14th of March 2013
ALICE, ATLAS, CMS & LHCb joint workshop on *DAQ@LHC*

Wojciech Ozga
dr Jose Antonio Coarasa Perez
on behalf of ALICE, ATLAS, CMS, LHCb

Offline usage of online farms

Online farms @ CERN

Re-usage of resources

Network integration

Hardware integration

Conclusions

OUTLINE

Online farms @ CERN

Designed for data taking

20TBytes/day (CMS)

Complex

varied hardware
different network schemas
hundreds of OS configurations

Large size

	Nodes	Switches	Switch ports
CMS	1264	150	7000
ATLAS	1493	150	3000
LHCb	1044	200	
ALICE			1800

Powerful

Cores	RAM [Tbytes]	Storage [Tbytes]	
13312	26.6	232	CMS
15272	33	315	ATLAS
15008	26		LHCb
2500			ALICE

Online farms vs Tier$_{[0\text{-}2]}$
CPU in HEP-SPEC06

	HLT farm	Tier0	Tier1	Tier2
sum	602k + ALICE	356k	603k	985k
CMS	195k	121k	150k	399k
ATLAS	197k	111k	260k	396k
ALICE		90k	101k	143k
LHCb	210k	34k	92k	47k

http://wlcg-rebus.cern.ch/apps/pledges/resources/

Requirements

Minimal Changes

no special hardware reconfiguration

Opportunistic Usage

Re-usage of resources when no data taking
Flexibility in running different tasks

VIRTUALIZATION

cloud approach followed by CMS, ATLAS, ALICE

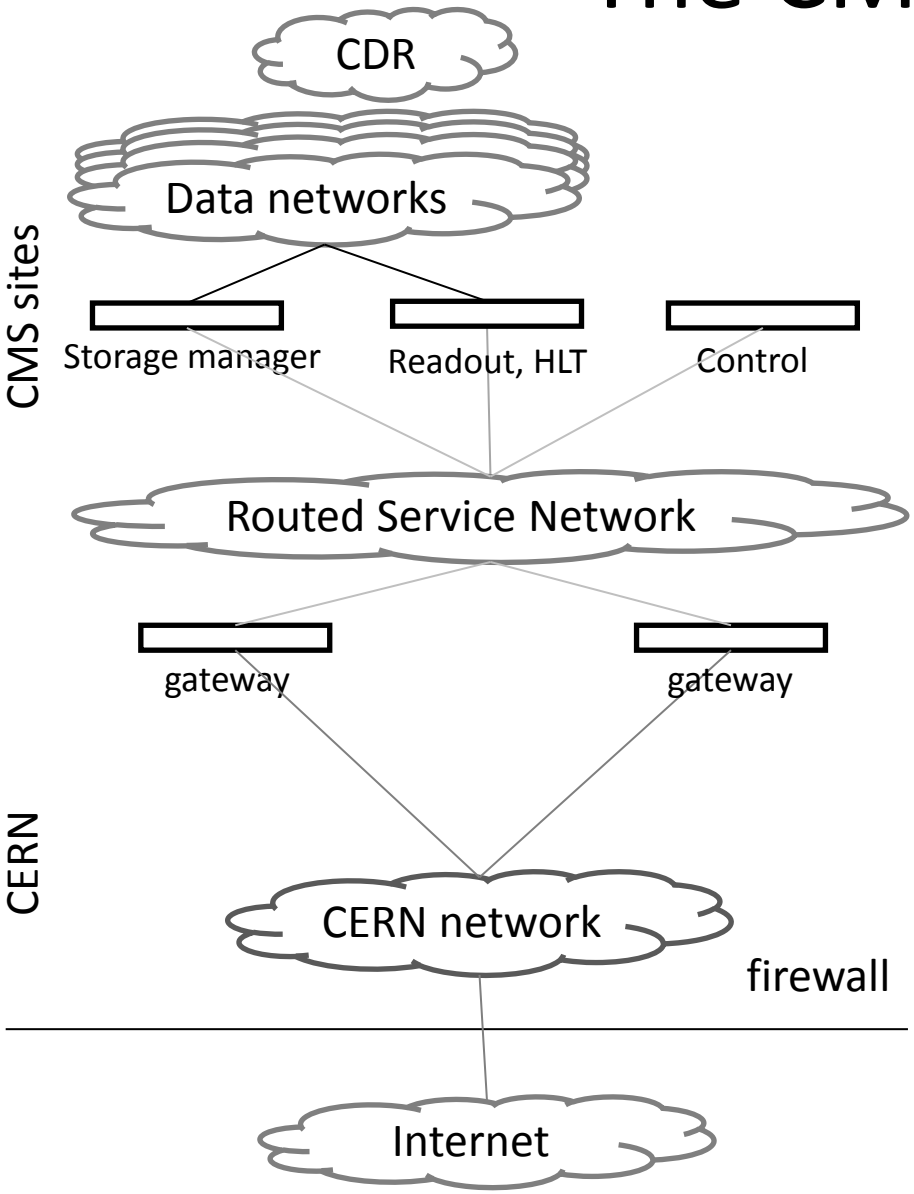
No impact on data taking

Nor during the setup phase
Neither during data taking on Heavy Ion runs

Full control

Full control on hardware and OS
no extra temporary sudoers

The CMS Online Cluster



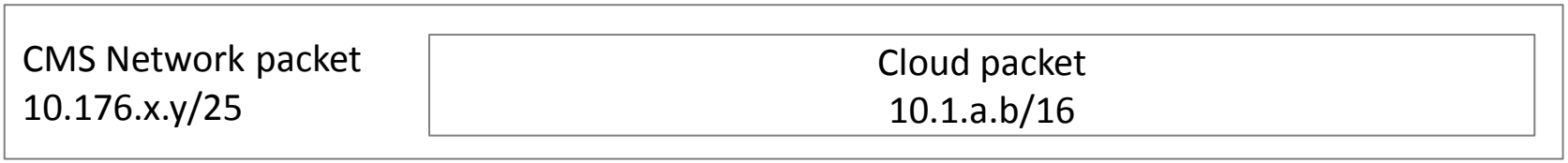
Private CMS networks

- Service networks (per rack)
  ~3000 1Gbit ports
- Data networks (VLAN's)
  ~4000 1Gbit ports
- Central Data Recording (CDR)
  Network to Tier0
- Oracle RAC networks
- Subdetectors networks

CERN network

Network virtualization (CMS)

Integration of cloud network with CMS network → GRE TUNNELING → network virtualization



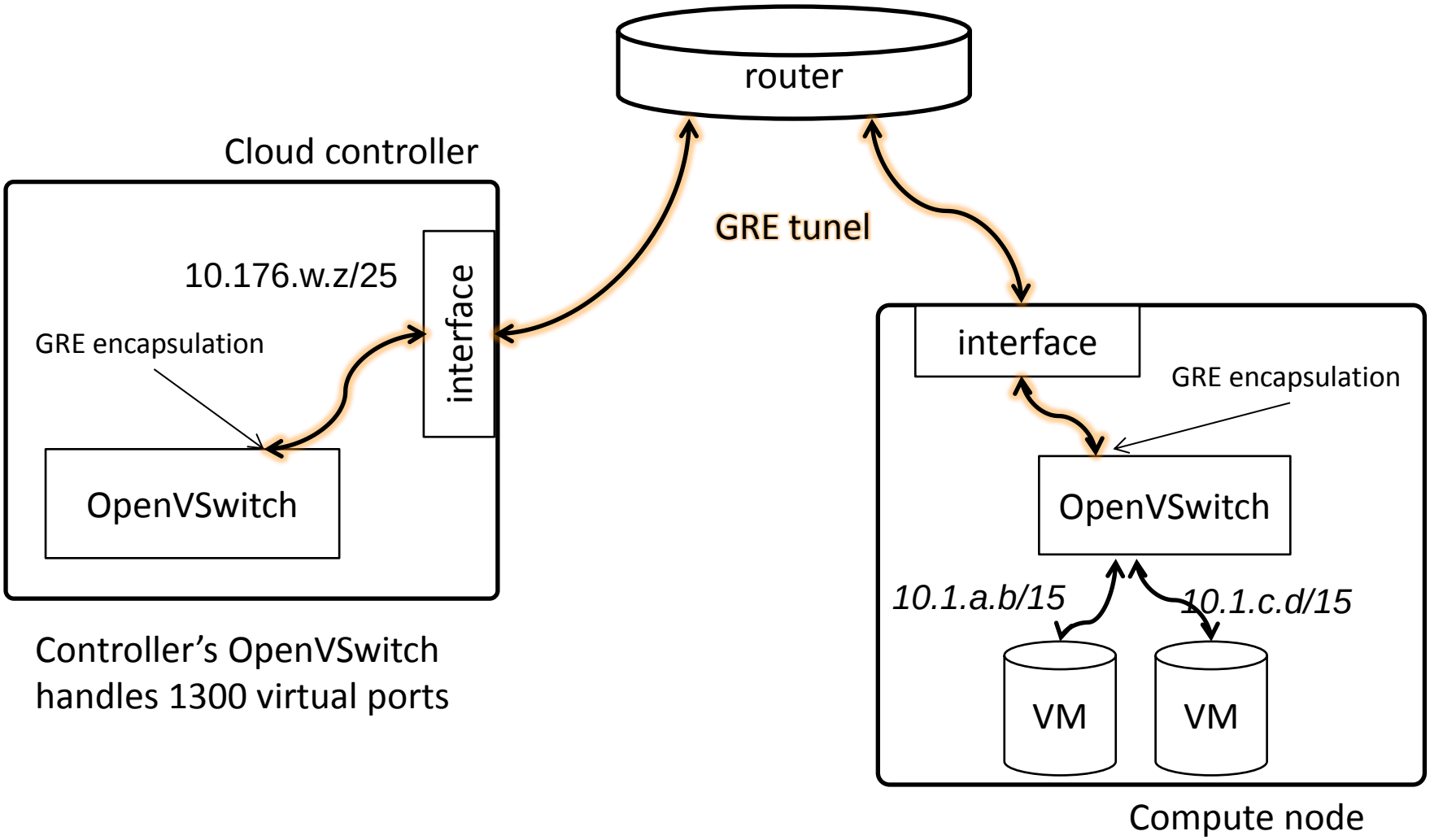
+

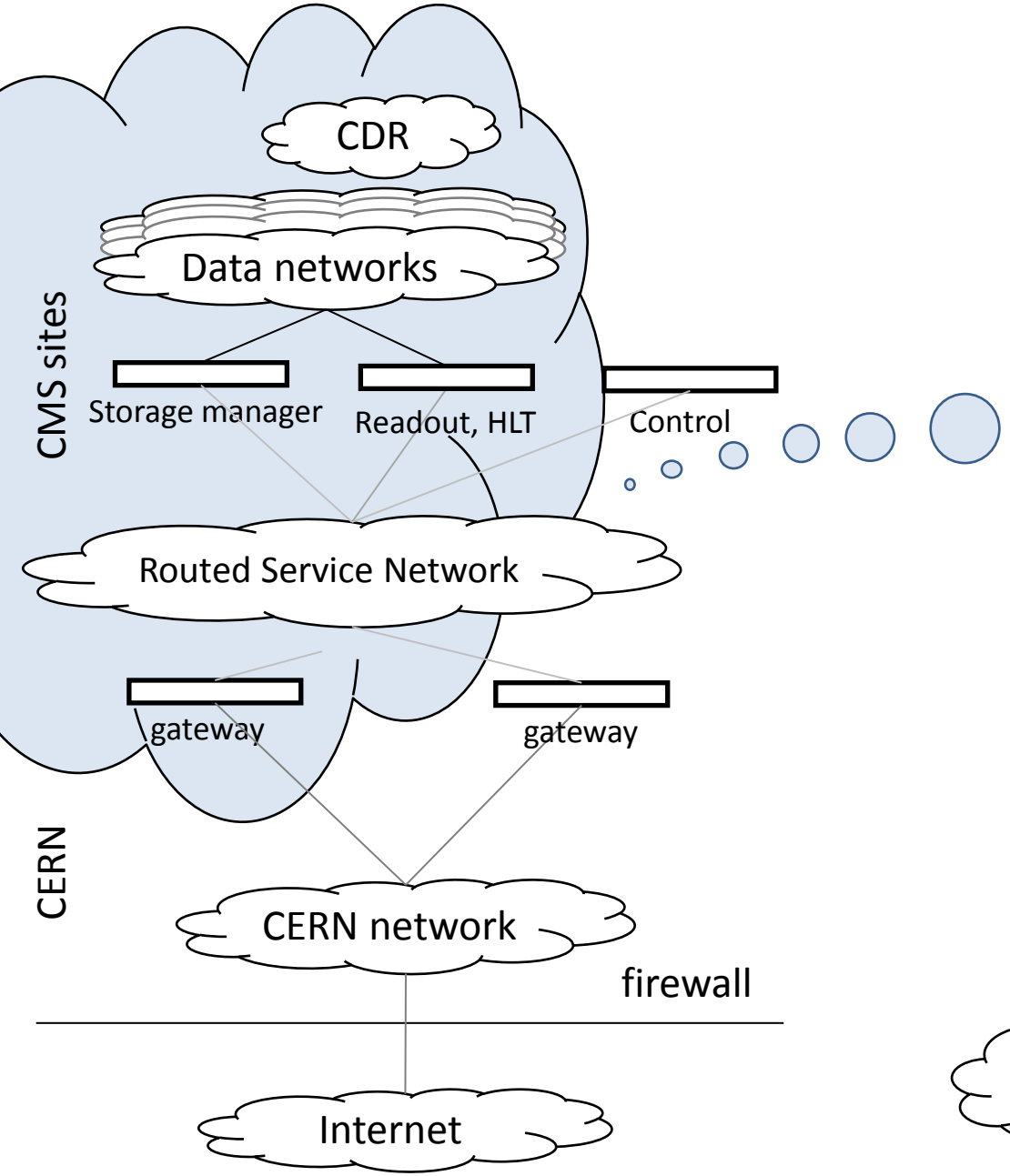
No router/switch re-configuration
No dependencies between production and cloud network

-

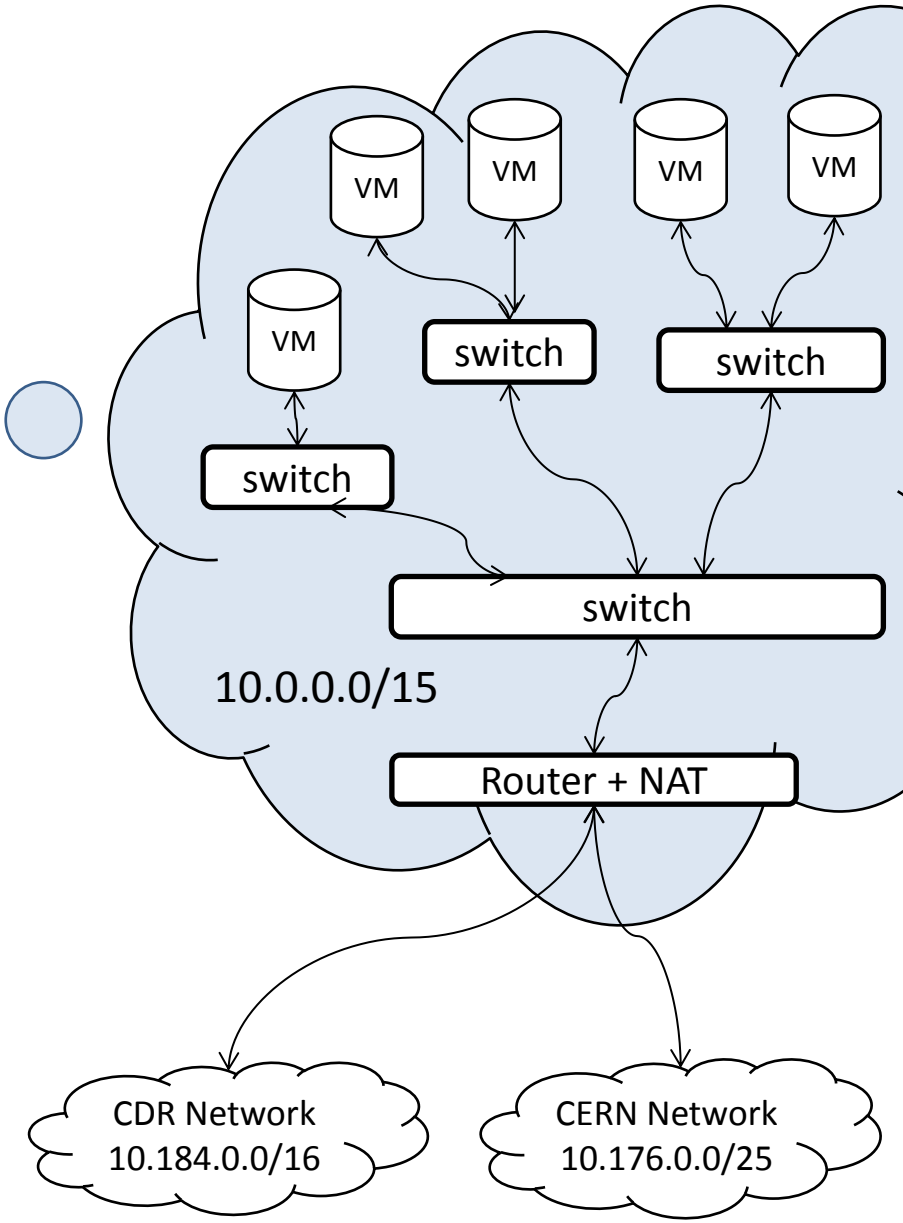
Encapsulation overhead
NATing to CERN network required

Network virtualization (CMS)





Physical network schema @ CMS



Logical cloud network schema

Hardware virtualization

Standard approach

Configure physical machine to run specific task

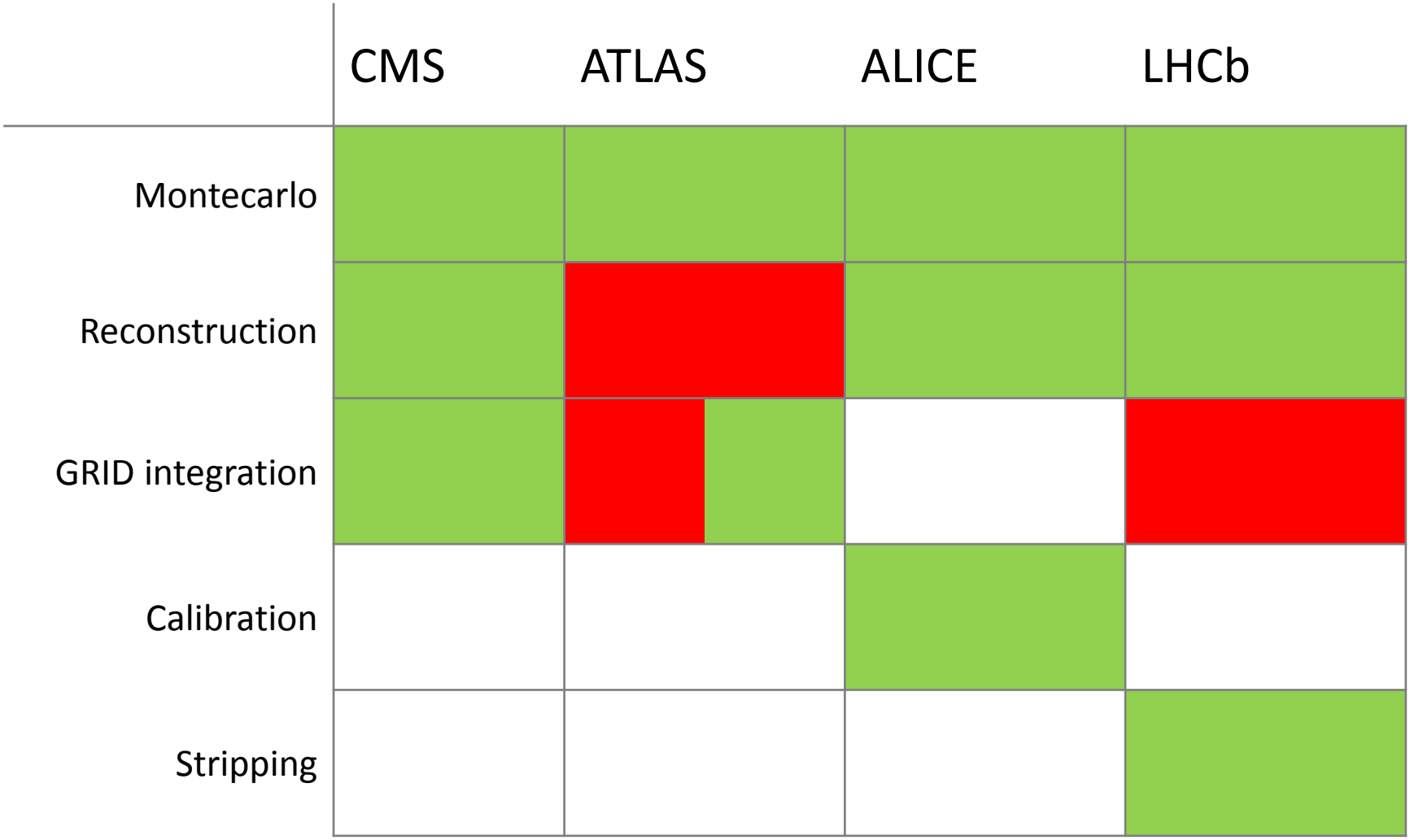
Number of configurations = Number of tasks x number of different hardware configurations

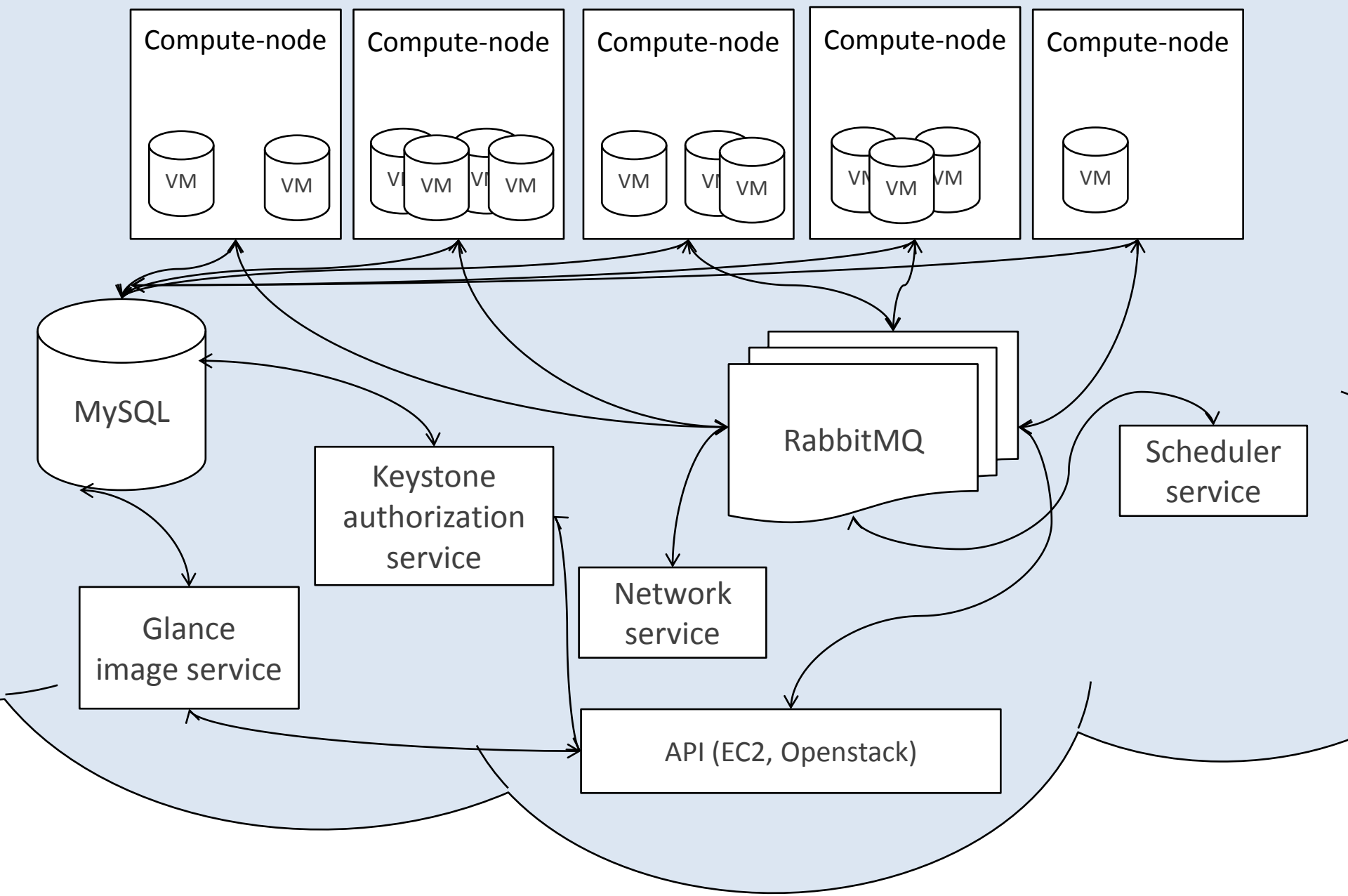
Cloud approach

Launch task specific **Virtual Machine** over hypervisor

Number of configurations = Number of tasks

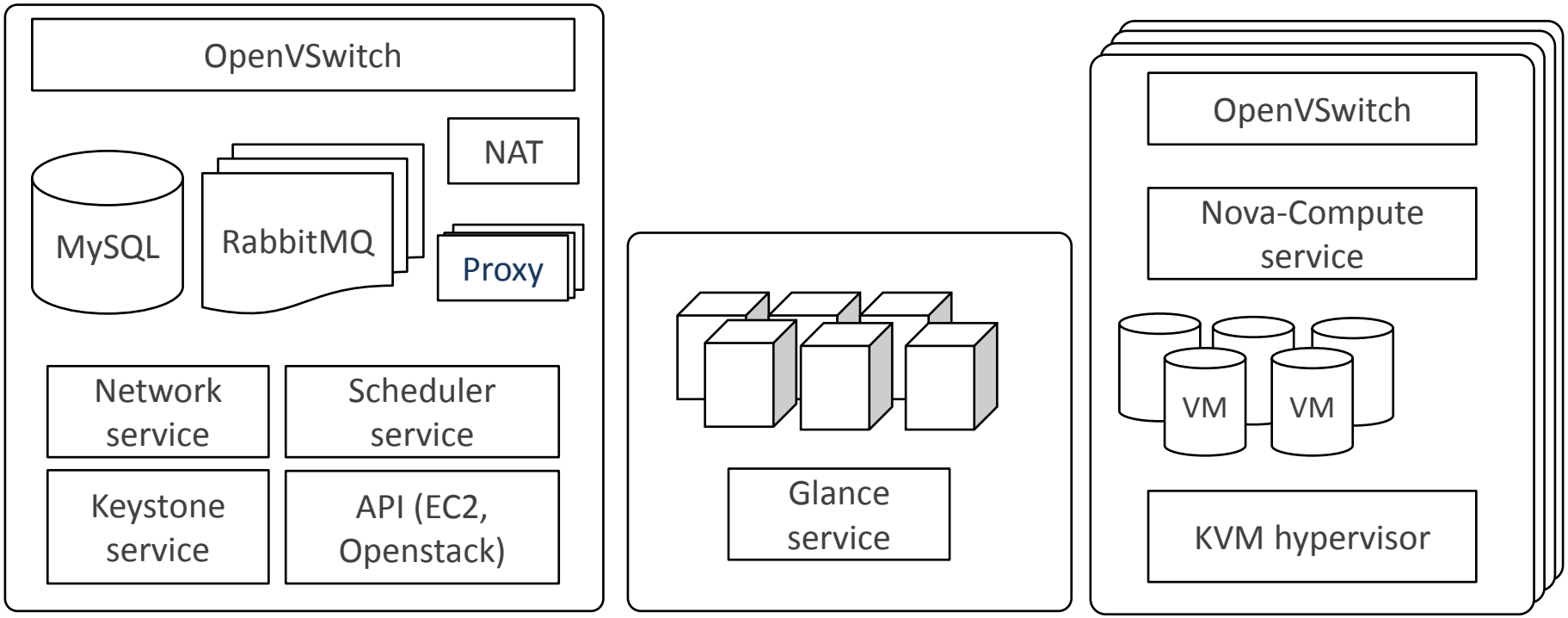
Offline tasks





Openstack Cloud Manager

Cloud Architecture (CMS, ATLAS)



1 x Fat Controller Node
48Gbytes, 8 CPU, 8x1Gbit Ethernet (CMS)

1 x VM Image Store
High bandwidth
Cache mechanisms

1300 x Compute node (CMS)
KVM virtualization
OVS network virtualization

"The best advice I know is 'Don't distribute.'"
Adam Bien

CMS Cloud usage

8th – 12th of October 2012

First big scale test: **Folding@Home**
Cloud controls 1300 hypervisors
768 completed Work Units in 12h

21st of December 2012 – 11th of January 2013

Test of stability, running **Folding@Home**
Stable run of 1000 VMs during 3 weeks
Deployment of 250VMs in ~5min
54625 completed Work Units

Since January 2013

Finalizing the integration as a GRID resource

Conclusions - CMS

An Overlay Cloud layer deployed on CMS HLT cluster
with Zero impact on data taking.

Man power dedicated to cloudify the CMS HLT cluster
was about ~1.5 FTE for ~6months

CMS online share knowledge about cloud installation

CMS HLT cluster computing power is being used in an opportunistic way
leaving full control over cloud availability to DAQ CMS

Conclusions - ATLAS

ATLAS follows cloud approach using OpenStack,
first test installation finished

ATLAS skips network virtualization due to NATing overhead

Cloudified ATLAS HLT farm will partly join GRID resources
not fully due to security issues

Work responsibilities:
TDAQ sysadmins + BNL: Infrastructure as a Service (IaaS)
ATLAS offline: virtual machines (VMs)

Conclusions - ALICE

ALICE follows cloud approach with emphasis on CernVM Cloud

first cluster started, instances under control

ALICE HLT plan to run offline tasks (simulation, reconstruction, calibration, …)

Conclusions – LHCb

LHCb HLT is ready to run MonteCarlo, Reco and Stripping

MonteCarlo is running since the begining of LS1

Thanks for contribution:

Marc Dobson (CMS)
Sergio Ballestrero (ATLAS)
Franco Brasolin (ATLAS)
Niko Neufeld (LHCb)

Thank you. Questions?