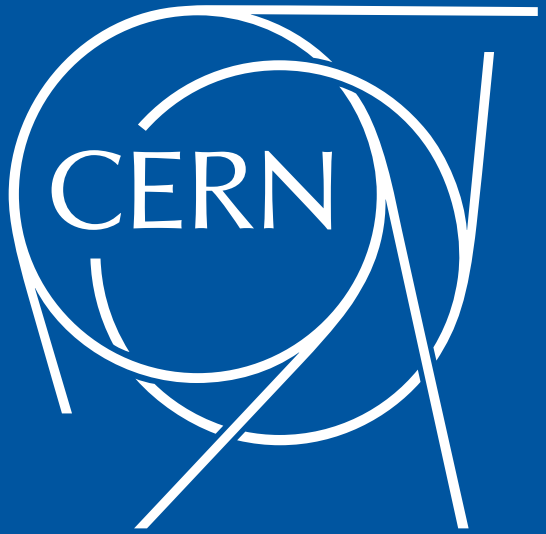
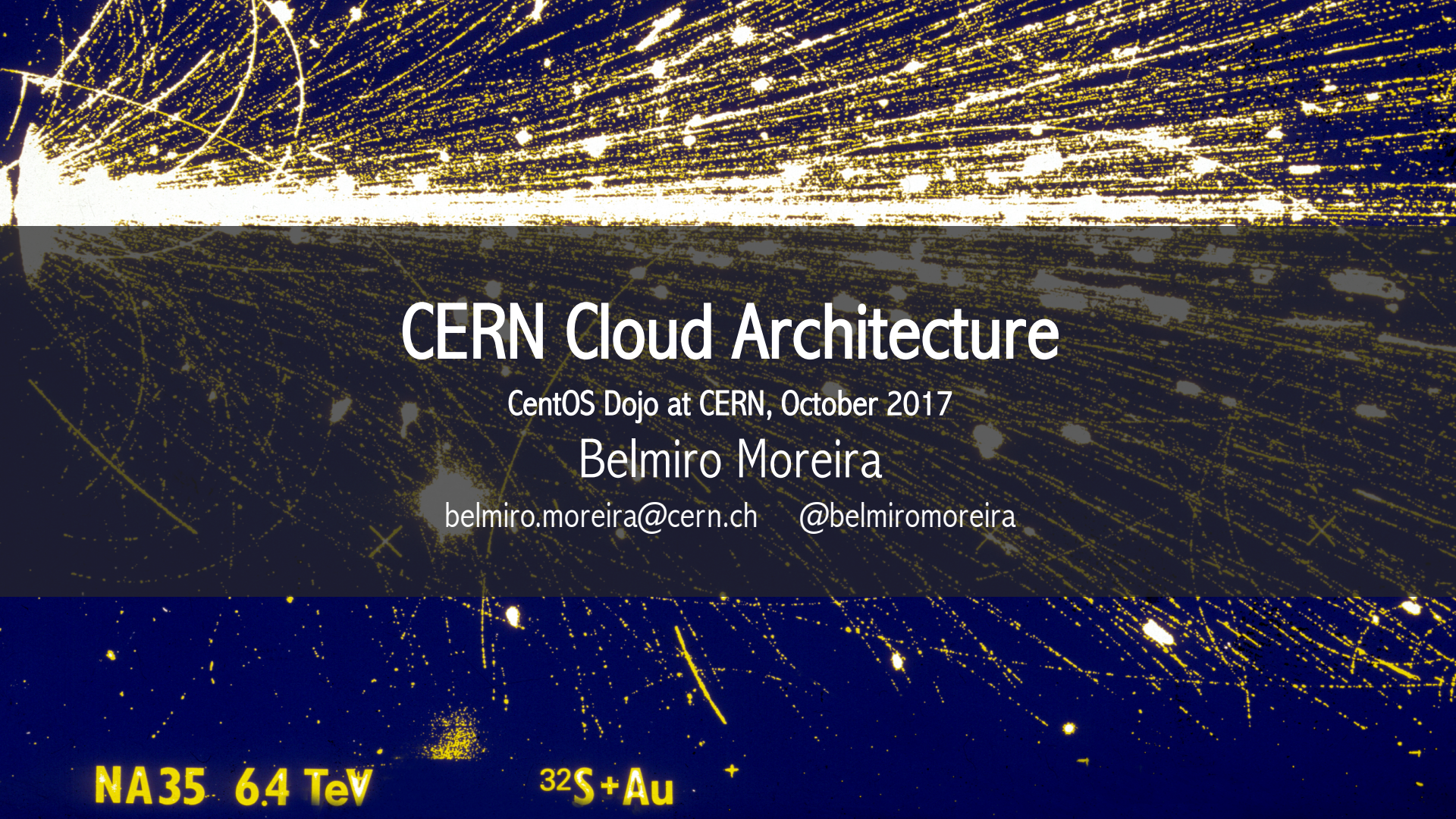

CERN



CERN Cloud Architecture

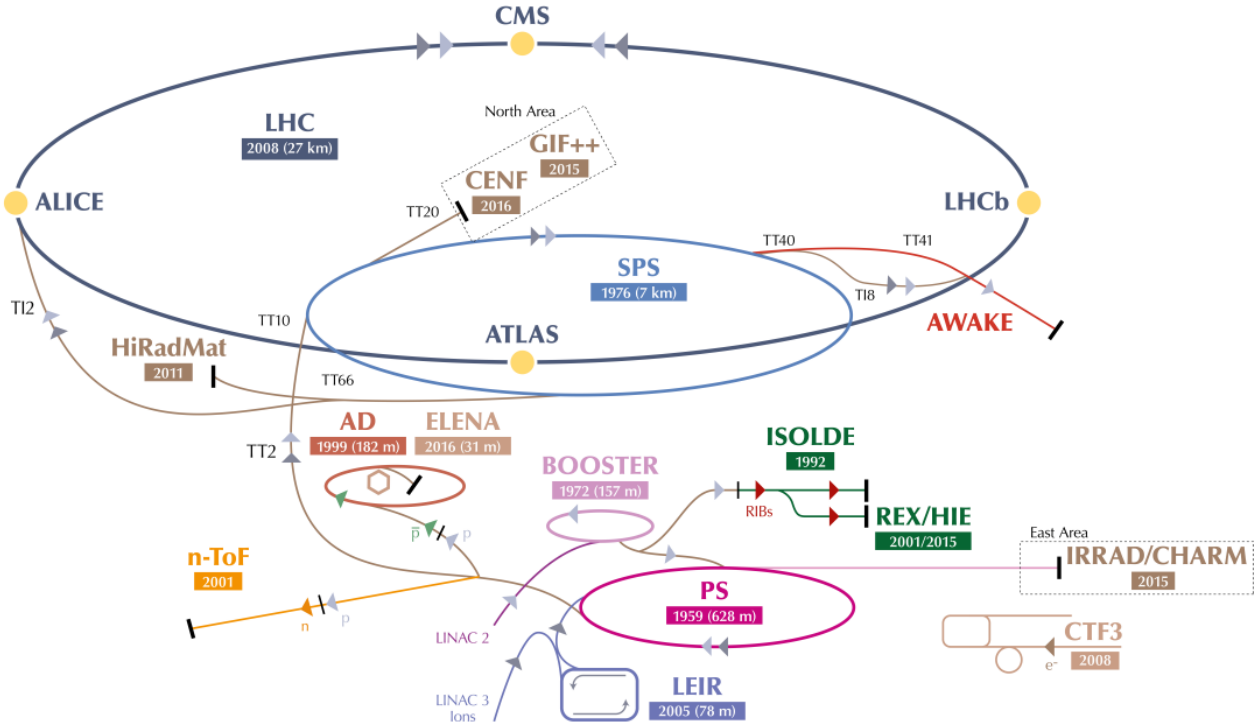
CentOS Dojo at CERN, October 2017

Belmiro Moreira

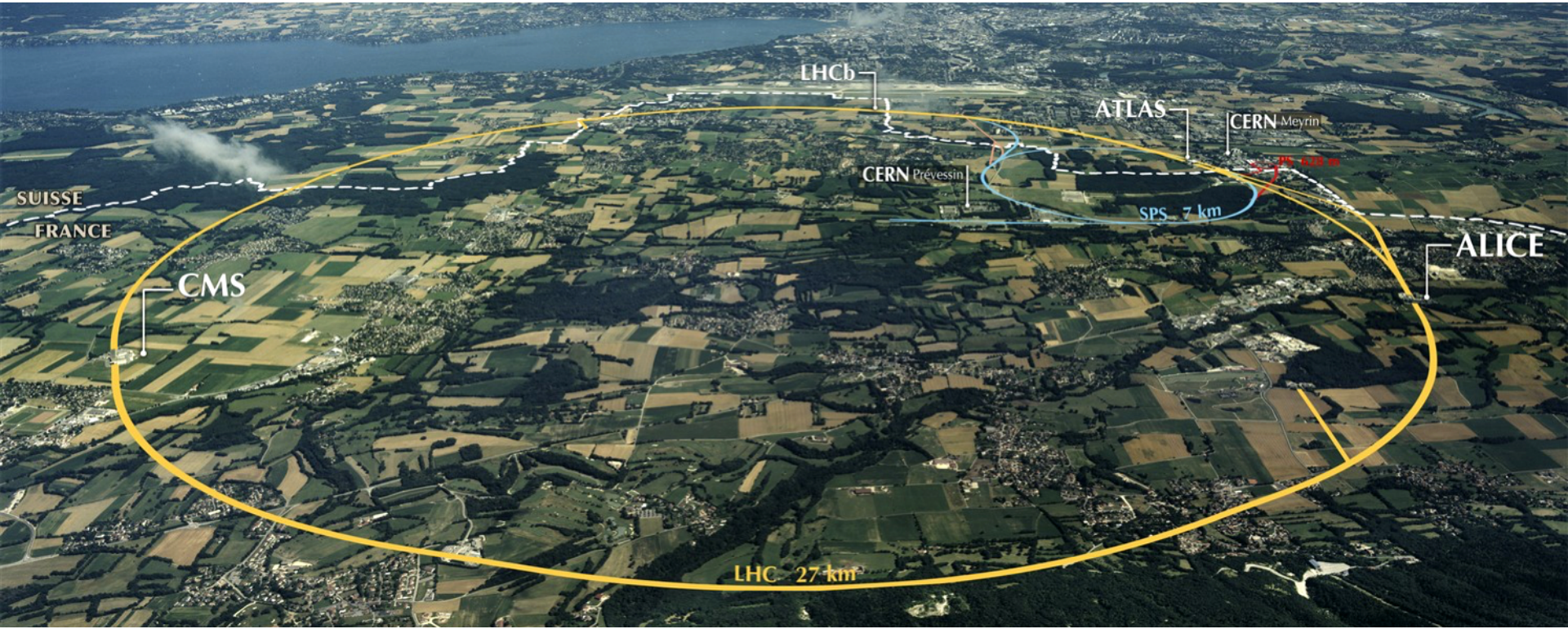
belmiro.moreira@cern.ch @belmiromoreira

What's CERN?

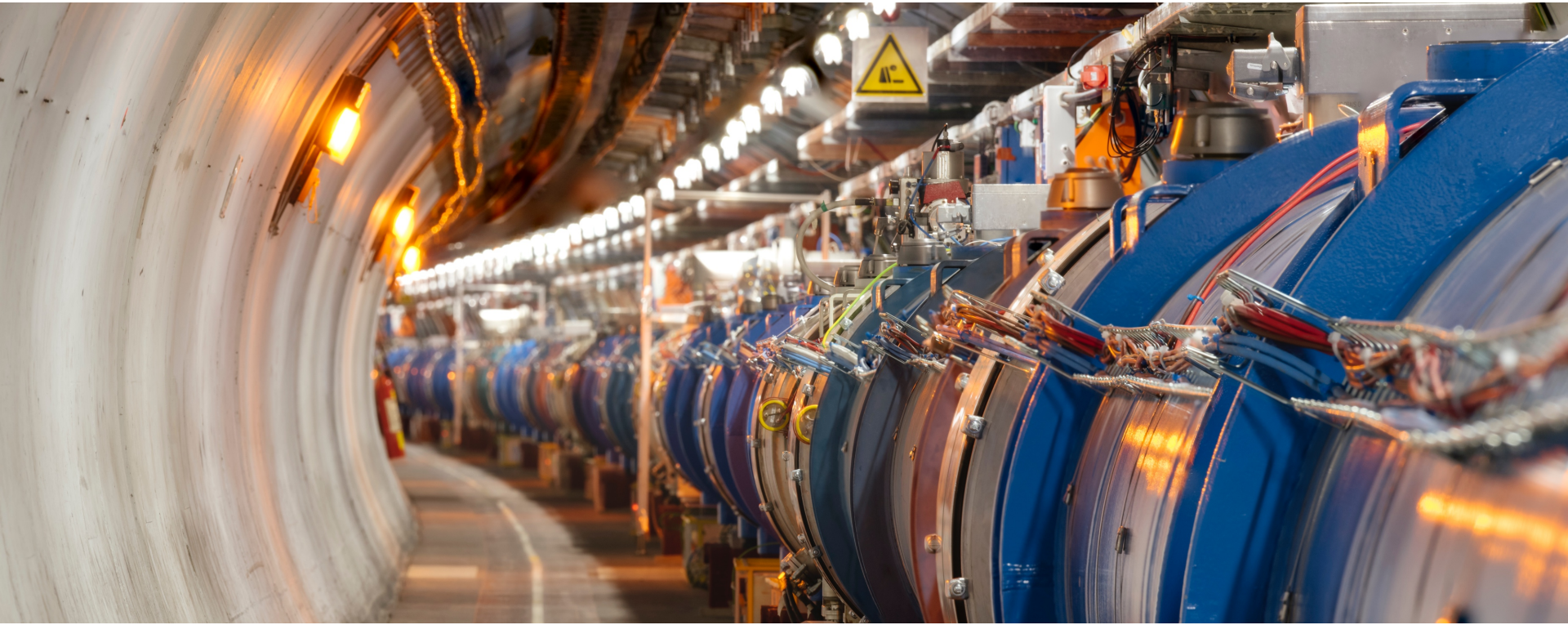
CERN - Accelerator complex



CERN - LHC and Experiments

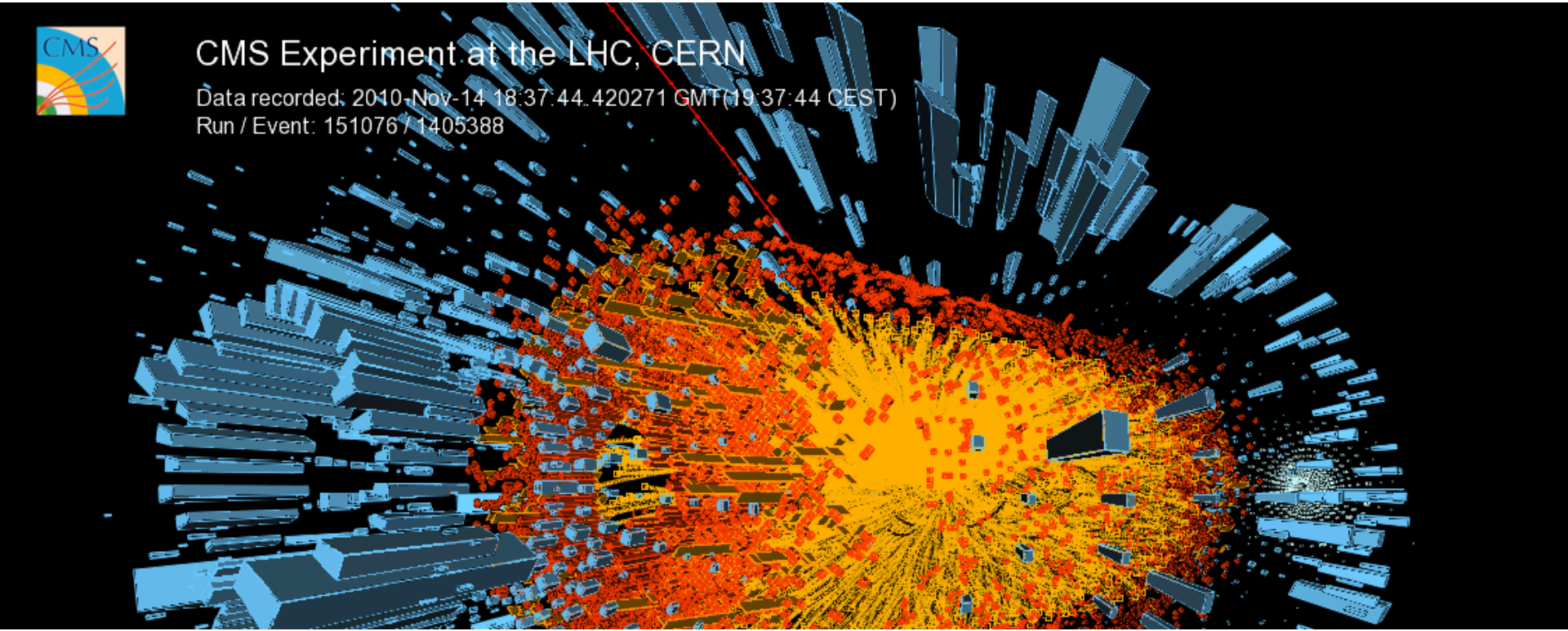


CERN - LHC and Experiments

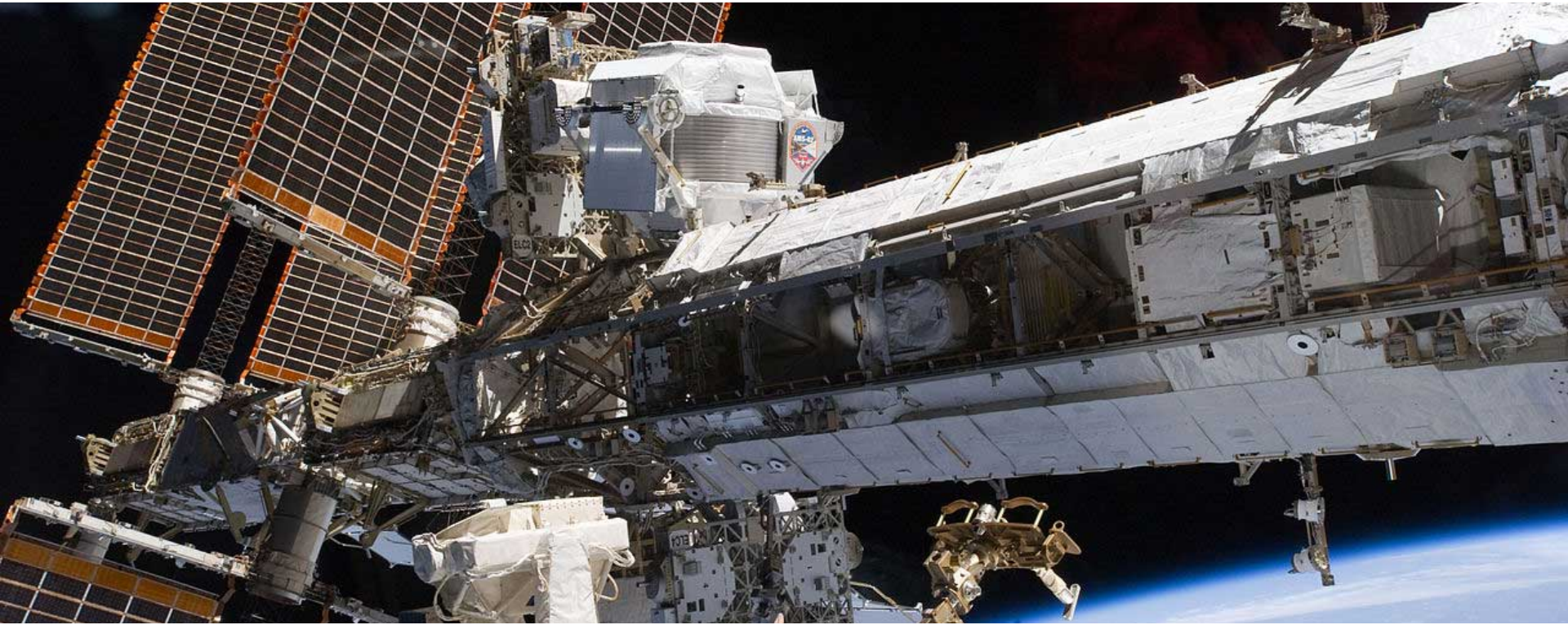


CERN - LHC and Experiments

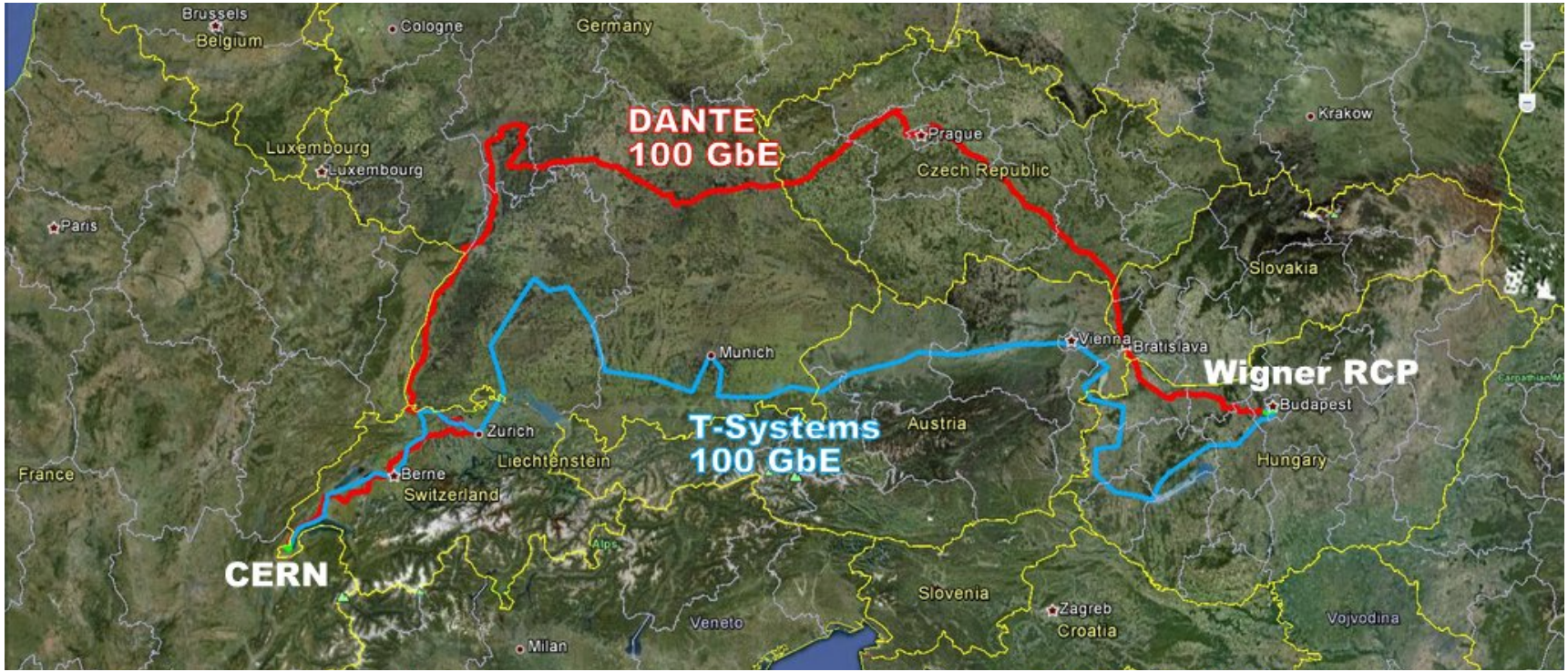
CERN - LHC and Experiments



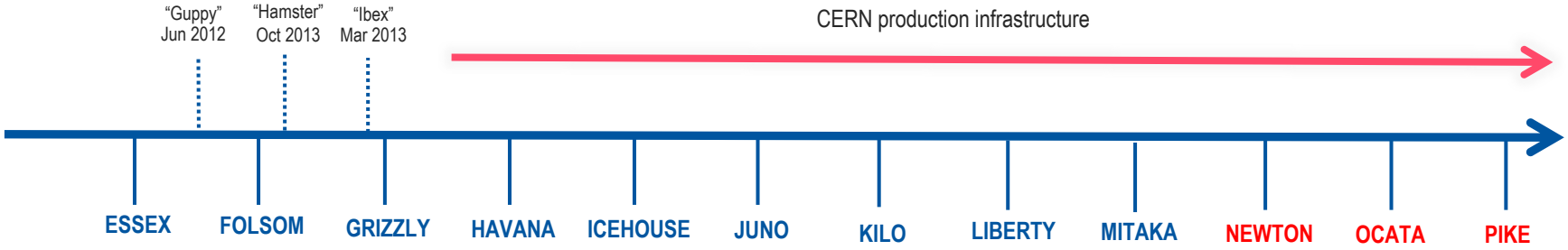
CERN - AMS



CERN Data Centres



OpenStack timeline at CERN



OpenStack Projects at CERN Cloud

IRONIC
an OpenStack Community Project

OSLO
an OpenStack Community Project

GLANCE
an OpenStack Community Project

NOVA
an OpenStack Community Project

NEUTRON
an OpenStack Community Project

EC2-API
an OpenStack Community Project

HORIZON
an OpenStack Community Project

RALLY
an OpenStack Community Project

OCTAVIA
an OpenStack Community Project

CINDER
an OpenStack Community Project

PUPPET OPENSTACK
an OpenStack Community Project

MAGNUM
an OpenStack Community Project

HEAT
an OpenStack Community Project

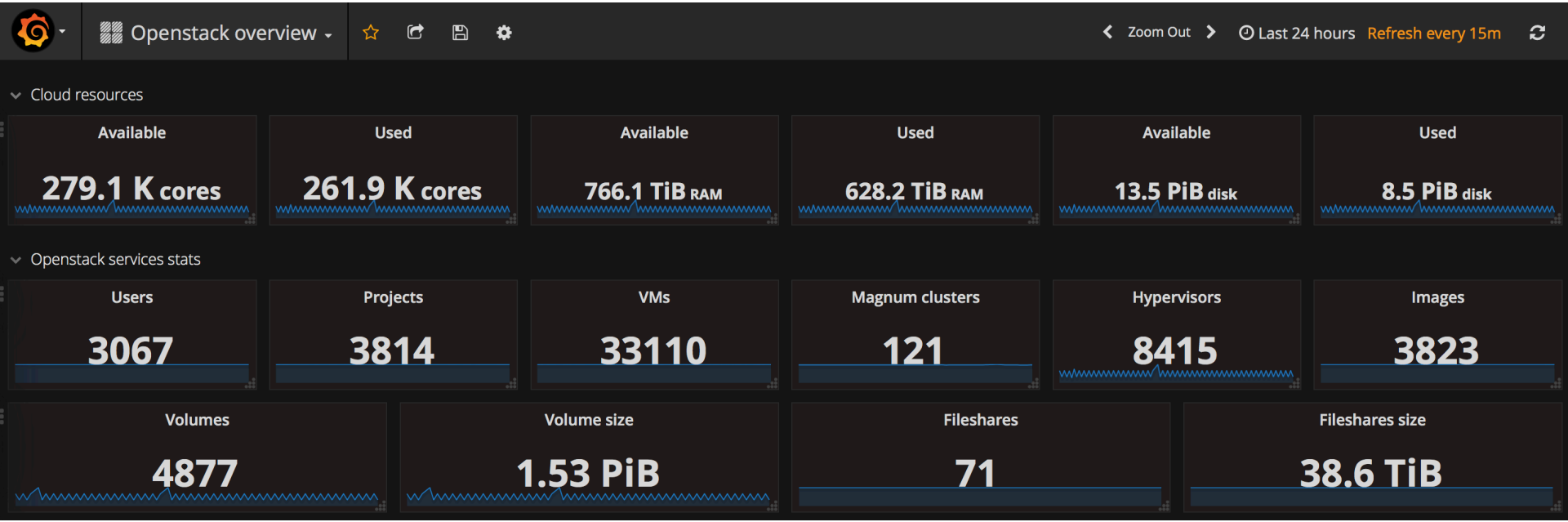
MANILA
an OpenStack Community Project

BARBICAN
an OpenStack Community Project

KEYSTONE
an OpenStack Community Project

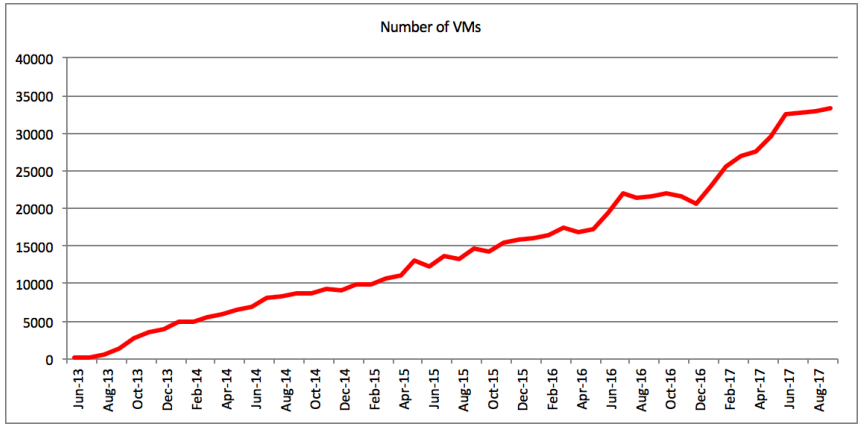
MISTRAL
an OpenStack Community Project

CERN Cloud by Numbers

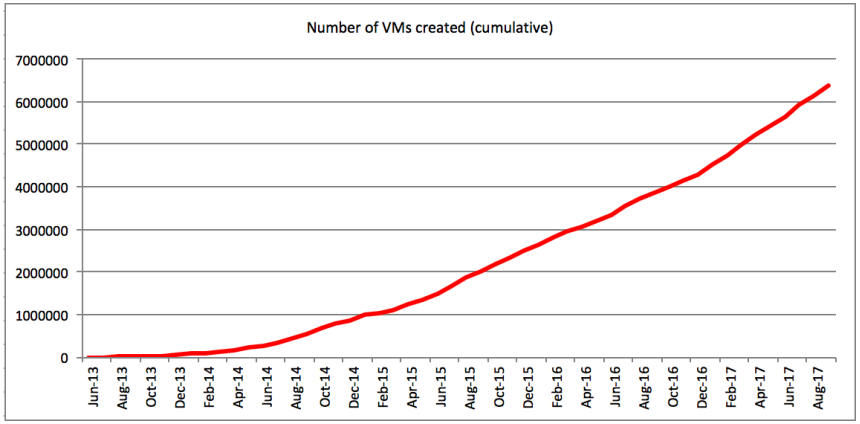


OpenStack timeline at CERN

- Evolution of the number of VMs created since July 2013



Number of VMs running

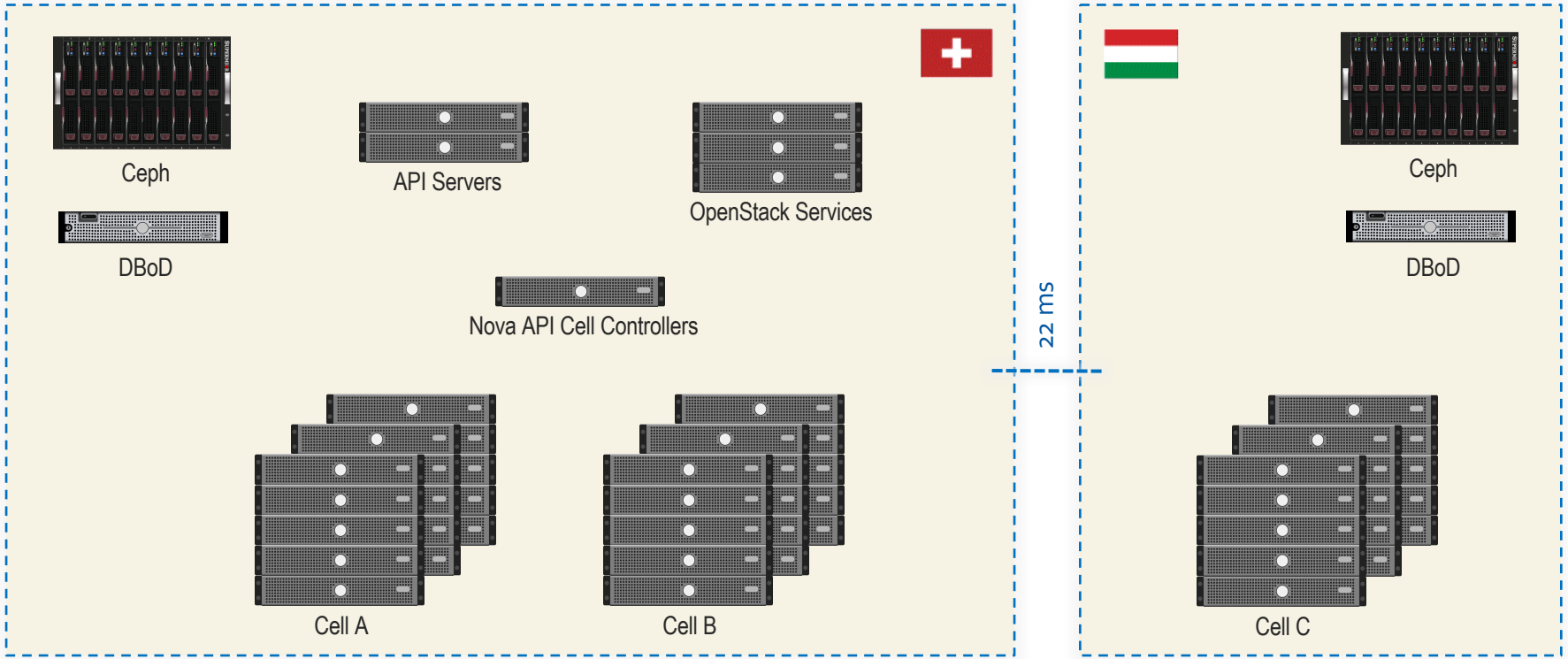


Number of VMs created (cumulative)

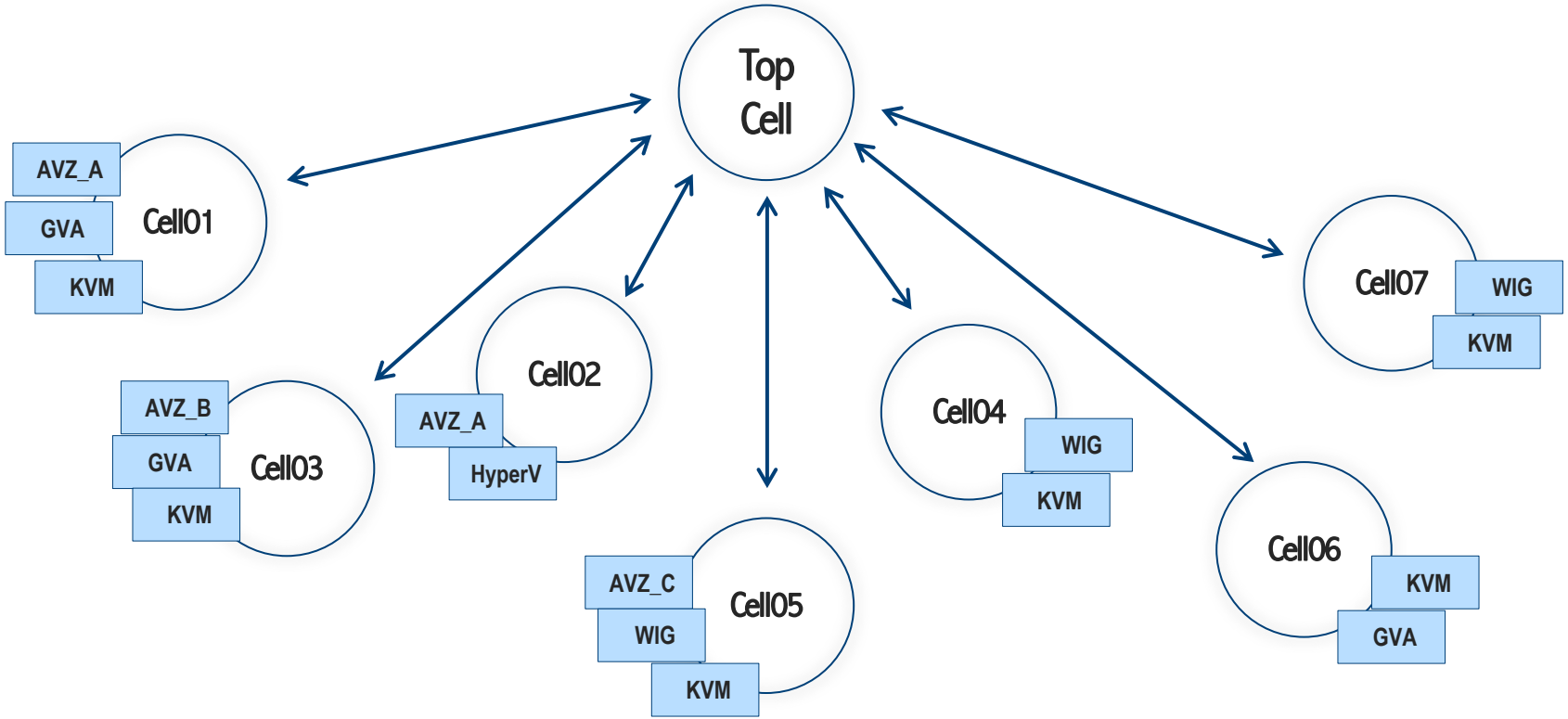
Infrastructure Overview

- One region, two data centres, 62 Cells
- Nova only has HA architecture on Top Cell (replicated message infrastructure)
- Other projects (Neutron, Cinder,…) have a distributed message infrastructure
- No HA for DBs
- Nova control plane and other projects (Neutron, Cinder, Glance, …) are usually VMs running in the shared infrastructure
- New nova cells use Neutron (Linux bridge; provider network)
  - Still 2/3 of the infrastructure is using nova-network
- CERN Centos 7.4
- 3 Ceph instances
- Deployment using OpenStack puppet modules and RDO

Architecture Overview



Cells


Top Cell
Cell01
AVZ_A
GVA
KVM
Cell02
AVZ_A
HyperV
Cell03
AVZ_B
GVA
KVM
Cell04
WIG
KVM
Cell05
AVZ_C
WIG
KVM
Cell06
KVM
GVA
Cell07
WIG
KVM

Workloads/Use cases

- Physics data analysis
  - Batch Jobs; experiments frameworks; …
- IT services
  - Terminal servers; Gitlab; Elastic search; Puppet; Foreman …
- Experiment services
  - Build nodes;
- Engineering Services
  - Microelectronics; chip design …
- Infrastructure services
  - Radiation control; Hostel booking; car renting …
- Personal
  - Desktop replacement; development; …

Virtual Machines - Ephemeral

- Physics data analysis
  - Optimize for compute efficiency
    - CPU passthrough, NUMA aware flavours
  - Still, very different workloads
    - IT Batch
      - LSF and HTCondor; longlived VMs; 8, 10 and 16-core VMs
    - CMS Tier-0: medium-long, 8-core VMs
    - LHCb Vcycle: short-lived, single core VMs
- Low-SLA
- 83% of CPU cores, 67% of VMs

Virtual Machines - Pets

- Service nodes, dev boxes, Personal VMs, …
- Performance less important than “durability”
  - Live migration is an important feature
- Persistent block storage is required
- Linux and Windows VMs
- 14% of CPU cores, 33% of VMs

Containers

- OpenStack Magnum
  - COE available at CERN: Kubernetes; Docker Swarm; Mesos
- Use cases
  - Batch Processing
  - End user analysis / Jupyter Notebooks
  - Machine Learning / TensorFlow / Keras
  - Continuous Integration / Deployment
  - …

Baremetal

- OpenStack Ironic
  - Deploy virtual/physical resources using the same OpenStack APIs
  - Reduce the time to provision a physical server
  - Traceability/accounting
- Containers on Baremetal
  - Interesting for workloads that require maximum performance
- Hardware lifecycle using Ironic

Challenges

- Nova CellsV2
- Migrate from Nova-network to Neutron
- Pre-emptible VMs (low SLA, opportunistic workloads)
- Ironic in production
- Software Defined Networks

belmiro.moreira@cern.ch
@belmiromoreira

www.cern.ch

http://openstack-in-production.blogspot.com