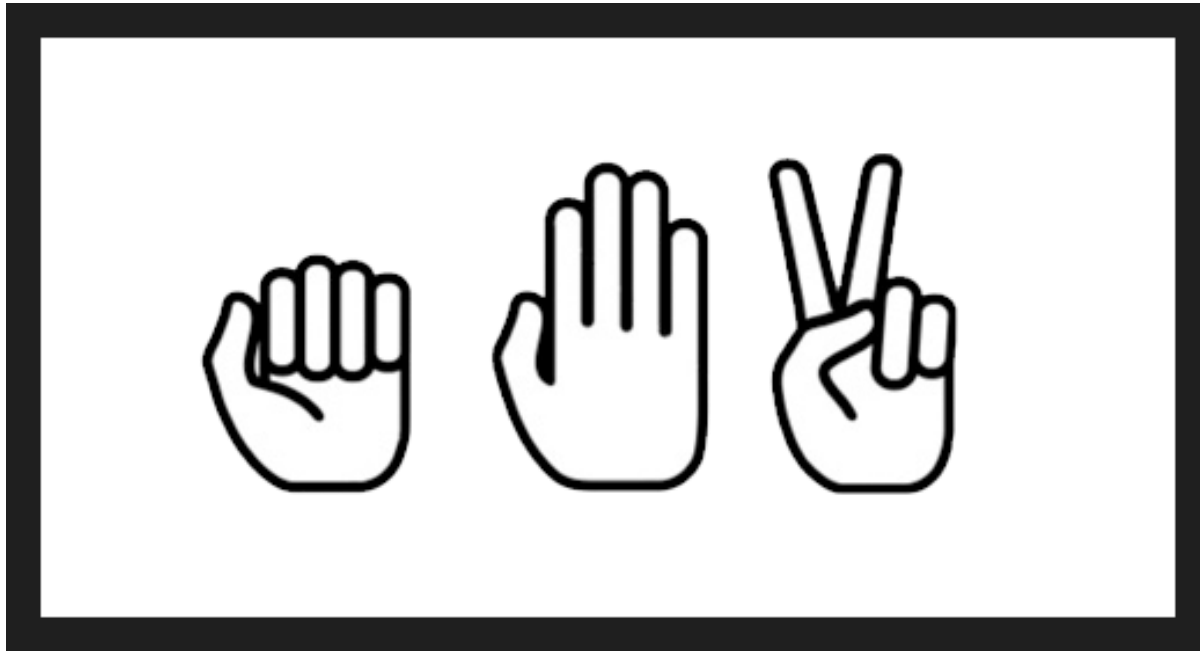


Resource Provisioning Services

Introduction and Plans



Who are we?

BRITO DA ROCHA, Ricardo Manuel
CASTRO LEON, Jose
DENIS, Marek Kamil
DOMINGUES CORDEIRO, Cristovao Jose
DOS SANTOS BOMPASTOR, Bruno Luis
FERMIN LOBO, Marcos
FERNANDEZ RODRIGUEZ, Daniel
GIORDANO, Domenico
MICHELINO, Davide
NOEL, Bertrand
PIGUEIRAS ARECES, Luis Oscar
RODRIGUES MOREIRA, Belmiro Daniel
TRIGAZIS, Spyridon
VAN ELDIK, Jan
VELTEN, Mathieu
WIEBALCK, Arne
ZILLI, Stefano

Plus collaborators

- BARC
- Master students
- Summer students
- “Visitors”

Section Mandate

“Provide Compute Services for CERN Tier-0 and WLCG”

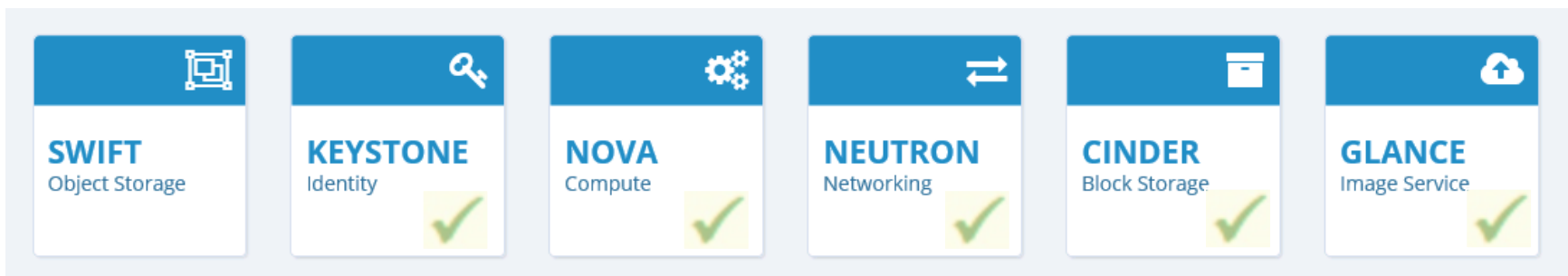
- CERN Private Cloud
- Extension into public clouds

OpenStack












OpenStack software controls large pools of compute, storage, and networking resources throughout a datacenter, managed through a dashboard or via the OpenStack API. OpenStack works with popular enterprise and open source technologies making it ideal for heterogeneous infrastructure.

-- *OpenStack*
<http://www.openstack.org>



Many optional services

 HORIZON Dashboard 	 CEILOMETER Telemetry 	 HEAT Orchestration 
 TROVE Database	 SAHARA Elastic Map Reduce	 IRONIC Bare-Metal Provisioning 
 ZAQAR Messaging Service	 MANILA Shared Filesystems 	 DESIGNATE DNS Service
 BARBICAN Key Management	 MAGNUM Containers 	 MURANO Application Catalog 
 CONGRESS Governance		

...AND MANY MORE

Deploying OpenStack at CERN

- Configuration infrastructure based on Puppet
- Community Puppet modules for OpenStack
- RDO - RPM Packages
 - Currently on Kilo release
- 5500 hypervisors
 - 4600 CC7
 - 180 Microsoft Windows 2012 R2
 - 50 RHEL6/7
 - 700 SLC6 -> being phased out

Usage

- ~2000 CERN users, 400+ shared projects
 - LCG experiments, IT services, development boxes, ...
- ~15000 Virtual Machines
 - Creation/deletion rate of ~200 VMs/hour
- Different workloads
 - Compute intensive, I/O intensive
 - Compute nodes are optimized for these workloads

Service additions during 2015

- Introduced OpenStack Heat
 - Cloud Orchestration
- Cinder Volumes
 - NetApp for Hyper-V
 - Ceph @ Wigner
- Keystone endpoint filtering
 - Allows for selective adding of services (Heat, S3)

Automation



- Automated service interventions, log the results
- Interacts with ServiceNow, OpenStack, ActiveDirectory, etc.
- Allows sharing of tasks with other teams without exposing credentials or procedures:
 - **SysAdmins and hardware repair teams:**
Workflows related to hypervisor maintenance (h/w intervention, notify users...)
 - **Cloud-Operations:** Project creation, Health reports, Quota update

Ongoing activities

Scaling the service

- A lot of box shuffling
 - Today: ~5500 compute nodes (145K cores)
 - Underway: 25K cores
 - Spring 2016: 60K cores
 - Retirements: being clarified
- Developing a VM migration framework
 - (block) live-migration where possible, cold migration where necessary
 - Automated planning & execution!

CPU Performance optimizations

- Developed an optimized configuration
 - NUMA-aware VM scheduling, 2MB huge pages
 - See Hepix talk for details
- Virtualization overhead pushed below 5%
 - For full-node VMs!
- Configuration successfully field-tested in Q4 2015
- Rollout has now started
 - ~2000 compute nodes to upgrade in coming months
 - In close collaboration with the Batch team

Metering, Accounting, Auditing

- Start using our Ceilometer data
 - Data analytics
 - Gnocchi – time series of samples
- Enable CADF
 - Cloud Accounting Data Federation
 - 7 “W”s of audit and compliance
 - Uses Ceilometer events
- Produce per-domain usage reports

CERN Cloud Procurement Roadmap

- ☑ First Procurement, March '15
 - Target a single VO, run **simulation** jobs, up to 3,000 vCPUs
- ☑ Second Procurement
 - Production activity currently running (**Nov.-Dec. '15**), Target **multi VOs**, simulation jobs. ~2,000 vCPUs
- ☑ Microsoft Azure IaaS evaluation
 - Granted credit for usage (**Apr.-Nov. '15**), up to 4,600 vCPUs in 3 Data Centres
 - Similar workloads as in Second Procurement
- ☐ IBM Bare metal server Evaluation
 - Granted credit for next 5 months (until **May '16**)
- ☐ Third Procurement
 - Production activity to start during Spring '16, Target multi VOs, **full chain** processing
 - Procurement phase is ongoing
 - Crucial components: Cloud Storage and Network
- Areas of work: Capacity Management, Cloud Monitoring, Benchmarking, Accounting

Neutron

- Deprecation of nova-network progressing slowly...
 - our Neutron deployment reflects this 😊
- Plan:
 - Deploy a compute cell w/ Neutron
 - Get experience
 - Develop migration plan for existing cells
 - Execute it.
- Note: initially we aim for feature parity. Later, we may investigate enhancements (like per-tenant networks)

CVI phase-out

- Phase-out strategy:
 - Help users to **recreate** VMs on OpenStack
 - Migrate VMs to OpenStack **where appropriate**
- Actively working with different customers:
 - 130 developer VMs managed BE/CO
 - 110 IT service and development VMs
 - 80 Engineering VMs managed
 - Contacting end-users for 250 “custom” VMs
- Aim: close down the service by mid-2016

Domains

- “Top-level” for projects
 - With delegation of administration privileges
- Introduce gently into our cloud
 - First: Indigo Datacloud (federated science cloud)
 - Set up Edugain Federation
 - Sort out access to resources, ownership, etc.
- Nested quota are coming along
 - Cinder: has landed in Liberty
 - Nova: should land in Mitaka

Container orchestration

- Use cases developing rapidly
 - Root-as-a-Service, EN/ICE, OpenShift, Indigo DataCloud
 - Requests for Kubernetes for orchestration
- Investigations into OpenStack Magnum are well underway
 - Scale-out using Heat
 - Deployment scenario's
 - Introducing OpenStack Neutron network component 😊
 - Atomic hosts
- Pre-prod service in Q1

Image management

- Working with Linux Support team to build Linux VM and container images using Koji
- New Glance feature “community images” allow for easier image sharing
- Could review creation and support of official images
- Similar for Windows images

Other work items

- OpenStack EC2API
 - Replaces Nova EC2 driver
 - Better implementation, better support
- Possible service extensions
 - Ironic – Bare Metal provisioning
 - Murano – Application Catalog
 - Manilla – Shared Filesystems

Major work items for this year

- Integrating Private and Public cloud services
- Operating and evolving these services
- Improvements in monitoring, benchmarking, accounting
- Introduce exciting new features 😊

Crystal clear? Cloudy?

