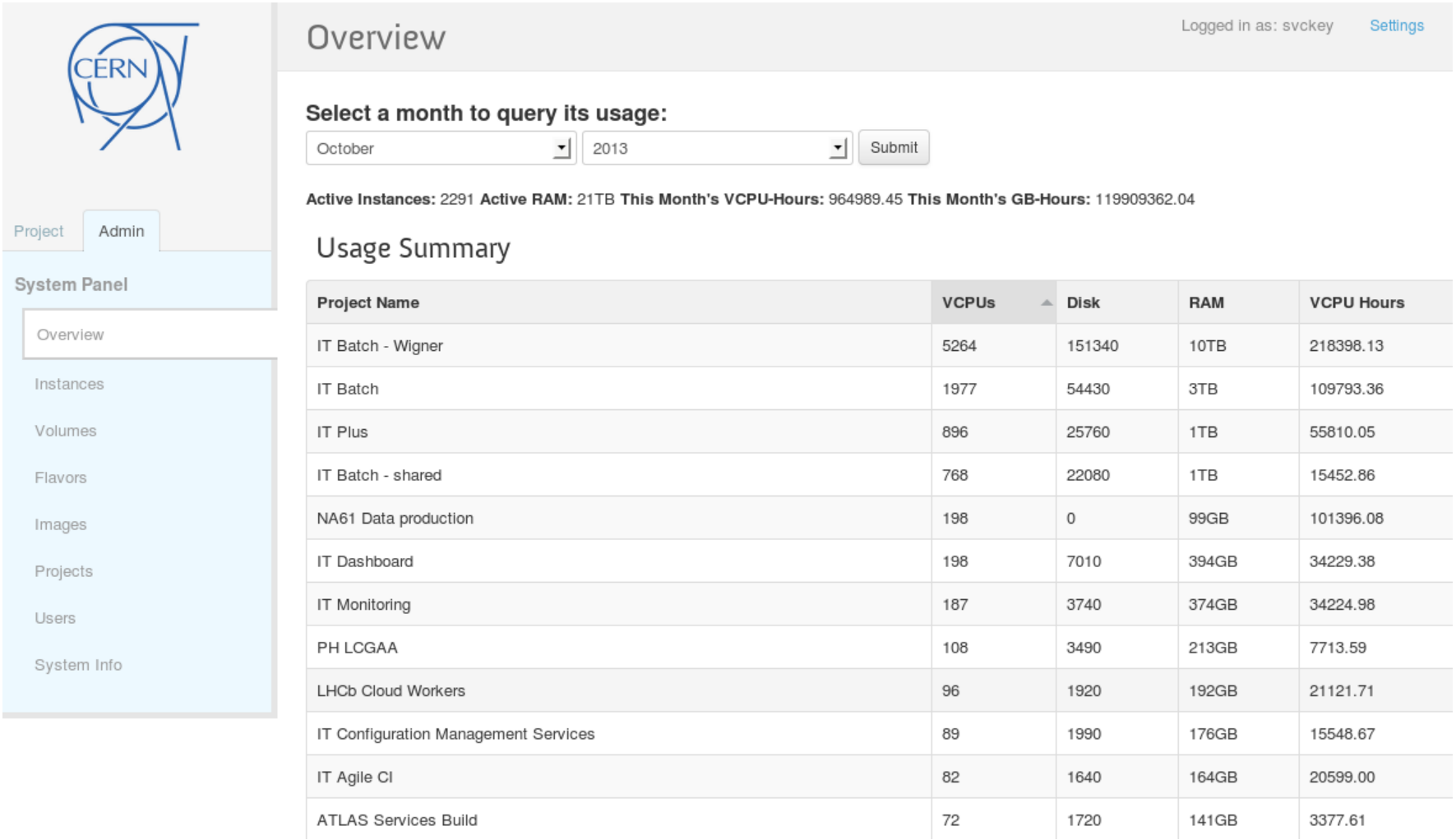
architecture



Grizzly features

- Horizon Dashboard,
- Nova cells: Geneva and Wigner,
- Nova compute node,
- Glance images,
- Cinder and Ceph,
- Ceilometer,
- Monitoring,
- EC2 and nova-api.

Admin interface

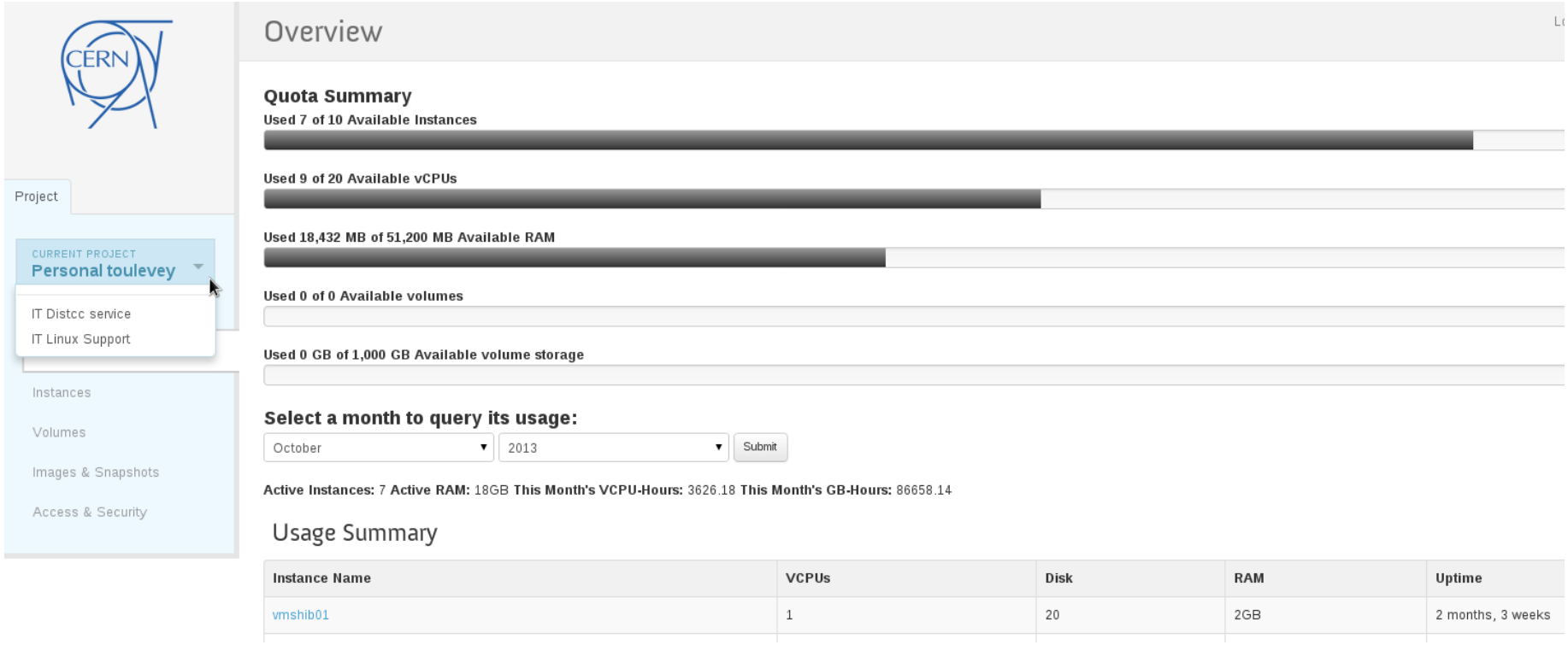


Project Name	VCPUs	Disk	RAM	VCPU Hours
IT Batch - Wigner	5264	151340	10TB	218398.13
IT Batch	1977	54430	3TB	109793.36
IT Plus	896	25760	1TB	55810.05
IT Batch - shared	768	22080	1TB	15452.86
NA61 Data production	198	0	99GB	101396.08
IT Dashboard	198	7010	394GB	34229.38
IT Monitoring	187	3740	374GB	34224.98
PH LCGAA	108	3490	213GB	7713.59
LHCb Cloud Workers	96	1920	192GB	21121.71
IT Configuration Management Services	89	1990	176GB	15548.67
IT Agile CI	82	1640	164GB	20599.00
ATLAS Services Build	72	1720	141GB	3377.61

Projects

Project Name	VCPUs	Disk	RAM	VCPU Hours
IT Batch - Wigner	5264	151340	10TB	218398.13
IT Batch	1977	54430	3TB	109793.36
IT Plus	896	25760	1TB	55810.05
IT Batch - shared	768	22080	1TB	15452.86
NA61 Data production	198	0	99GB	101396.08
IT Dashboard	198	7010	394GB	34229.38
IT Monitoring	187	3740	374GB	34224.98
PH LCGAA	108	3490	213GB	7713.59
LHCb Cloud Workers	96	1920	192GB	21121.71
IT Configuration Management Services	89	1990	176GB	15548.67
IT Agile CI	82	1640	164GB	20599.00
ATLAS Services Build	72	1720	141GB	3377.61

User dashboard



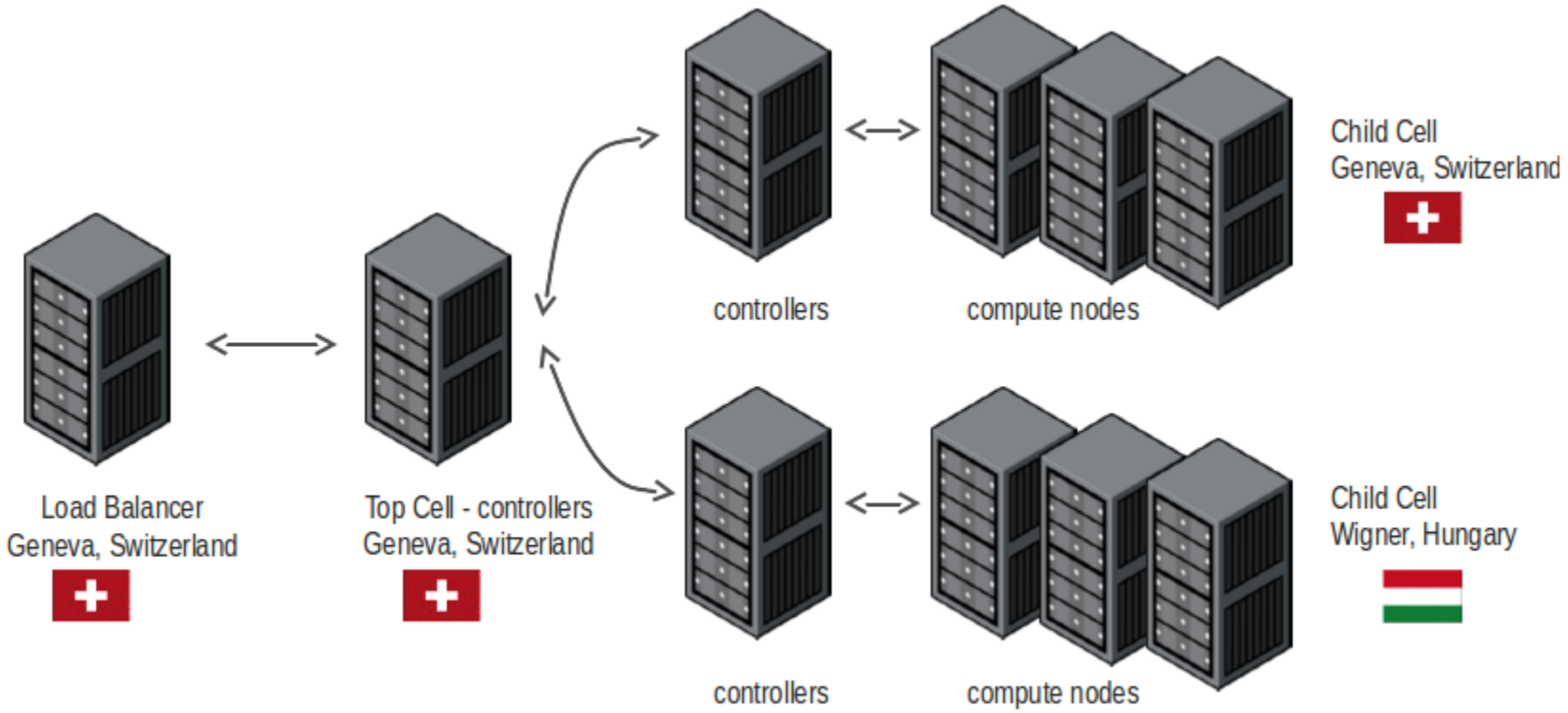
Nova cells

- Allow running the cloud in a distributed fashion and to scale.
- Introducing a second level of scheduling.
- Cells are configured in a tree. The top-level cell has a host that runs the nova-api service.

Future improvements:

- Live migration (Havana)
- Security groups (Havana with Neutron)
- Cinder volumes (Havana)

Nova cells


Load Balancer
Geneva, Switzerland
Top Cell - controllers
Geneva, Switzerland
controllers
compute nodes
Child Cell
Geneva, Switzerland
controllers
compute nodes
Child Cell
Wigner, Hungary

Nova compute nodes

KVM and Hyper-V compute nodes share the same infrastructure

- Choice of hypervisors should be tactical
  - Performance
  - Compatibility/Support with integration components
- Hypervisor selection is based on "Image" properties

Hyper-V status

Supporting "Windows on Windows" use case.

- Server consolidation activities

Hyper-V hypervisor are running on Grizzly:

- Fully puppetized
- No external patch needed on Grizzly

Hyper-V driver still lacks some functionality:

- console access
- resize support
- ephemeral disk support
- ceilometer metrics support

Glance

- Images for all CERN supported operating system.
  - Windows 7 / 2008
  - Scientific Linux 5 / 6
- User's images
  - Possibility to upload images.

Glance on ceph


```
default_store = rbd
rbd_store_user = images
rbd_store_pool = images
```

- authx for security (key lifecycle is a challenge as always)
- need librbd (from EPEL)
- issue found with default max processes on SLC6. (When OSD > 1024)

Glance images

CERN IT Department

Images & Snapshots

Logged in as: toulevey | Settings | Help | Submit a ticket | Sign Out

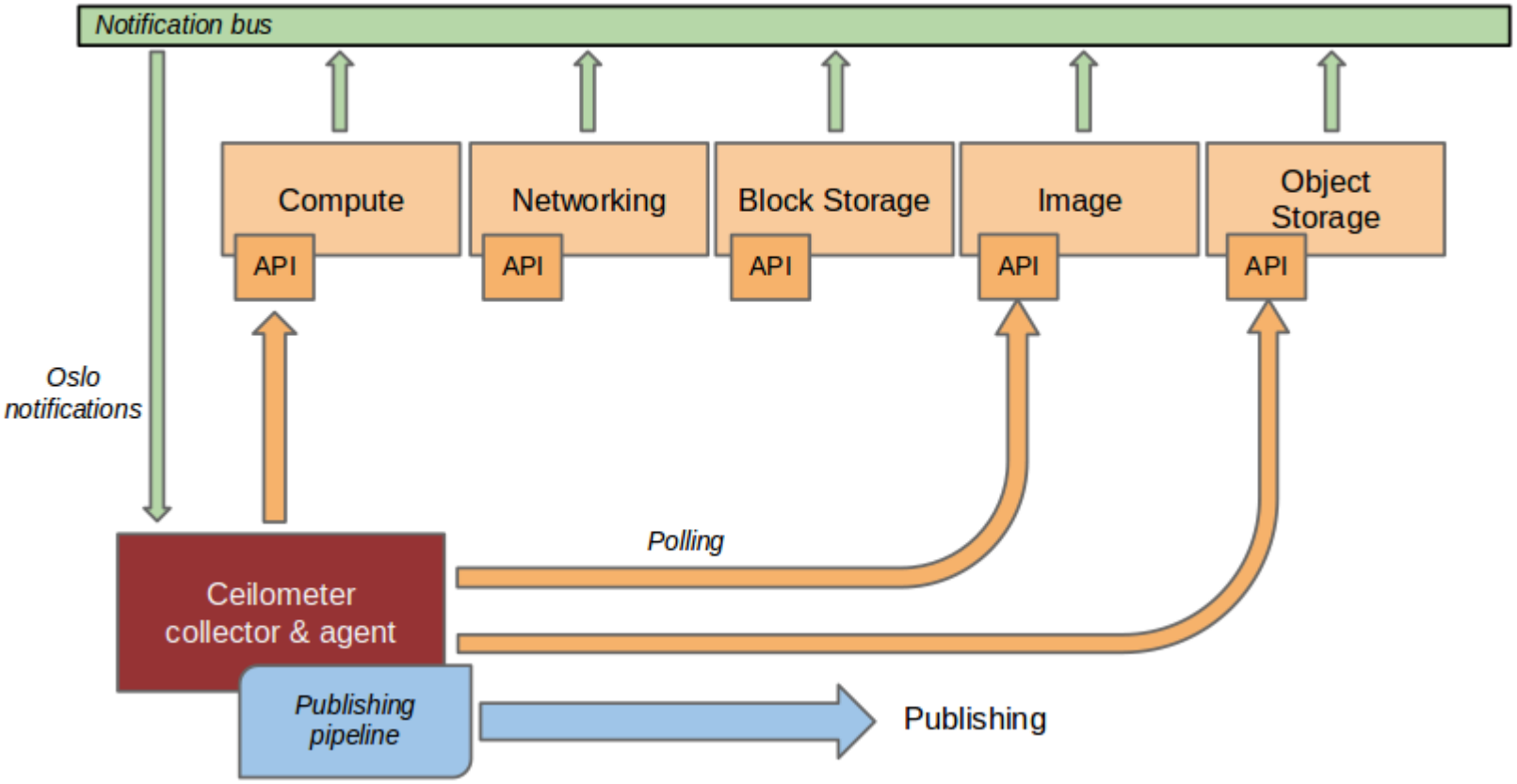
Images

Project (1) | Shared with Me (0) | Public (14) | + Create Image | Delete Images

Image Name	Status	Public	Format	Actions
Windows 7 - x64 [130924]	Active	Yes	VHD	Launch
SLC5 Server - i386 [130920]	Active	Yes	QCOW2	Launch
SLC6 Server - i386 [130920]	Active	Yes	QCOW2	Launch
SLC6 Server - x86_64 [130920]	Active	Yes	QCOW2	Launch
SLC5 CERN Server - i386 [130920]	Active	Yes	QCOW2	Launch
SLC5 Server - x86_64 [130920]	Active	Yes	QCOW2	Launch
SLC5 CERN Server - x86_64 [130920]	Active	Yes	QCOW2	Launch
SLC6 CERN Server - x86_64 [130920]	Active	Yes	QCOW2	Launch
SLC6 CERN Server - i386 [130920]	Active	Yes	QCOW2	Launch
Windows Server 2008 R2 - x64 [130904]	Active	Yes	VHD	Launch
SLC5 Server - i386 [130624]	Active	Yes	QCOW2	Launch
SLC5 Server - x86_64 [130624]	Active	Yes	QCOW2	Launch
SLC6 Server - x86_64 [130624]	Active	Yes	QCOW2	Launch
SLC6 Server - i386 [130624]	Active	Yes	QCOW2	Launch

Displaying 14 items

Ceilometer



- Ceilometer2ssm: http://cern.ch/go/W9Zw

Cinder volumes on Cinder

- 3 PB cluster available(10Gbps),
  - SLC6 with qemu-kvm patched by Inktank.
  - Grizzly Cinder with Cells support is missing, backport from Havana needed.
- Initial testing with FIO, 4GB files, libaio, bs 256k (1Gbps)

	numjobs=1	numjobs=4	numjobs=8
Rand RW	99MB/s	95MB/s	95MB/s
Rand R	103MB/s	107MB/s	100MB/s
Rand W	108MB/s	106MB/s	106MB/s

- Caching not enabled, available with qemu-kvm but not as an openstack volume option yet.
- Will be handed over to users soon

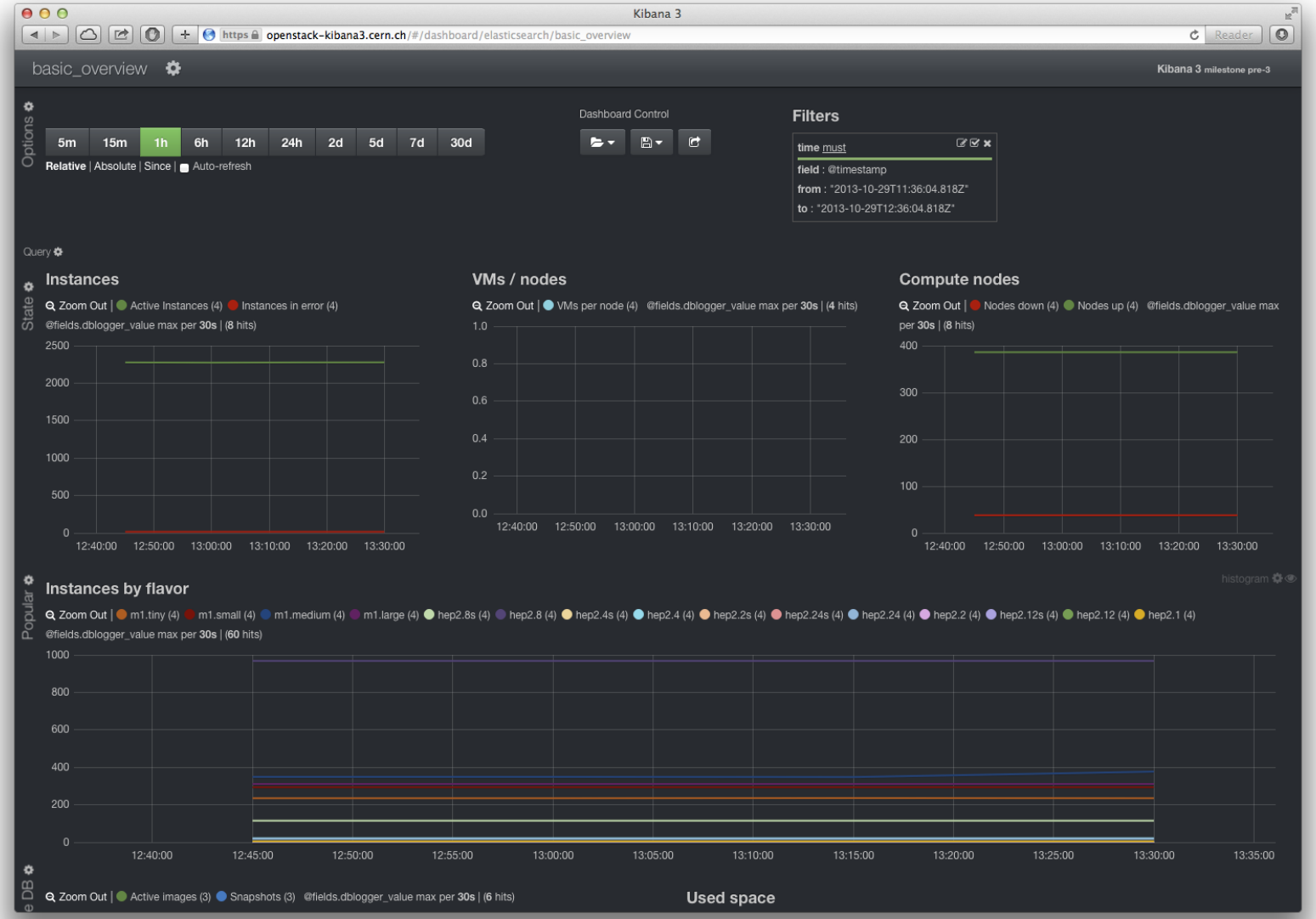
Monitoring: what ?

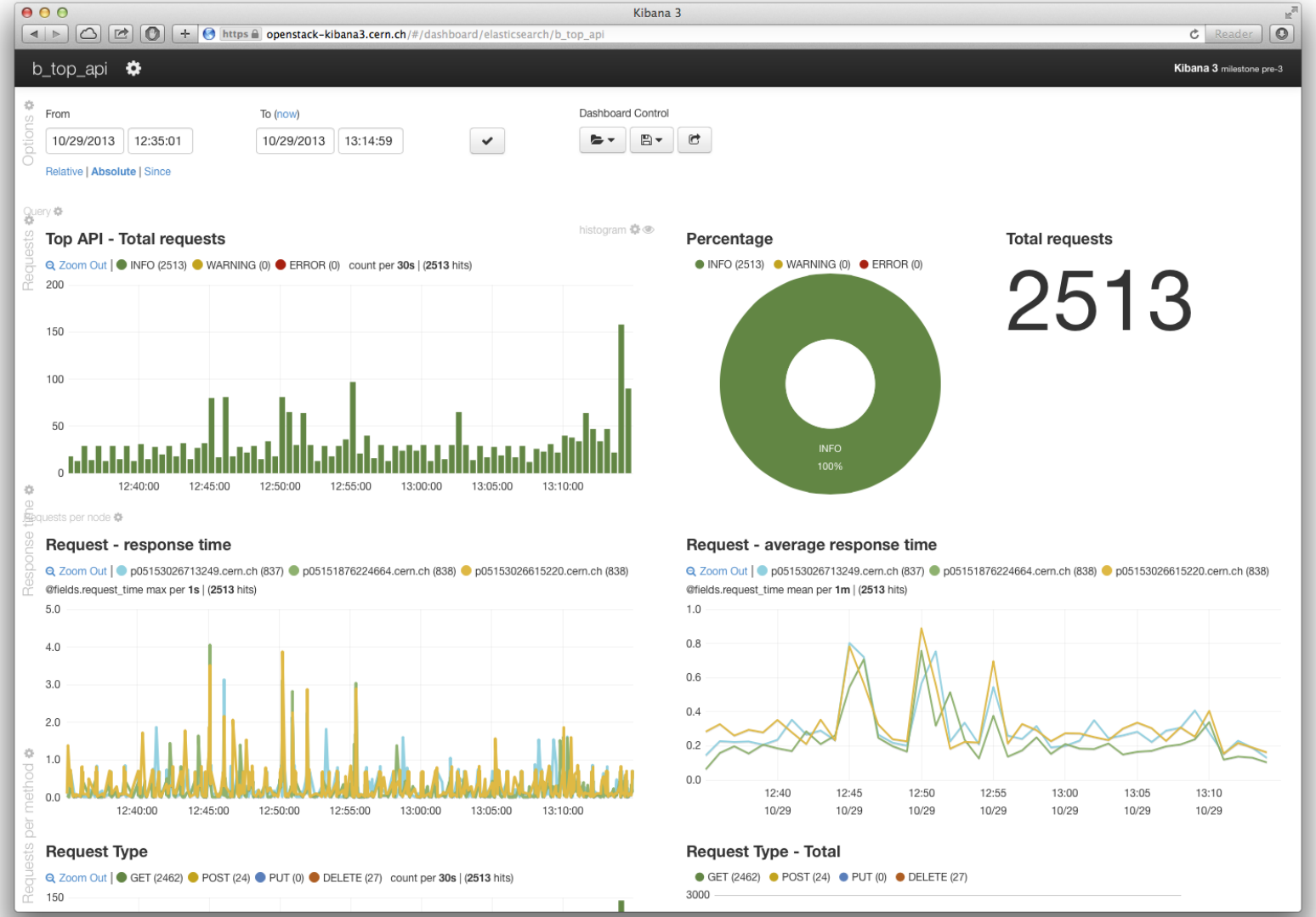
- ERRORs, WARNINGs – log visualization
- identify in “real time” possible problems
- preserve all logs for analytics
- visualization of cloud infrastructure status
  - service managers
  - resource managers
  - users

Tools: Apache Flume, Elastic search, Kibana

Monitoring: Kibana

CERN IT Department




Kibana 3
https openstack-kibana3.cern.ch/#/dashboard/elasticsearch/b_top_api
Reader
b_top_api
Kibana 3 milestone pre-3
Options
From
10/29/2013
12:35:01
To (now)
10/29/2013
13:14:59
Relative | Absolute | Since
Dashboard Control
Query
Requests
histogram
Top API - Total requests
Zoom Out | INFO (2513) WARNING (0) ERROR (0) count per 30s | (2513 hits)
200
150
100
50
0
12:40:00
12:45:00
12:50:00
12:55:00
13:00:00
13:05:00
13:10:00
Percentage
INFO (2513) WARNING (0) ERROR (0)
INFO
100%
Total requests
2513
Response time
requests per node
Request - response time
Zoom Out | p05153026713249.cern.ch (837) p05151876224664.cern.ch (838) p05153026615220.cern.ch (838)
@fields.request_time max per 1s | (2513 hits)
5.0
4.0
3.0
2.0
1.0
0.0
12:40:00
12:45:00
12:50:00
12:55:00
13:00:00
13:05:00
13:10:00
Request - average response time
Zoom Out | p05153026713249.cern.ch (837) p05151876224664.cern.ch (838) p05153026615220.cern.ch (838)
@fields.request_time mean per 1m | (2513 hits)
1.0
0.8
0.6
0.4
0.2
0.0
12:40 10/29
12:45 10/29
12:50 10/29
12:55 10/29
13:00 10/29
13:05 10/29
13:10 10/29
Requests per method
Request Type
Zoom Out | GET (2462) POST (24) PUT (0) DELETE (27) count per 30s | (2513 hits)
150
Request Type - Total
GET (2462) POST (24) PUT (0) DELETE (27)
3000

EC2 and nova-api

EC2

- CMS
- Atlas
- NA61

Nova-api

- Puppet managed VMs in the datacentre: ai-bs-vm, ai-kill-vm, etc...

Students projects

Monica Talach - Monitoring Dashboard investigation

Gary Mcgilvary - Cinder with Ceph and Netapp

Davide Michelino - Implementation and test of OpenStack Heat

Andrea Giardini - DBaaS with Openstack Trove

Slides : http://cern.ch/go/tJ7x

Next challenges @ CERN

- Growing infrastructure,
  - Adding 100 Hypervisor/week
- In-place migration from Grizzly to Havana,
- Nova-network deprecation,
- External authentication Kerberos, X509,
- Keystone domains,
  - Delegate management to communities/experiments.
- Federated Cloud.
  - A reference architecture for federation of OpenStack clouds.

Federating Openstack

- A CERN/Rackspace Collaboration
  - 1 year of joint research,
  - blueprint and code contributions to the open source communities.
- Presentations and white papers to allow others to build on our findings.

Federating Openstack

- Deploy a Rackspace private cloud at CERN in parallel with the CERN Private cloud,
- Investigate OpenStack cloud federation in areas such as Authentication, Images, Networking and Metering,
- Demonstrate burst workload from private clouds to Rackspace public cloud.

QUESTIONS ?

HTTPS://OPENSTACK.CERN.CH

AND ASSOCIATED DOC

HTTP://CERN.CH/GO/LWN8

Thank you !