

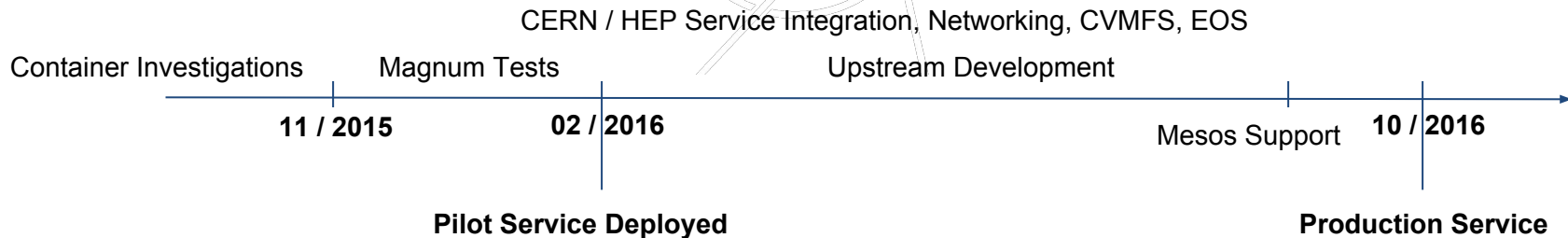
OpenStack Magnum & CVMFS

Ricardo Rocha

(on behalf of the CERN cloud team)

Goals and Timeline

- Integrate containers in the CERN cloud
 - Shared identity, networking integration, storage access, ...
- Agnostic to container orchestration engines
 - Docker Swarm, Kubernetes, DCOS/Mesos
- Fast, easy to use

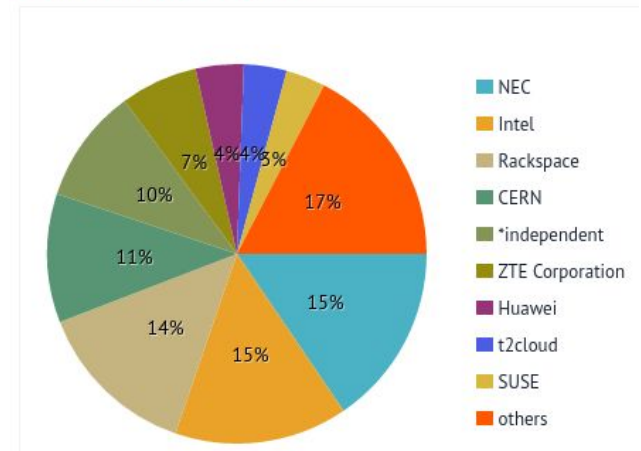


OpenStack Magnum

#openstack-containers

- Kubernetes, Docker Swarm, Apache Mesos, DC/OS (experimental)aaS
- Deep integration of OpenStack with Container technologies:
 - Compute Instances
 - Networks, Load Balancers
 - Storage
 - Security
 - Native Container API
 - Lifecycle cluster operations
 - Scale cluster up and down
 - WIP (Upgrades, Healing)

Contribution by companies



OpenStack Magnum Overview

<http://clouddocs.web.cern.ch/clouddocs/containers/quickstart.html>

- Clusters are described by *cluster templates*
- Shared/public templates for most common setups, customizable by users
 - COE and container runtime versions (kubernetes, docker, etc)
 - Integrated monitoring
 - Filesystem access
 -

```
$ openstack coe cluster template list
+-----+-----+
| uuid | name |
+-----+-----+
| .... | swarm |
| .... | swarm-preview |
| .... | kubernetes |
| .... | kubernetes-preview |
| .... | dcos |
| .... | dcos-preview |
+-----+-----+
```

```
$ openstack coe cluster template show kubernetes
...
| coe | kubernetes |
| master_flavor_id | m1.medium |
| flavor_id | m1.medium |
| server_type | vm |
| image_id | fedora-atomic-26 |
| labels | {"kube_tag": "1.9.0"} |
| network_driver | flannel |
```

OpenStack Magnum Overview (2)

- Create a cluster in a single command (no matter what size)

```
$ openstack coe cluster create --name mykub --cluster-template kubernetes-preview --node-count 10

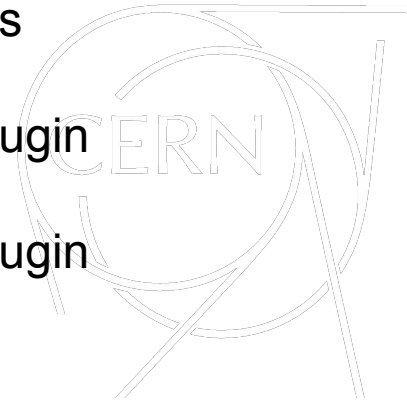
$ openstack coe cluster list
+-----+-----+-----+-----+-----+
| uuid | name          | node_count | master_count | status          |
+-----+-----+-----+-----+-----+
| .... | mykub        | 10         | 1             | CREATE_COMPLETE |
+-----+-----+-----+-----+-----+

$ $(openstack coe cluster config mykub)

$ kubectl get pod
```

Storage Integration

- OpenStack Cinder and CephFS
 - Built in kubernetes
- CVMFS
 - Docker volume plugin
- EOS
 - Docker volume plugin



docker-volume-cvmfs

- Access to CVMFS for running containers
- Docker volume driver
 - Runs as an additional container
- Kubernetes integration via FlexVolume

```
$ docker volume create -d cvmfs --name=cms.cern.ch
```

```
$ docker run -it --rm --volume-driver cvmfs -v cms.cern.ch:/cvmfs/cms.cern.ch centos:7 /bin/bash  
[root@874cbf8199d0 /]# ls /cvmfs/cms.cern.ch  
CMS@Home bootstrap_slc5_amd64_gcc462.log cmssw.git
```

<https://gitlab.cern.ch/cloud-infrastructure/docker-volume-cvmfs>

docker-volume-cvmfs

```
$ cat nginx-cvmfs.yaml
...
spec:
  containers:
  - name: nginx-cvmfs
    image: nginx
    ports:
    - containerPort: 80
    volumeMounts:
    - name: lhcb
      mountPath: /cvmfs/lhcb.cern.ch
  volumes:
  - name: lhcb
    flexVolume:
      driver: "cern/cvmfs"
      options:
        repository: "lhcb.cern.ch"

$ kubectl create -f nginx-cvmfs.yaml
```


Speeding up access to images: CVMFS graphdriver

- Problem: some users have large container images
- We recently integrated the docker cvmfs graphdriver from CernVM
- Available as a flag in the Magnum templates
- Hope was to increase significantly the container startup time
 - File based instead of Layer based access

```
$ openstack coe cluster template show swarm-docker-cvmfs  
| labels | {"cvmfs_storage_driver": "true"}
```

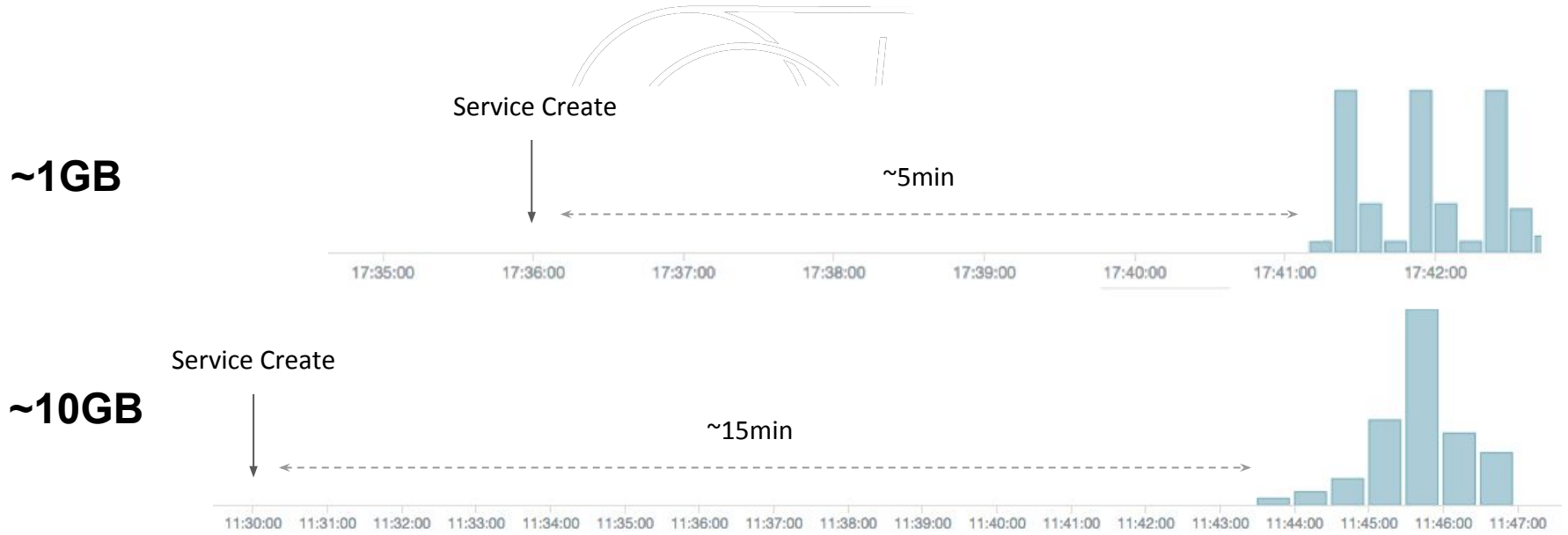
- “CernVM-FS Graph Driver Plugin for Docker”, yesterday by Nikola

CVMFS Graph Driver - Local Tests

- Fresh Swarm Cluster of ~100 nodes
 - VMs on our production cloud
- Performance Benchmark Suite
 - docker service create --mode global ...
 - 100 replicas pulling from the same registry at ~same time
 - Medium sized image of ~1GB and very large image of ~10GB

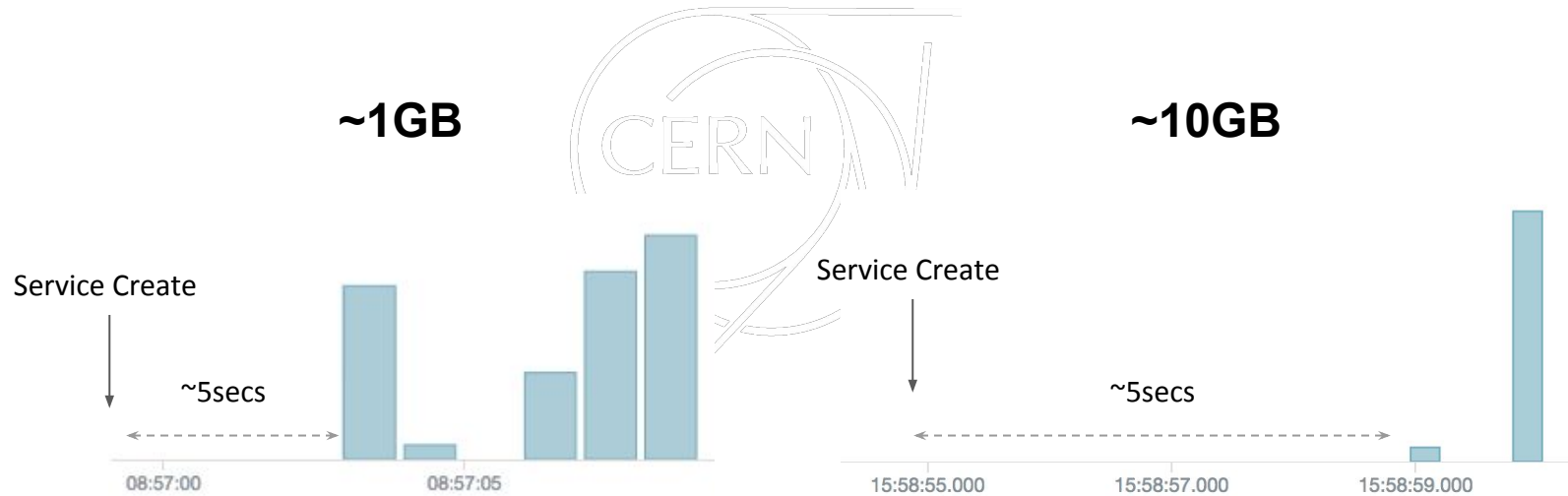
CVMFS Graph Driver - Local Tests

- Regular image results (no CVMFS)



CVMFS Graph Driver - Local Tests

- CVMFS aware thin Image



CVMFS Graph Driver - External Tests

- Access to external cloud resources (RHEA, Exoscale, OTC)
 - Via HNScienceCloud
- Still fetching images from CERN's gitlab registry
 - Higher latency
 - Better cache usage (local squid cache on each cloud)
 - Tests done for 10GB images

CVMFS G

- Access
 - V
- Still fe
 - H
 - B
 - T

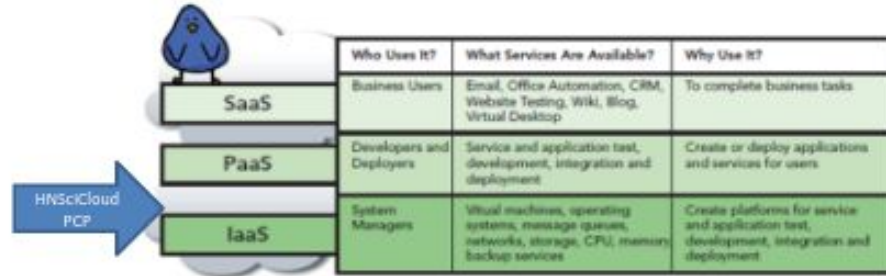
What is being procured



A hybrid cloud platform for the European research community



Combining services at the IaaS level to support science workflows

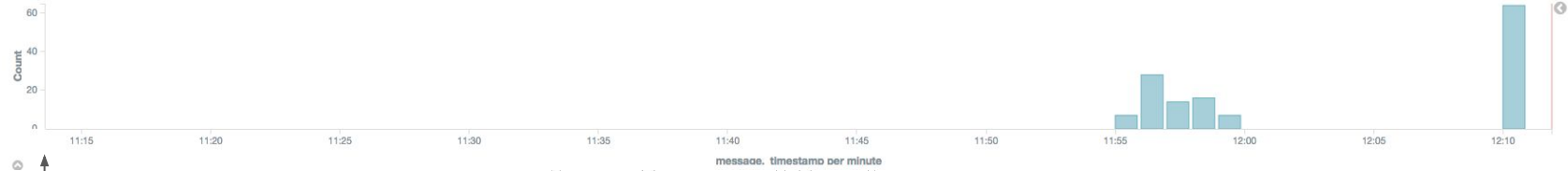


Source: Cloud Computing for Science, DST Solutions, David Bunker, Vitor Marau and Caron Bentley

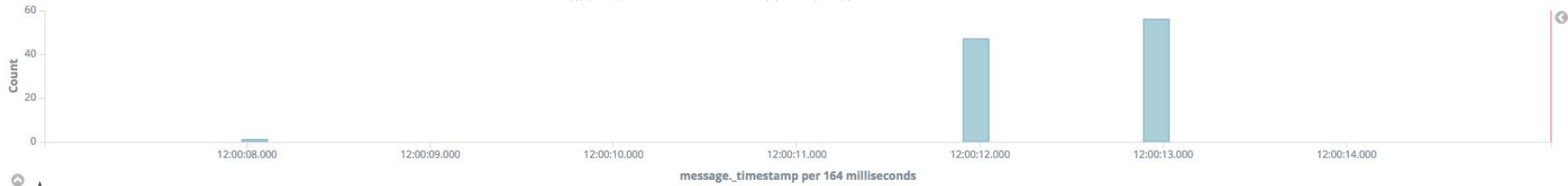
The R&D services to be developed are to be integrated with Resources in data centres operated by the Buyers Group,

GEANT GEANT network and eduGAIN fed. identity mgmt eduGAIN

CVMFS Graph Driver - External Tests



Service Create

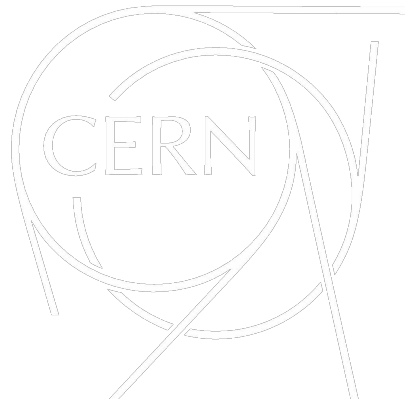


Service Create

Future Work

- OpenStack Magnum improvements
 - Upgrades, Auto-healing
 - Support for other container runtimes (containerd, cri-o)
 - Federation Support
- Adapt the plugin to CSI (Container Storage Interface)
- CVMFS graphdriver integration with containerd
 - Snapshot Driver?
- This was fun to integrate and test!
 - Many thanks to Jakob and Nikola

Backup Slides

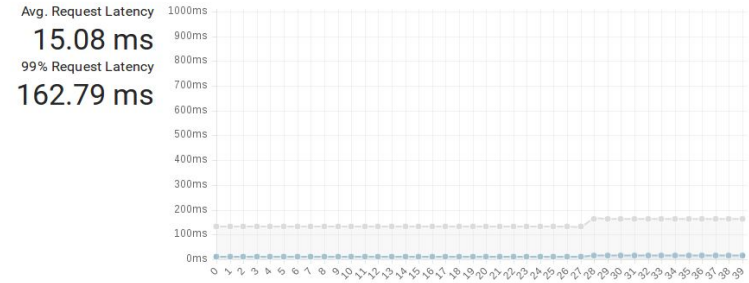
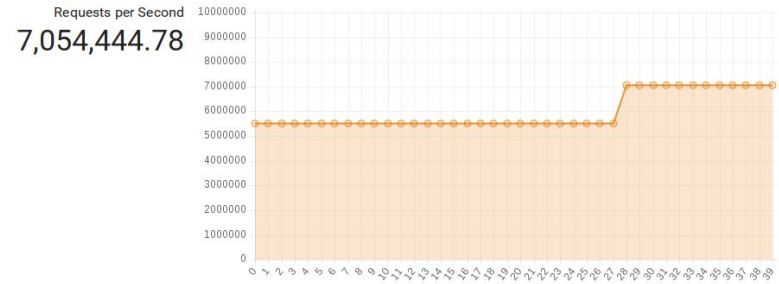


Performance Evaluation

- August 2016
 - Kubernetes stress test
 - Managed **7 million requests / sec**
- And an analysis of cluster deployments

Cluster Size (Nodes)	Deployment Time (min)
2	2.5
32	4
128	5.5
512	14
1000	23

Kubernetes 1M 10M Reqs/Second



Server Availability 100% # Servers 500 # Loadbots 9,449