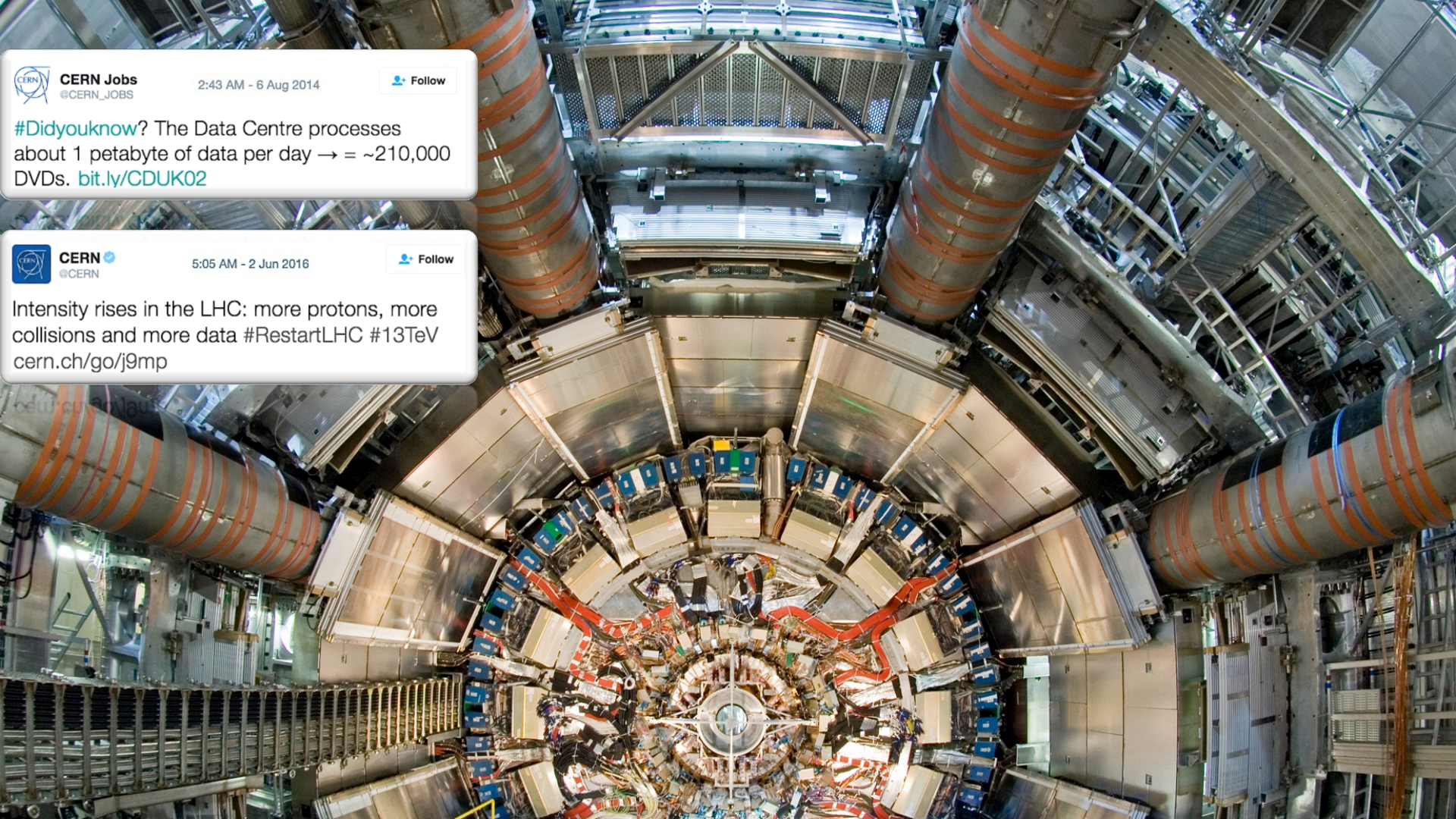
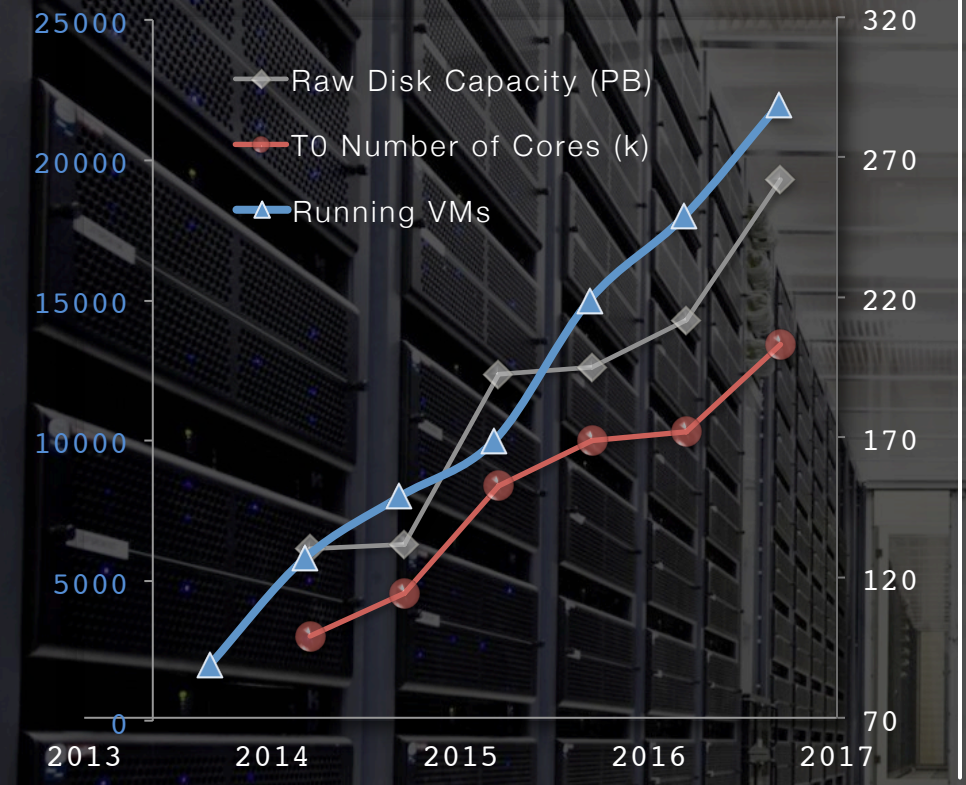
CERN Computing in Commercial Clouds

C. Cordeiro, on behalf of the CERN IT Cloud Integration Team

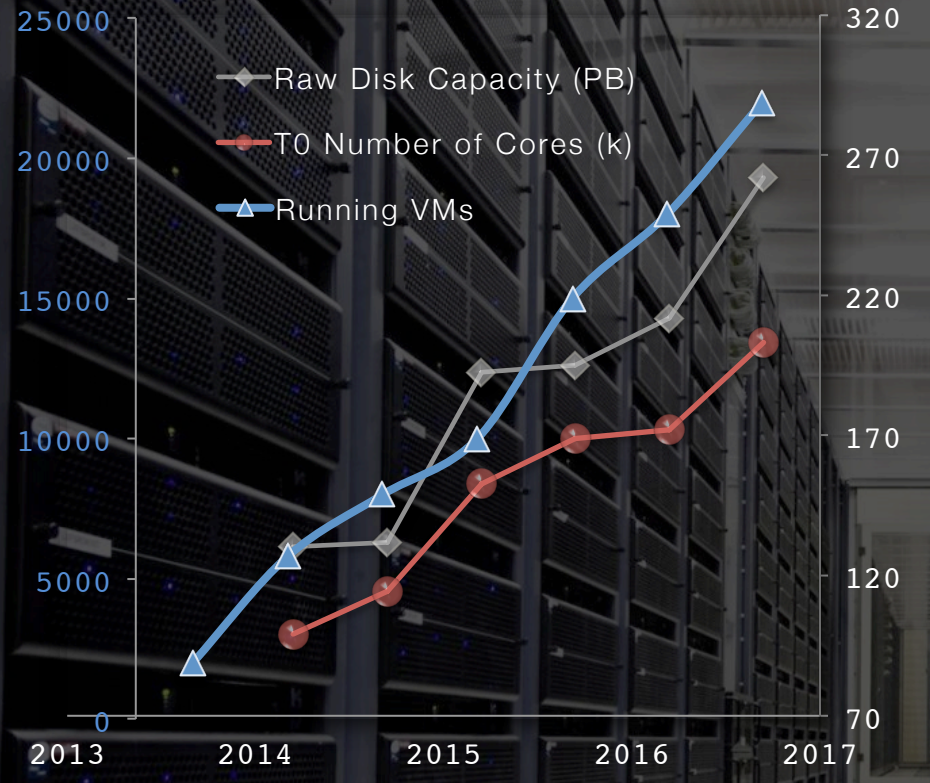
22nd International Conference on Computing in High Energy Physics
October 2016


CERN Jobs
@CERN_JOBS
2:43 AM - 6 Aug 2014
Follow
#Didyouknow? The Data Centre processes about 1 petabyte of data per day → = ~210,000 DVDs. bit.ly/CDUK02
CERN
@CERN
5:05 AM - 2 Jun 2016
Follow
Intensity rises in the LHC: more protons, more collisions and more data #RestartLHC #13TeV cern.ch/go/j9mp

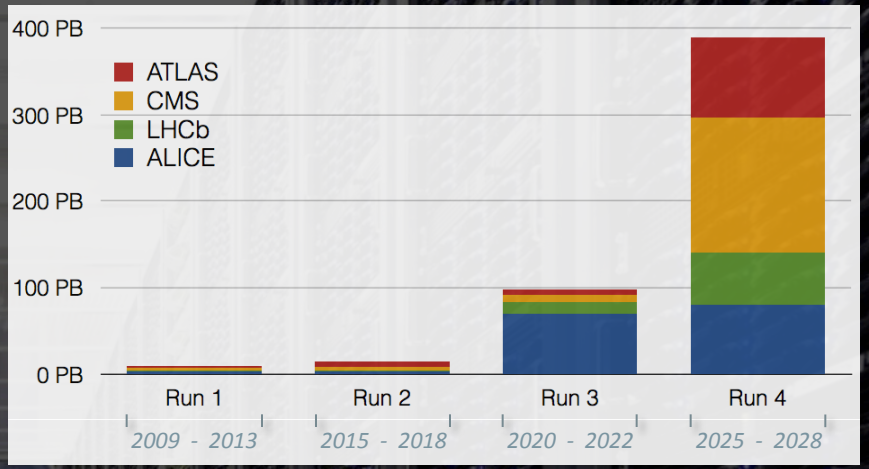
COMPUTING HISTORY



COMPUTING HISTORY



FUTURE CHALLENGE

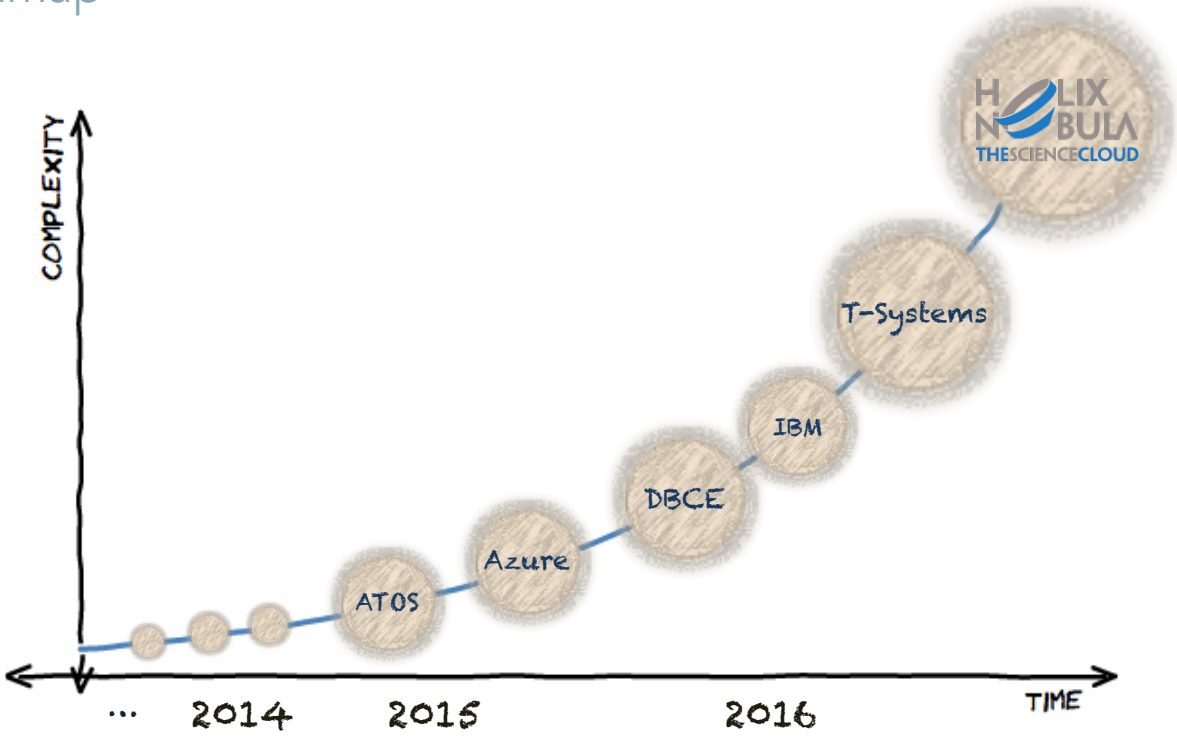


Moore helps, but need more


Elastic
scalable
on-demand
IaaS
grid
commercial
VMs
CERN
resources
Extra
Capacity
technology
remote
services
flexible
Public Cloud
virtual
trend
hybrid
Cloud Computing
storage
WLCG
data
Externally Managed

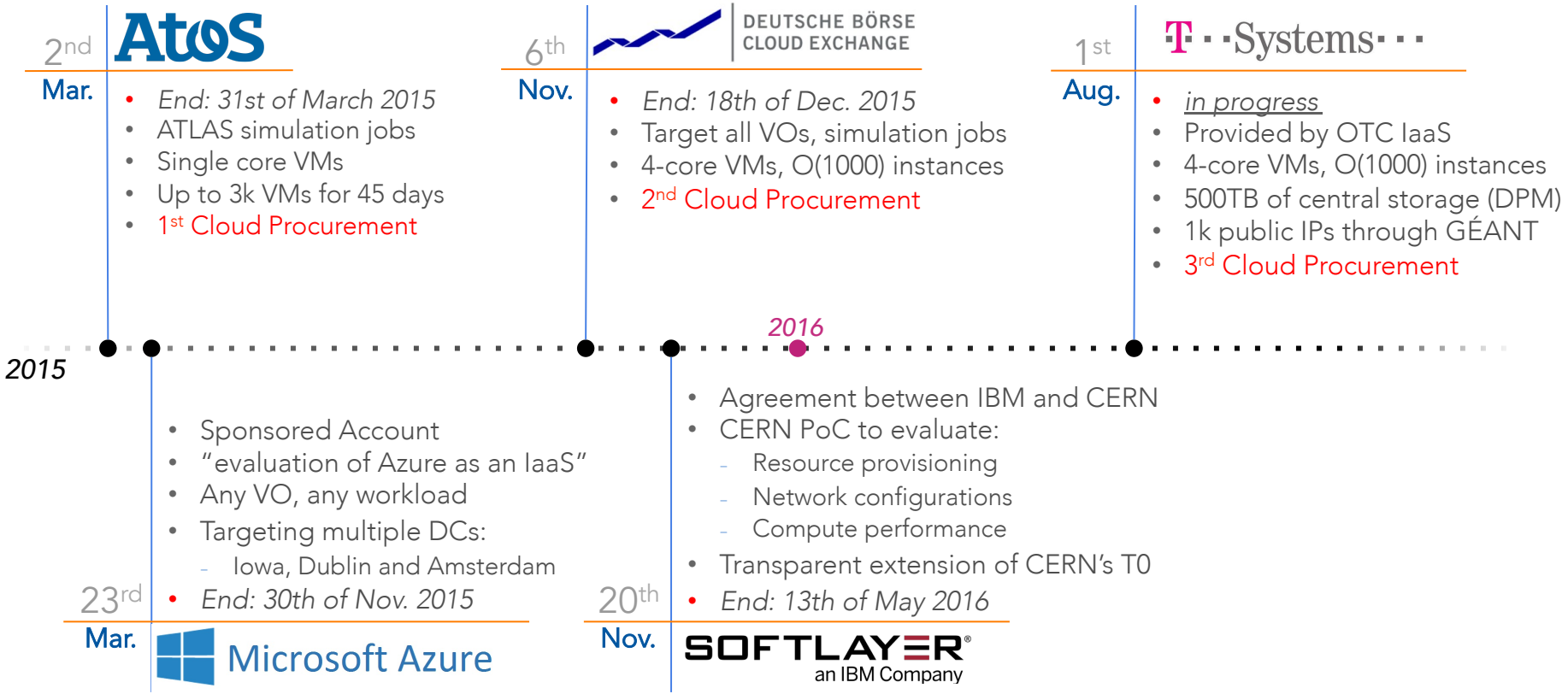
Integrating Clouds into the WLCG

complexity roadmap




Supporting Services
Image Management
Capacity Management
Quota Management
FUNCTIONAL
MINDSET
FOR COMMERCIAL CLOUDS
Monitoring
Data Access and Network
Pilot Job Framework
Accounting

…since 2015@CERN



3000 running VMs (CentOS 6) for 5 weeks
= ~1.2M CPUh of processing

~11.5M ATLAS GEANT4 simulation events processed ⇔ ~160k files produced

~97% job wall clock spent on successful runs

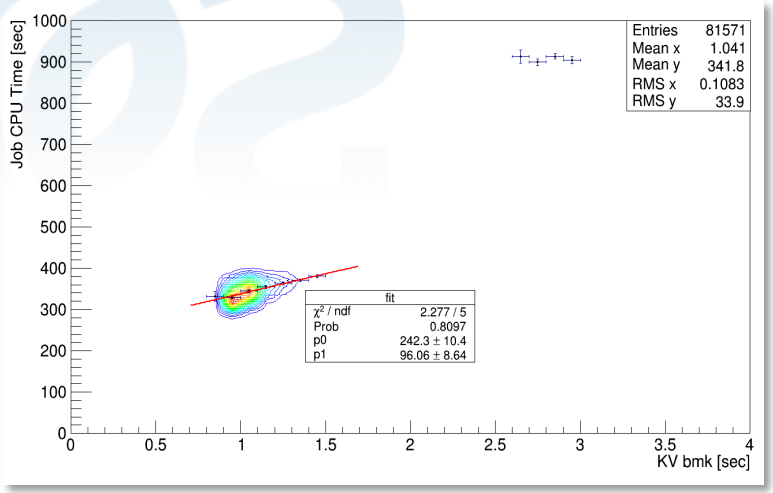
30k CPU benchmarks (KV) ran, collected and correlated with job CPU time

+

Custom resource manager with workload based auto-scaling and interface with SlipStream

+

Ganglia based resource monitoring and accounting



3 VOs included: LHCb, ATLAS and CMS, running several job types

3 different DCs targeted simultaneously

~4800 vCPUs reached

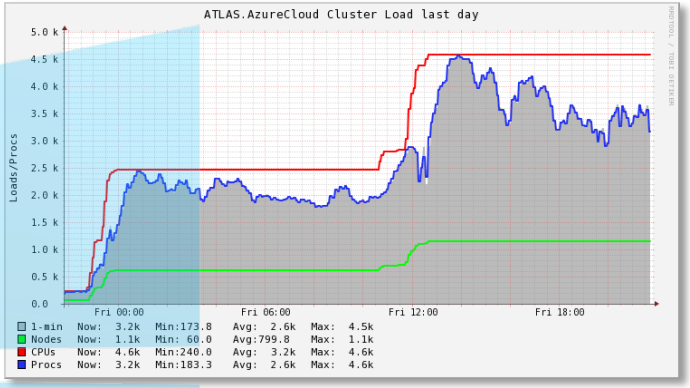
>240k benchmarks collected and analyzed in a *Performance* ***vs*** *Effective Cost* study

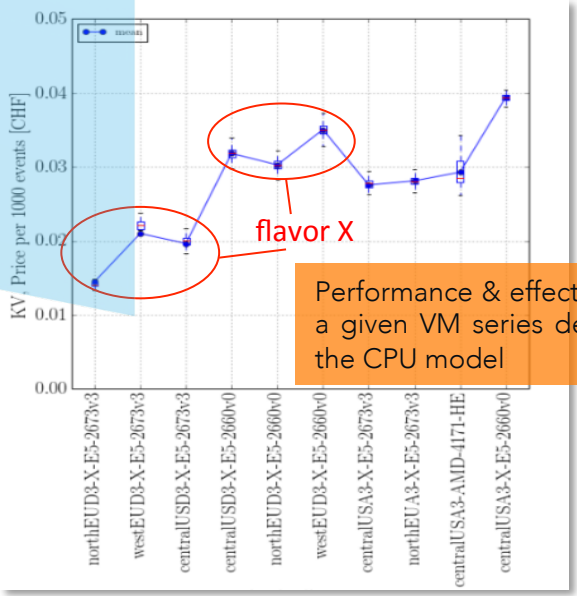
+

Adoption of CernVM and Squid caches

+

Template based provisioner wrapping the Azure Resource Manager CLI





Performance & effective cost of a given VM series depends on the CPU model

5 different cloud providers through DBCE marketplace, shared amongst 4 VOs

50% full capacity provided and used

~1.5M aggregated corehours of processing

<5% job efficiency difference vs *ALICE* jobs running at CERN

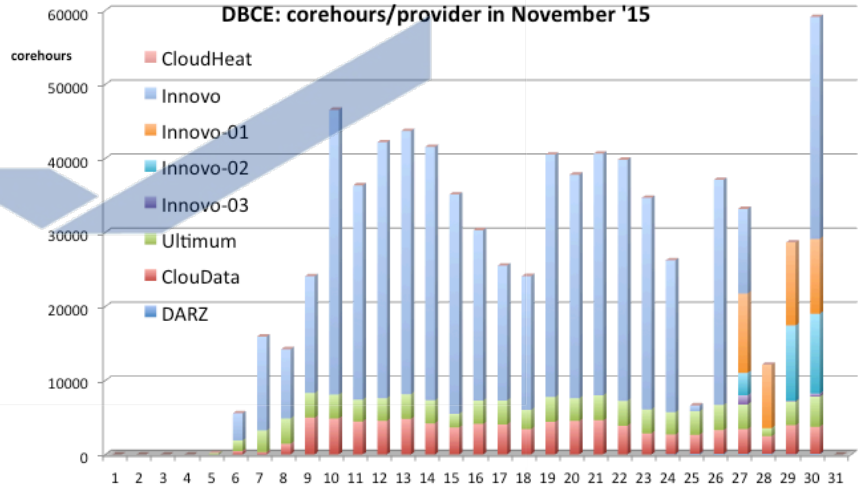
84% walltime spent on successful *ATLAS* jobs

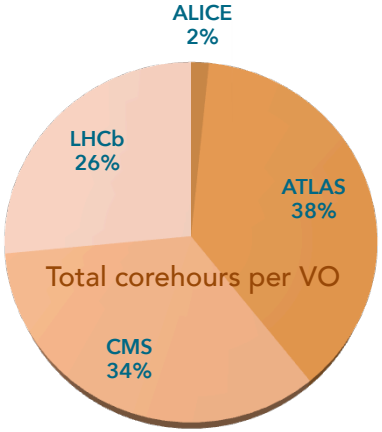
87% of the *CMS* jobs finished successfully

~24 years of CPU time spent on successful *LHCb* full chain processing workloads

+

Provisioning through VCycle and continuous CPU benchmarks to validate and compare the presumed and perceived performance





RESOURCES

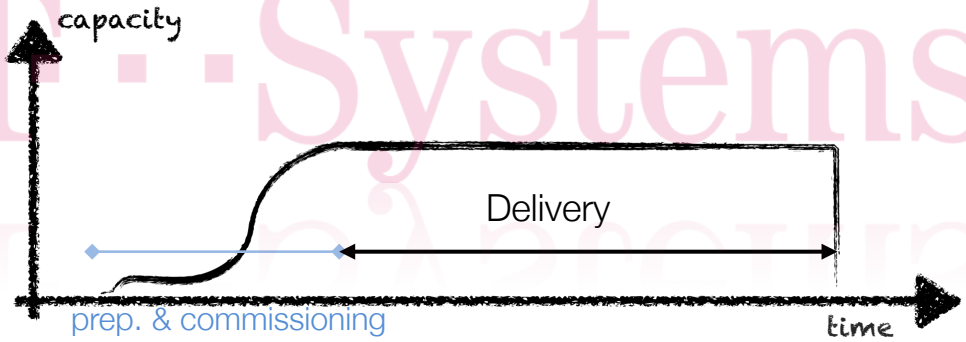
- Terraform for provisioning
  - For virtual resources only
- Setup Puppet on contextualization
  - Additional facts for external hosts
- Job management through HTCondor
- Adoption of SLC
- Hosts under *softlayer.cern.ch* domain
- Monitoring and alerts delivered through *Lemon*
- ATLAS simulation workloads to validate the infrastructure

NETWORK

- Evaluate basic performance
- Evaluate VPN adoption
  - ~2x higher throughput w/o VPN
  - Tested using perfSONAR in bare metal box with 10Gbps NIC
  - Comparable to regular Tier 1

COMPUTE PERFORMANCE

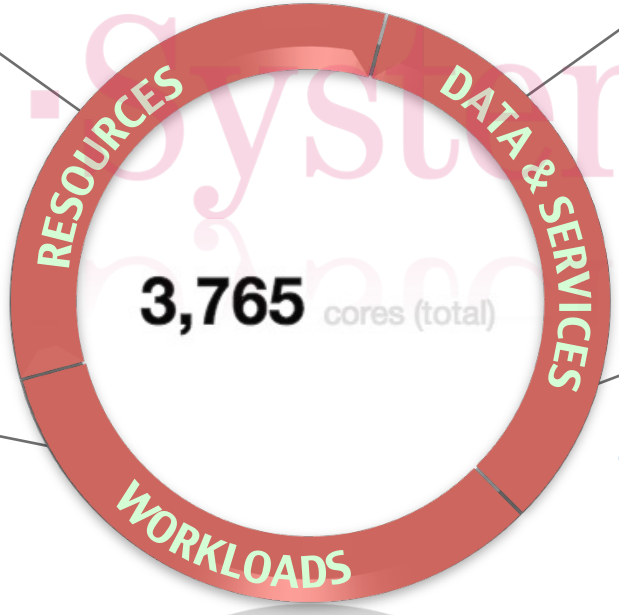
- HEP-SPEC06 on bare metal
  - Comparable with same HW at CERN, within 1 σ
- Synthetic benchmarks on VMs and bare metal



- Network reserved peak at least 10Gb/s
- Bandwidth of at least 1 Gb/s among VMs
- >800 IOPS **&** >50 MB/s streaming r/w, per VM
- >15000 IOPS **&** >125 MB/s streaming r/w, per data disk
- Pass the benchmark reference

881 puppetized workernodes

- 4 cores, 8GB RAM, 100GB disk
- Public IP through GÉANT
- Scientific Linux CERN 6.8
- Lemon monitoring
- Terraform provisioning

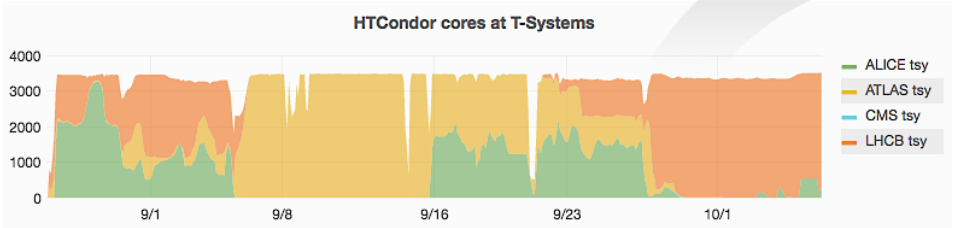


2 DPMs with **480TB**

- **1** for ATLAS and CMS, **1** for ALICE
- Total of **50** VMs - 48 disknodes
  - 2 filesystems of 5TB per disknode

Job management - **HTC**ondor

- Transparent job submission
- Same HTCondor batch instance as T0
- Avg CPU efficiency ≅ **80%**



Data Analytics from Monitoring

- Over **35M** monitoring events processed

Top 10 message.condition	Count
high cpu_idle	8271041
disk usage growing too fast	7975263
high memory swap usage	7956145
high cpu io wait	7183900

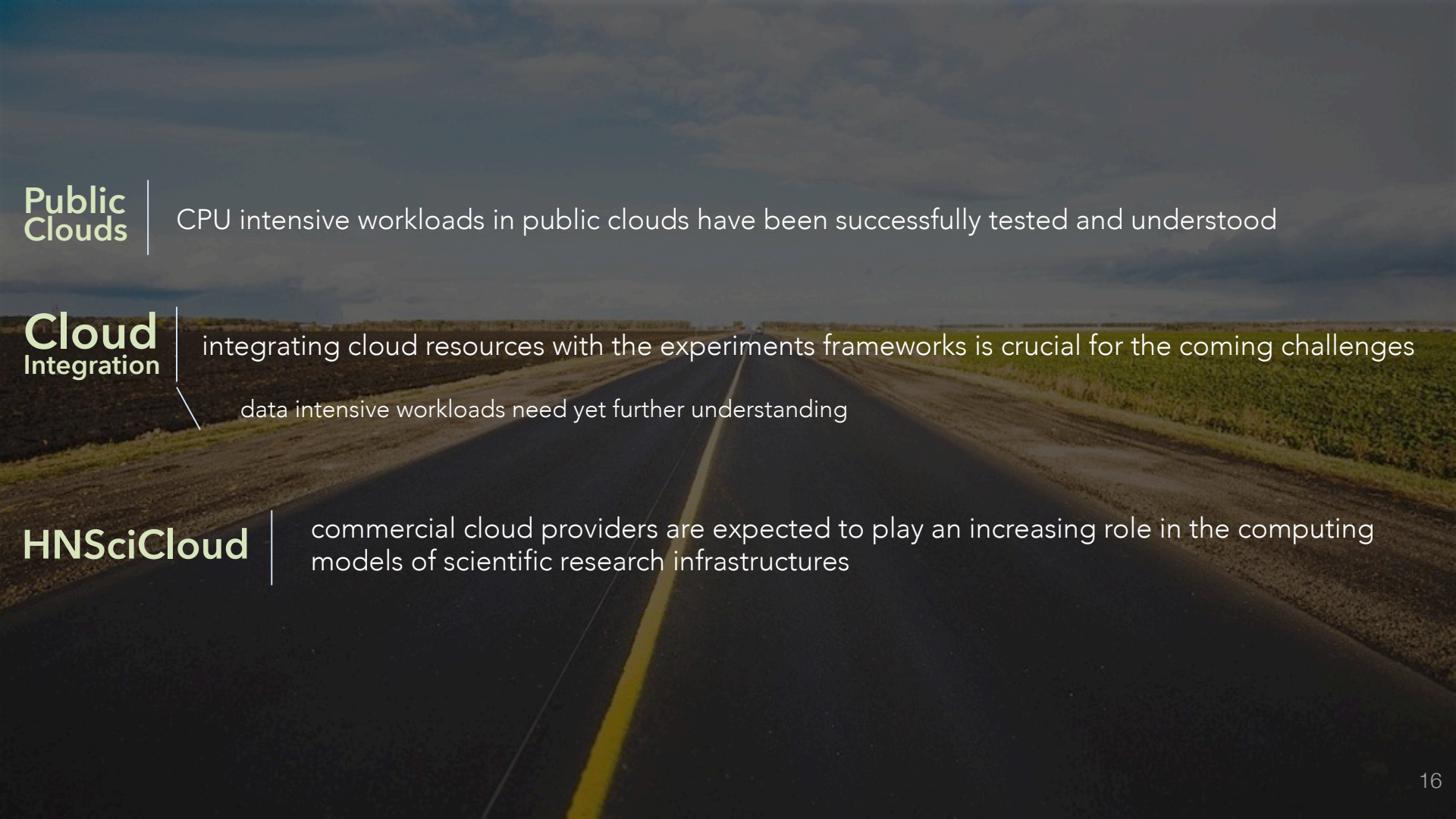
HELIX NEBULA
THESCIENCECLOUD

> *The Helix Nebula – the Science Cloud is a €5.3 million Pre-Commercial Procurement (PCP) tender for the establishment of a European hybrid cloud platform to support the deployment of high-performance computing and big-data capabilities for scientific research.*

- A collaborative commitment of:
  - procurement funds
  - manpower
  - use-cases
  - in-house IT resources

http://www.hnscicloud.eu/past-events

https://indico.cern.ch/event/505613/contributions/2230727/



Public Clouds | CPU intensive workloads in public clouds have been successfully tested and understood

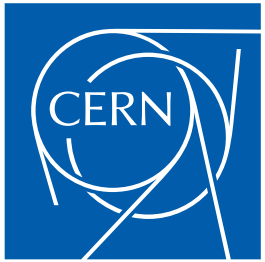
Cloud Integration | integrating cloud resources with the experiments frameworks is crucial for the coming challenges

- data intensive workloads need yet further understanding

HNSciCloud | commercial cloud providers are expected to play an increasing role in the computing models of scientific research infrastructures

References

- "HL-LHC Computing", M.Krzewicki, *Aix-Les-Bains HL-LHC Workshop*
- "Accessing commercial cloud resources within the European Helix Nebula cloud marketplace", G. Domenico et al, *2015 Journal of Physics: Conference Series **664** 022019*
- "CERN-IT evaluation of Microsoft Azure cloud IaaS", Giordano D., Cordeiro C., Di Girolamo A., Field L., Riahi H., Schovancova J., Valassi A., Villazon L. (2016, March 29). *Zenodo*. *http://doi.org/10.5281/zenodo.48495*
- "Benchmarking cloud resources", D. Giordano et al, *CHEP 2016 ID:28, https://indico.cern.ch/event/505613/contributions/2227960/*
- "Real-time complex event processing for cloud resources", C. Cordeiro et al, *CHEP 2016 ID:27, https://indico.cern.ch/event/505613/contributions/2230885/*
- "The HNSciCloud project", H. Meinhard, *CHEP 2016 ID:397, https://indico.cern.ch/event/505613/contributions/2230727/*
- "Future approach to tier-0 extension", B. Jones, G. McCance, *CHEP 2016 ID:22*


CERN