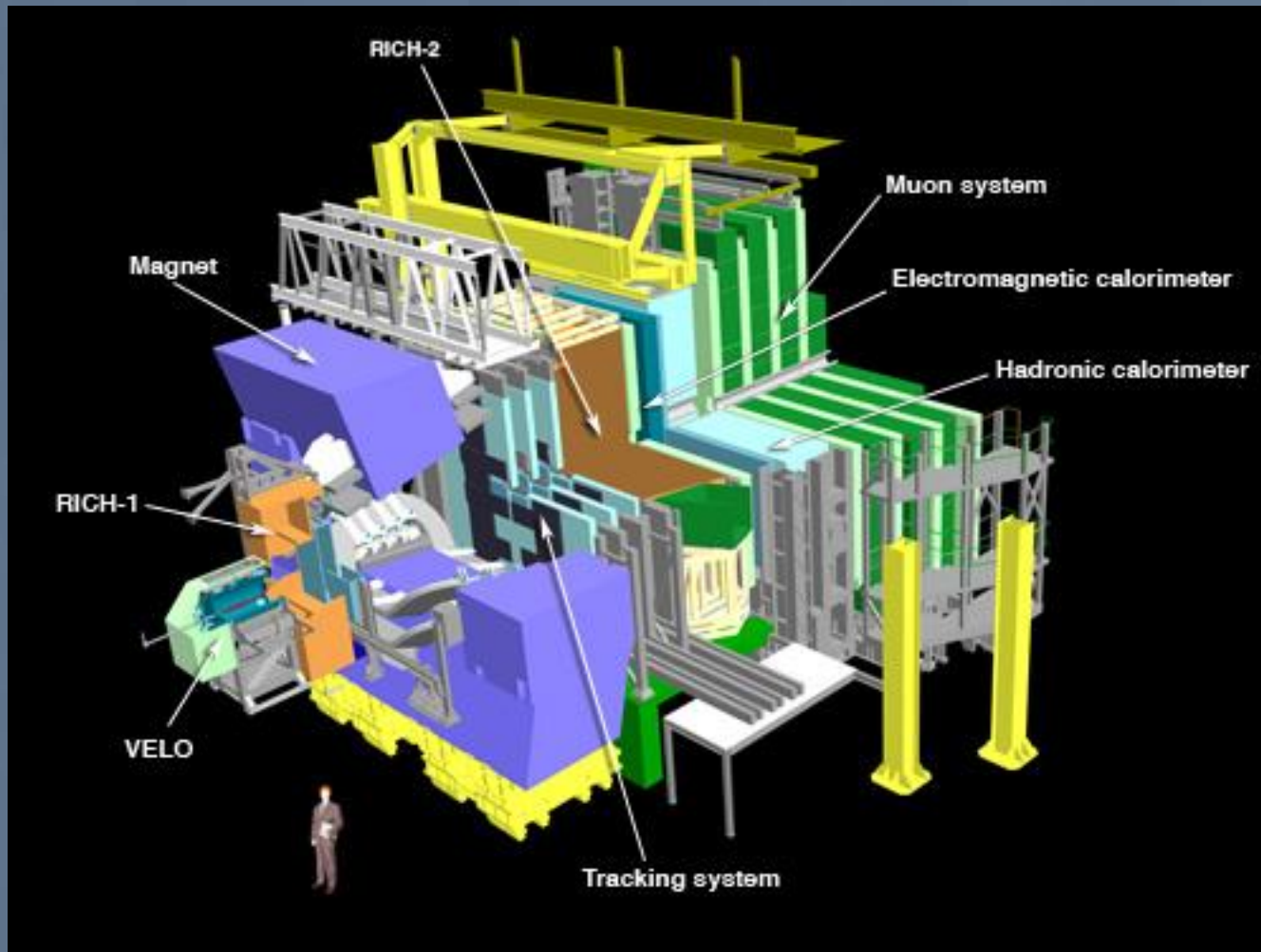# RHEV Platform at LHCb
Red Hat at CERN 17-18/1/17

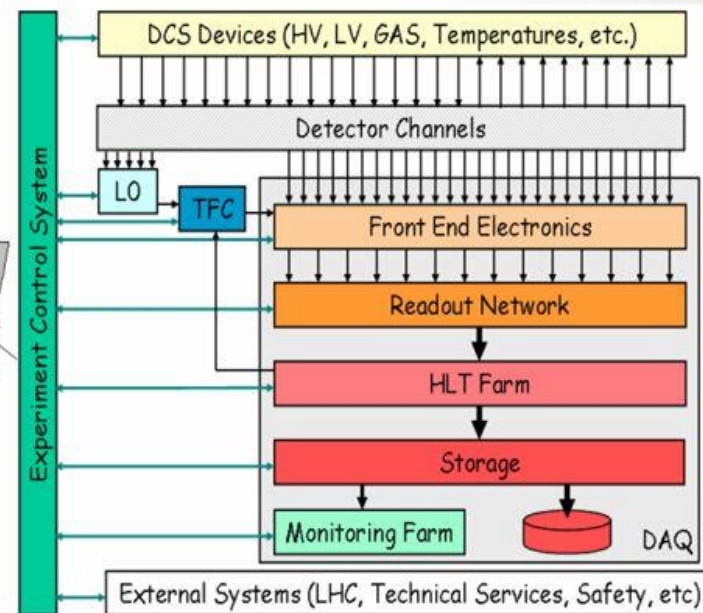Francesco Sborzacchi INFN

LHCb Experiment

- 1000 scientists
- 1000000 sensors (channels)
- > 2000 servers & embedded computers
- 4000 port network

Experiment Control System

LHCb & Virtualization

- Completely isolated network
  - Data Acquisition System
  - Experiment Control System

- Why do we virtualize
  - Improve manageability
  - High Availability
  - Hardware usability
    - Better usage of hardware resources
    - Move away from the model "one server = one application"

DCS Devices (HV, LV, GAS, Temperatures, etc.)

Detector Channels

L0   TFC   Front End Electronics

Readout Network

HLT Farm

Storage

Monitoring Farm   DAQ

Experiment Control System

External Systems (LHC, Technical Services, Safety, etc)

- Enrico.Bonaccorsi@cern.ch – Control System Virtualization for LHCb – ICALEPCS 2013 – San Francisco   • 2

LHCb Virtualization

- Red Hat Enterprise Virtualization
- ~400 VMs
- 2 FC Storage Domains
- 1 NFS Storage Domain
- ~ 4TB Memory
- ~ 10TB Hybrid Storage
- 3 different DATA CENTER

Data Center GENERAL

Data Center PLUS

Data Center CONTROL

LHCb
ONLINE

Data Center Definitions

- ## RHEV-GENERAL
  - General Purpose and infrastructure servers : Gateways, DC, DHCP, DNS, UIs, Terminal Services, Puppet, Web…
- ## RHEV-PLUS
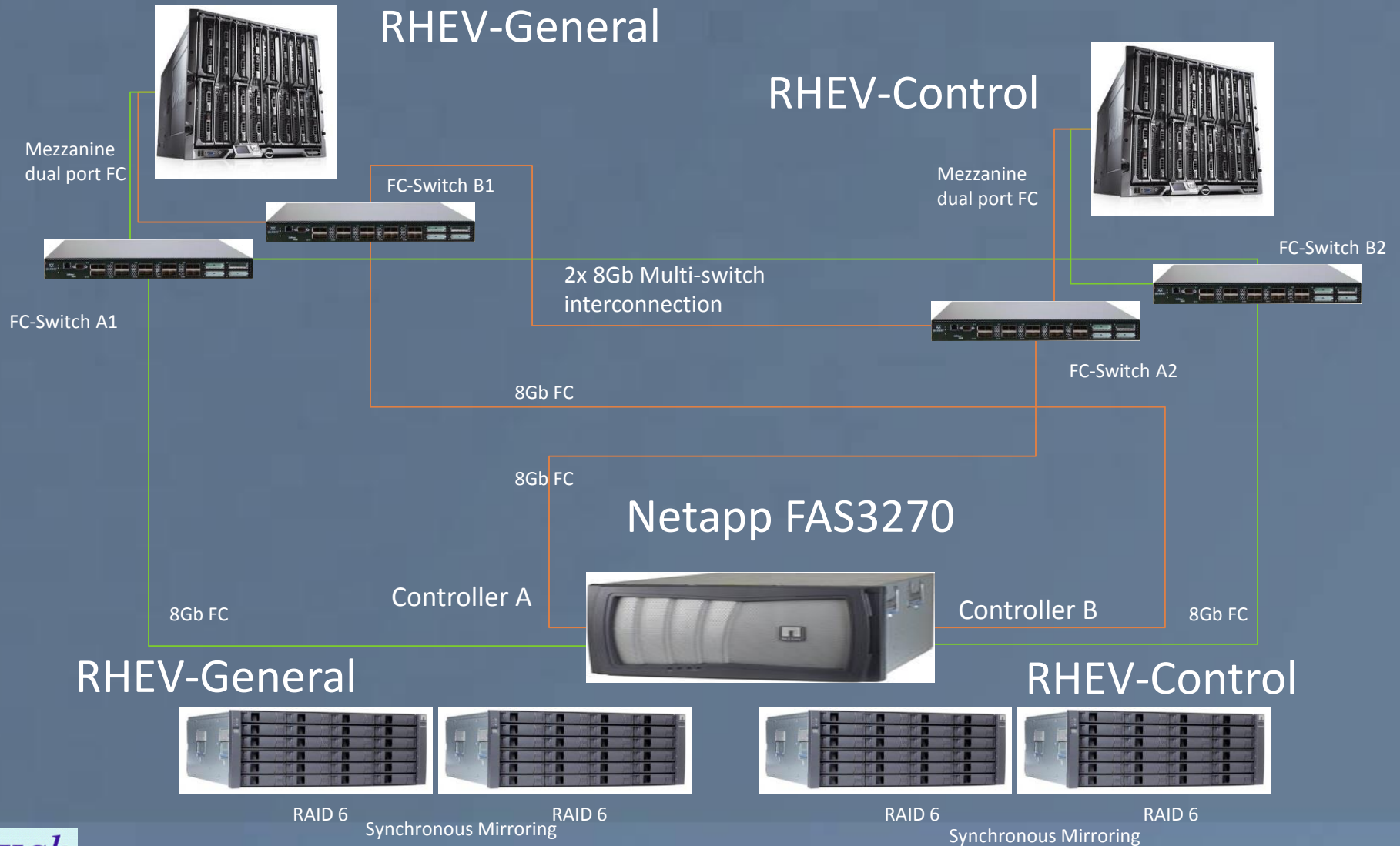 - Computing nodes for interactive user process
- ## RHEV-CONTROL
  - ECS (Experiment Control System): configuration, control, monitoring of the online system.
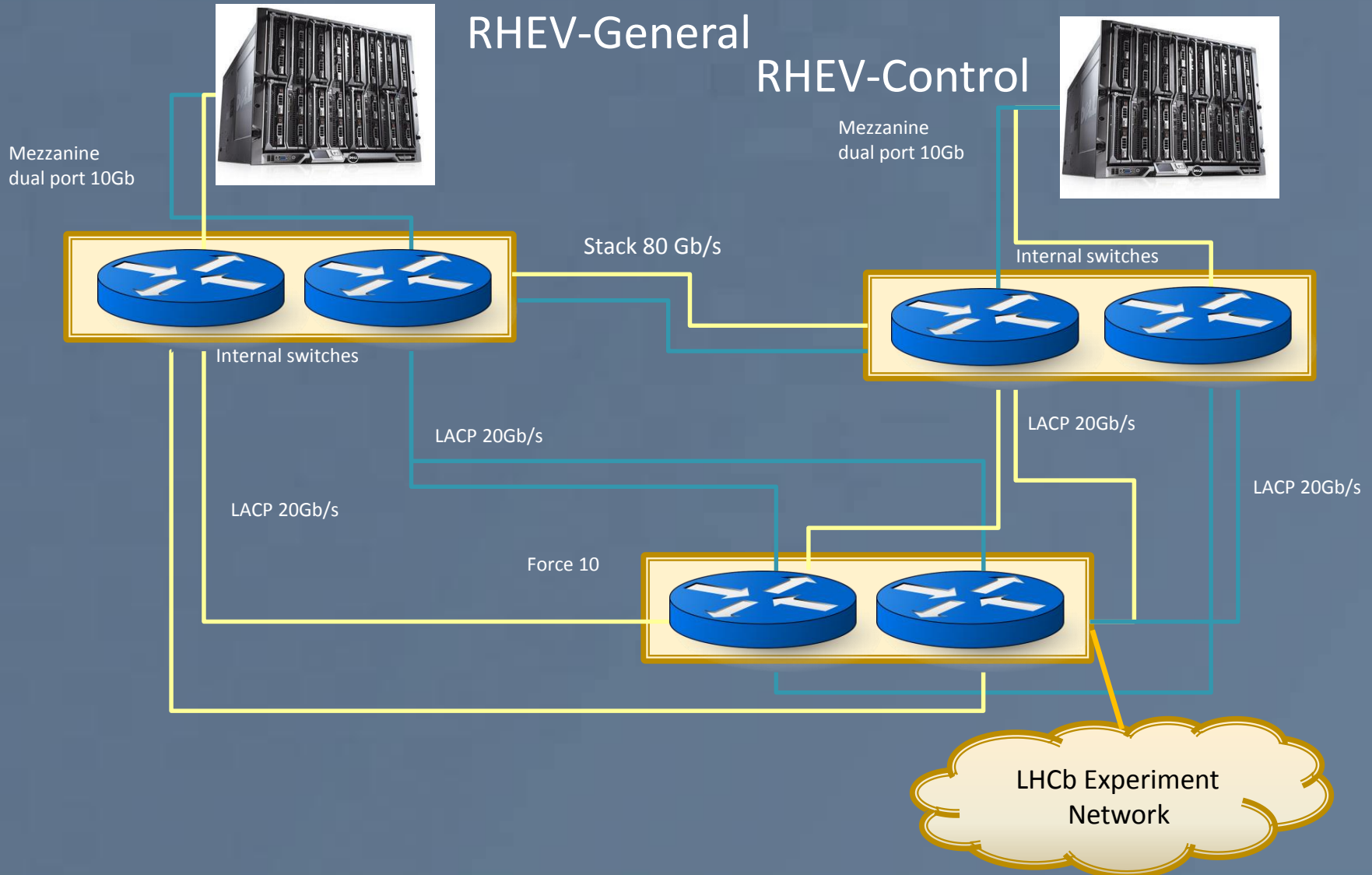
Current Infrastructure

- 3 Independent Clusters
- 2 RHEV MANAGER (3.6 – upgrade to 4.0 soon)
- 30 Hypervisors (RHEV-H 7.2)
- 4x 10Gb/s Ethernet switches (20Gb/s for live migration per Host, 80Gb/s max capacity)
- 4x 8Gb/s FC switches (Shared Storage)
- 2x NetApp 3270 FC+NFS
  - 4x Hybrid Storage SSD+SATA
  - Fully redundant controllers and storage (raid6 ,sync mirror)

Storage Area Network

RHEV-General

RHEV-Control

Mezzanine dual port FC

FC-Switch B1

Mezzanine dual port FC

FC-Switch B2

FC-Switch A1

2x 8Gb Multi-switch interconnection

8Gb FC

FC-Switch A2

8Gb FC

Netapp FAS3270

Controller A

Controller B

8Gb FC

RHEV-General

8Gb FC

8Gb FC

RHEV-Control

RAID 6

RAID 6

Synchronous Mirroring

RAID 6

RAID 6

Synchronous Mirroring

Ethernet Network

RHEV-General

RHEV-Control

Mezzanine
dual port 10Gb

Mezzanine
dual port 10Gb

Stack 80 Gb/s

Internal switches

Internal switches

LACP 20Gb/s

LACP 20Gb/s

LACP 20Gb/s

LACP 20Gb/s

Force 10

LHCb Experiment
Network

Ethernet Network PLUS

RHEV-PLUS

Mezzanine
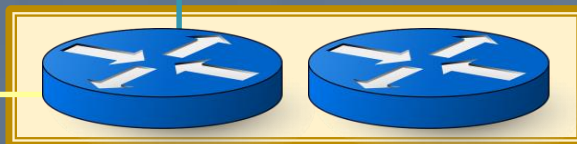dual port 10Gb

Netapp FAS3270

NFS 10 Gb/s

LACP 20Gb/s

Force 10

Ethernet 10Gb/s

LHCb Experiment
Network

Resource Optimization

- Kernel Shared Memory
  - Merging of the same memory pages (Shared >50%)
- Physical Memory Guaranteed per VM
- Bandwidth and Priority def. for live migration
- Storage DEDUPLICATION
  - Currently savings of 70%

RHEV key features

- High Availability
  - Allows Critical VMs to be started/migrated automatically
- System Scheduler
  - automatically balances the VM load among hosts in a cluster
- Maintenance
  - Running VMs live migrated (no downtime)
  - HV and HW intervention without interruption
- Deploy
  - Rest-API and Templates for easy and fast new installation

RHEV: bad things

- Major release upgrades
  - Upgrade between major release encountered a LOT of problems
    - Most of the time we had to connect directly to the DB engine and use SQL query directly to make the things work.

- Tuning of the MANAGER host not well documented
  - Performance and usability completely out of control when the number of VMs >100
  - DB engine not optimized at all (and not mentioned anywhere on the manual)
- Tuning of the HV host
  - Live migration out of the box limited and "hard coded" to 32Mb/s (out of our 20Gb/s capabilities)

Conclusion

What is crucial for virtualization:
- AVAILABILITY
- Shared Storage
- Flexibility

Why do we like RHEV:
- NO FAILURES during 4 Years of services
  - Most of the down-time due to HW or external reason
- Very good performance and QOS
- Extremely Positive Balance in terms of resources optimization and HW reduction