CERN

Computing infrastructure for LHC data processing

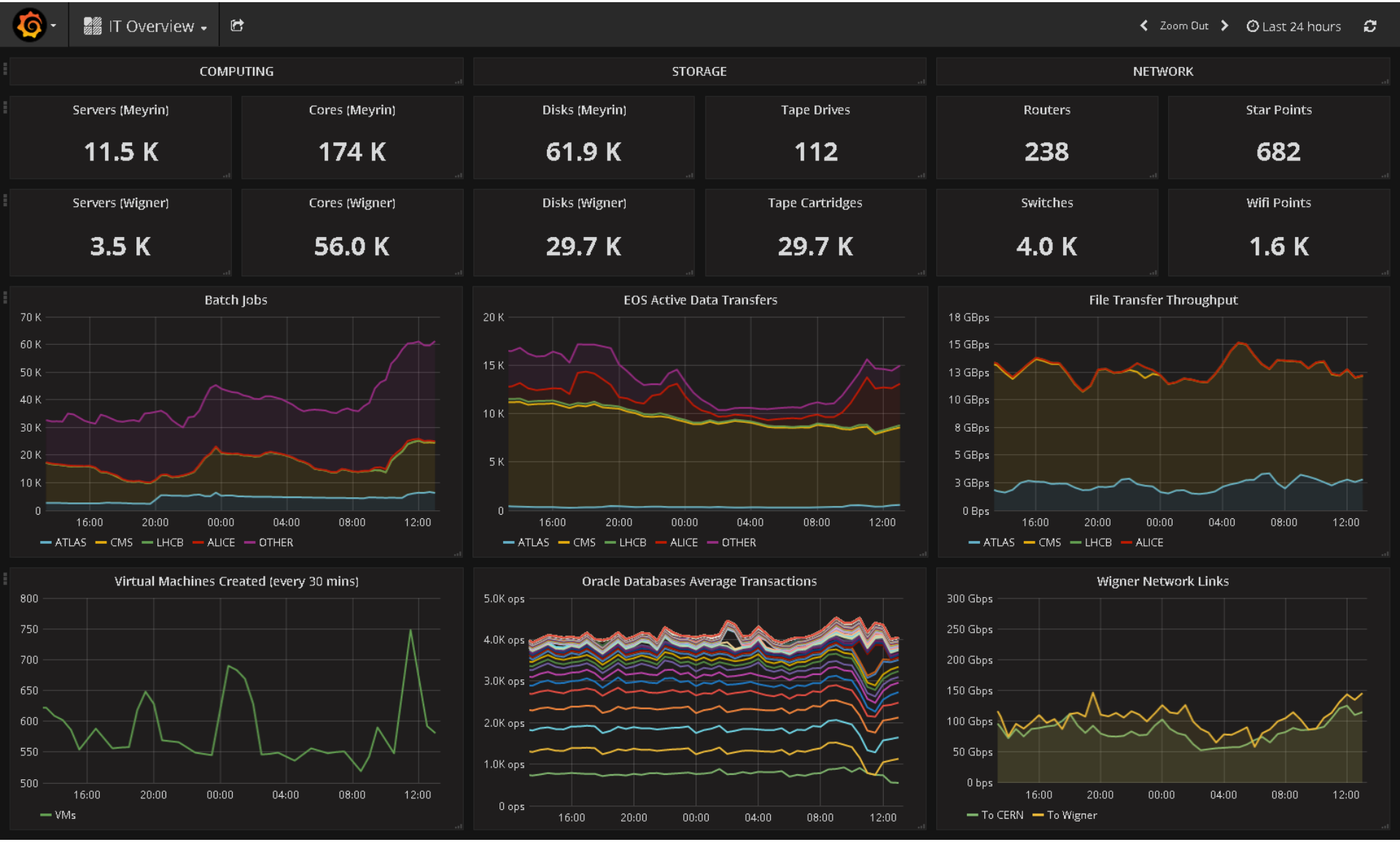
Vladimír Bahyl
CERN IT department

About the CERN IT Department

- Enable the laboratory to fulfill its mission
- Main data centre on Meyrin site (2/3)
- Wigner data centre in Budapest (since 2013 – 1/3)
- Connected via three dedicated 100Gbs links
- Where possible, resources at both sites (plus disaster recovery)

Drone footage of the CERN CC

Computing Centre in numbers



Services provided by IT department

CERN Service Portal

easy access to services at CERN

Search: type here

Home | News | Service Information | Navigate Catalogue | Contacts | My Profile | Site Guide

Service Availability Overview 05 Feb, 2018 13:22

Service Available | Service Degraded | Service Unavailable | Informational Message | No Information

Public screens view | Minimal view | FAQ

Batch Services
- Batch
- BOINC
- HPC

Collaboration Services
- Conference Rooms
- E-Mail
- Eduroam
- Lync
- Sharepoint

Computer Security Services
- Certificate Authority
- Single Sign On and Account Management

Data Analytics Services
- HADOOP

Database Services
- Accelerator Database
- Administration Database
- Database on Demand
- Database Replication
- Experiment Database
- General Purpose Database

Desktop Services
- Windows Desktop

Development Services
- Git
- JIRA
- SVN

Document Management Services
- CDS

Engineering Software Services
- Electronics Design Software
- Mathematics Software
- Mechanical Design Software

GRID Services
- File Transfer
- GRID Compute Element
- GRID Development
- GRID Information
- GRID Infrastructure Monitoring
- MyProxy
- VOMS

Infrastructure Application Services
- Indico Event Application Support

Interactive Services
- LXPLUS
- Windows Terminal Servers

IT Infrastructure Services
- ACRON
- Centralised Elasticsearch
- Configuration Management
- Linux Operating System
- Load Balancing
- Messaging
- Monitoring
- Server Provisioning

Network Services
- Campus Network
- CIXP
- Datacenter Network
- Network Database and Registration
- Network for Projects and Experiments
- Technical Network
- WIFI
- WLCG Network

Printing Services
- Printing and Copying

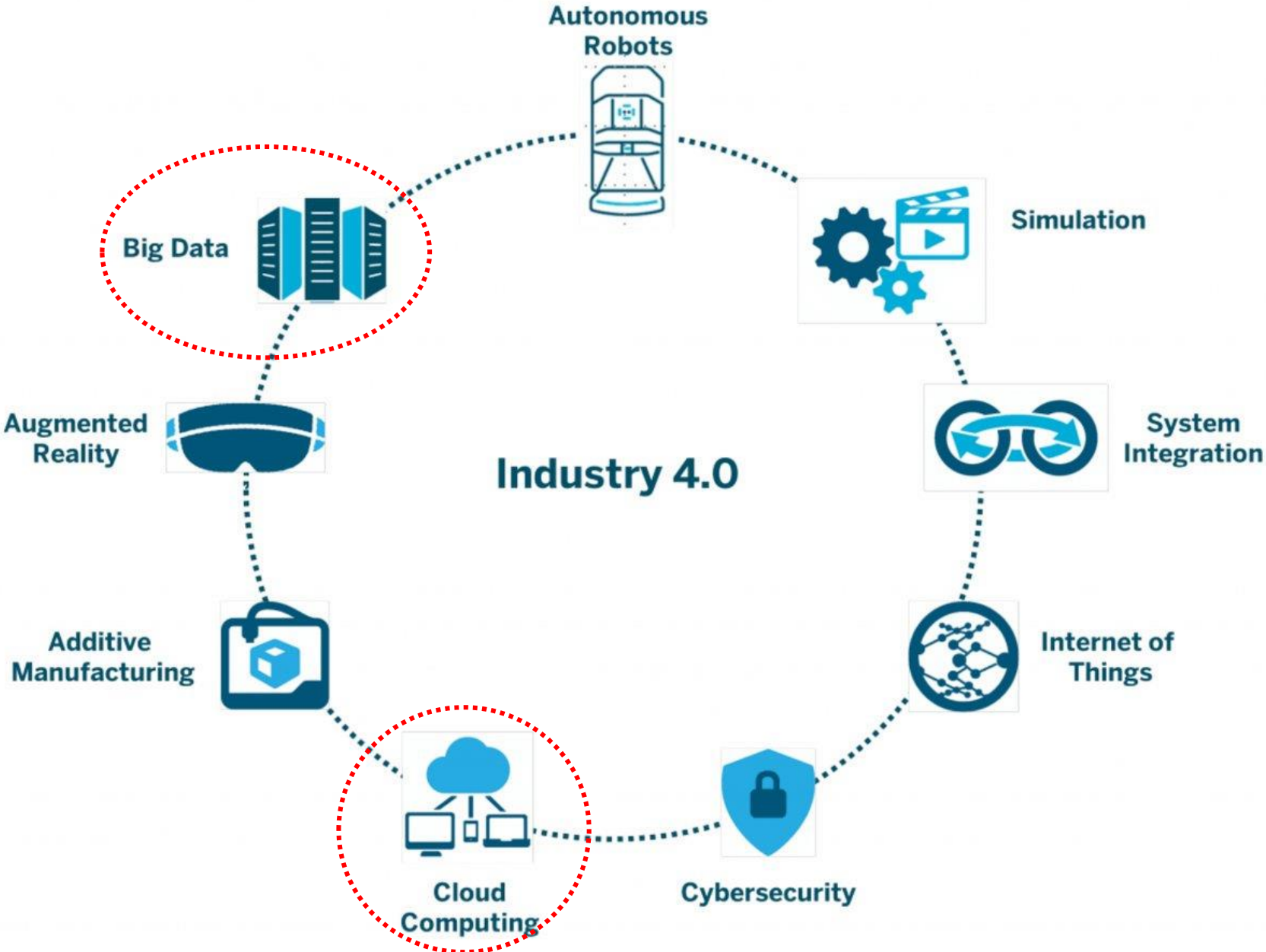
Storage Services
- AFS
- Backup and Restore
- CASTOR
- Ceph
- CERNBox
- CVMFS
- DFS
- EOS
- FILER

Telephone Services
- E-Fax
- Fixed Line Phone

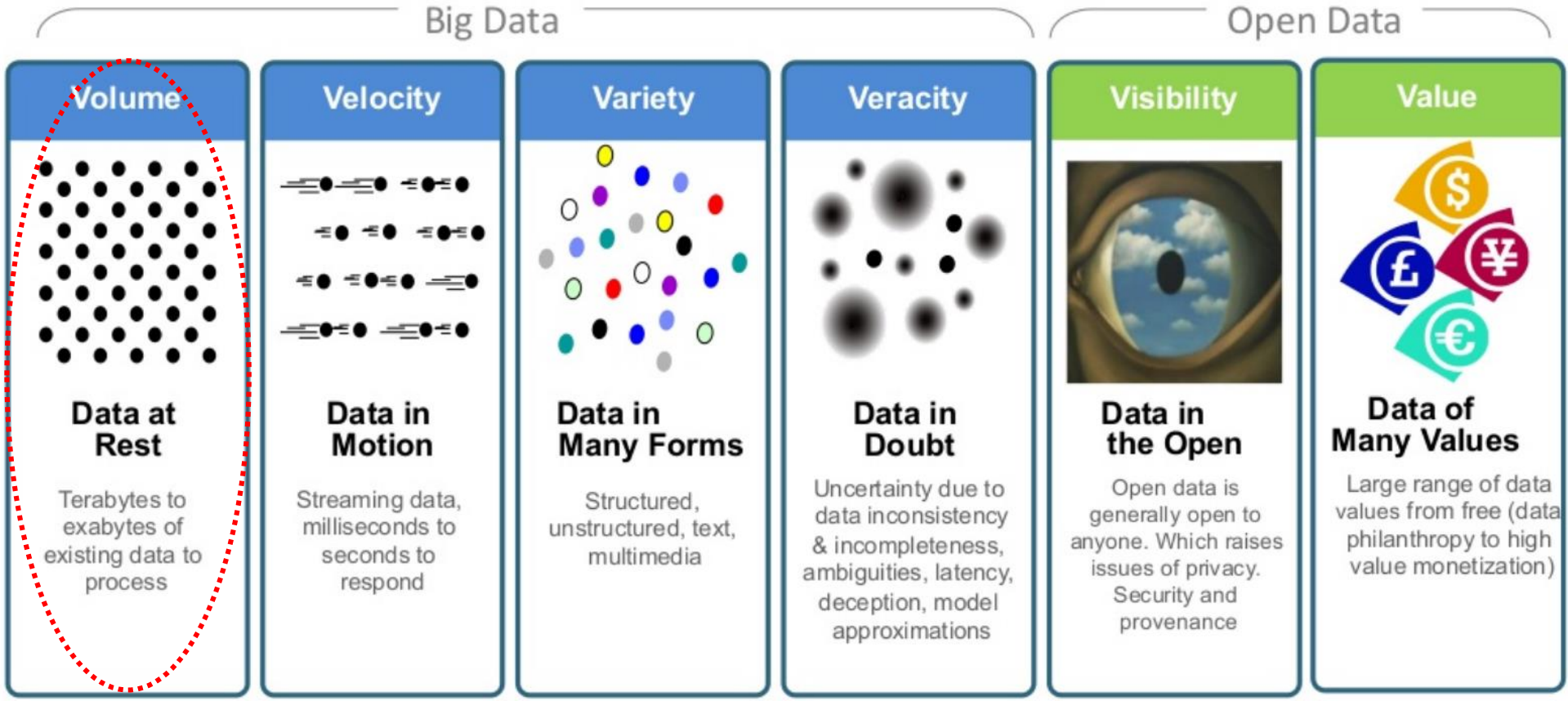
Text and Media Services
- Alerter
- MultiMedia
- Public Information Display

Web Services
- AFS Web Hosting
- CERN Search
- Databases Applications
- Drupal
- IIS Web Hosting
- PaaS Web Application Hosting
- Twiki

CERN IT & Industry 4.0



CERN IT & Big Data

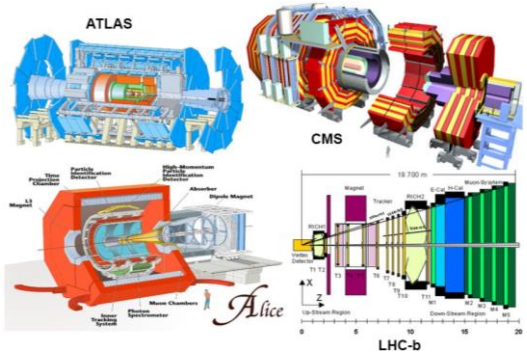


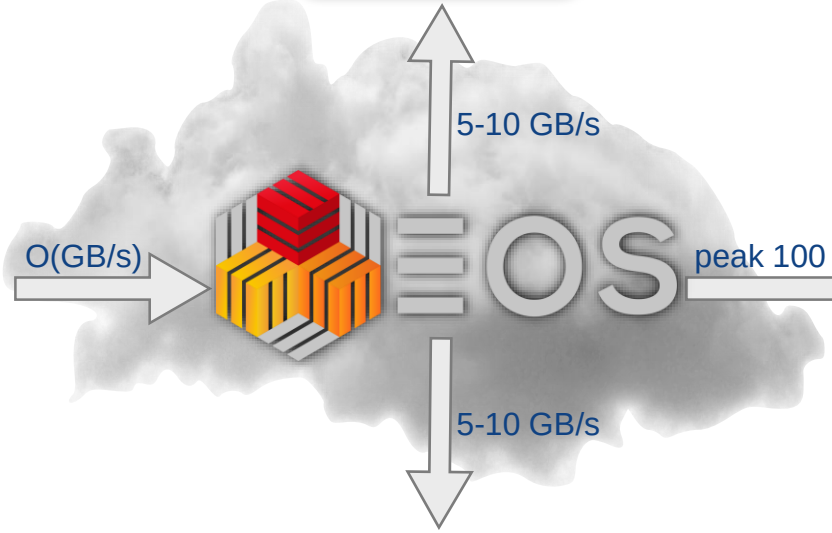
Big Data on Disk

Raw Disk capacity : 270 PB
Files stored : 2.4 billion

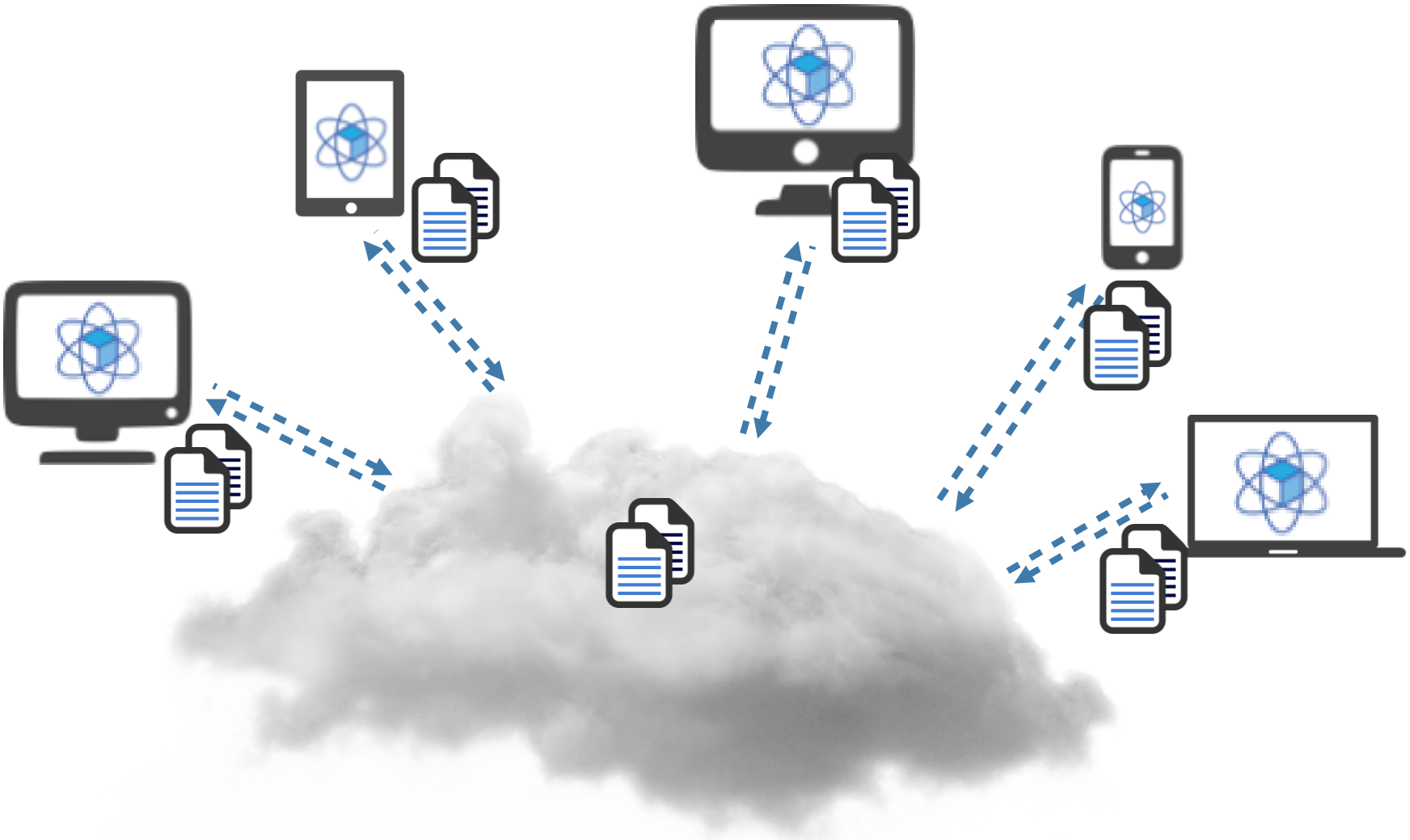
Tape Archive

LHC & Detectors





CERNBox



CERNBox powered by EOS

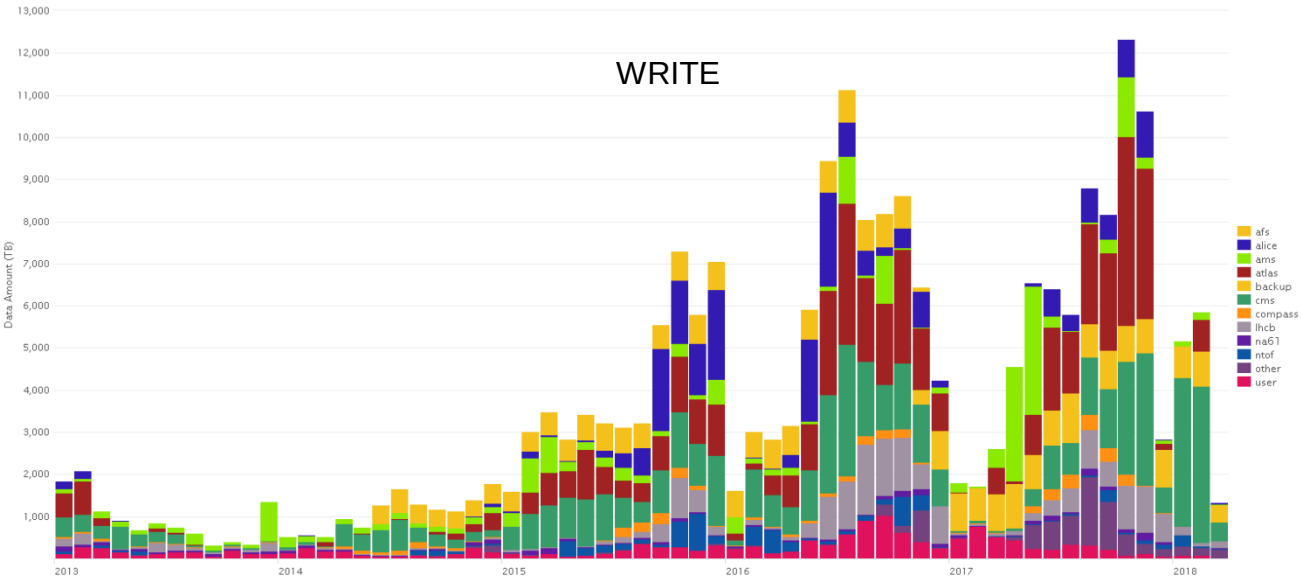
Big Data on Tape

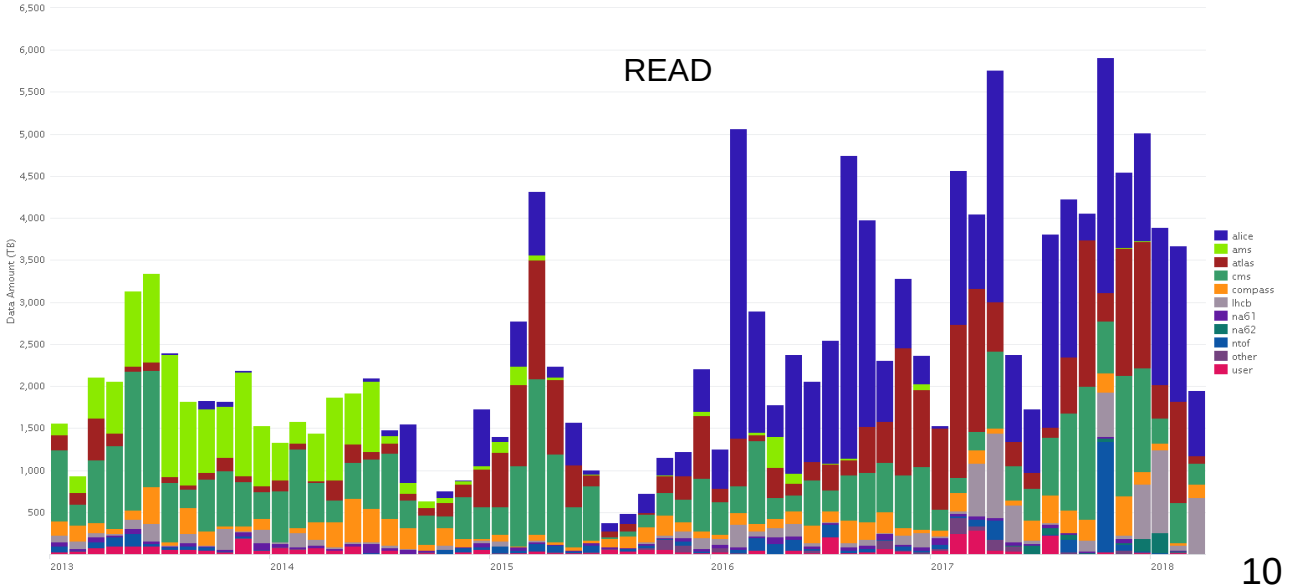
CASTOR archive:

- IBM
  - 1 x TS4500, 1 x TS3500
  - 46 x TS1155
  - 13000 x JD media (15 TB)
    6000 x JC media (7 TB)
- Oracle
  - 2 x SL8500
  - 20 x T10000D
  - 10000 x T2 media (8 TB)
- 10 PB disk cache
- ~220 PB of data on tape
  ~50 PB of free space
- Over 12 PB of new data per month
- Peaks of up to 8 GB/s to tape
- Lifetime of data: infinite

TSM backup:

- IBM
  - 2 x TS3500
  - 55 x TS1140
  - 200 x JC, 12000 x JB
- 8 PB; ~2300 M files
- 18 x TSM 7.1.4 servers





Big Data in the LHC Grid (HEP cloud)

CERN Tier-0 (Geneva & Budapest):

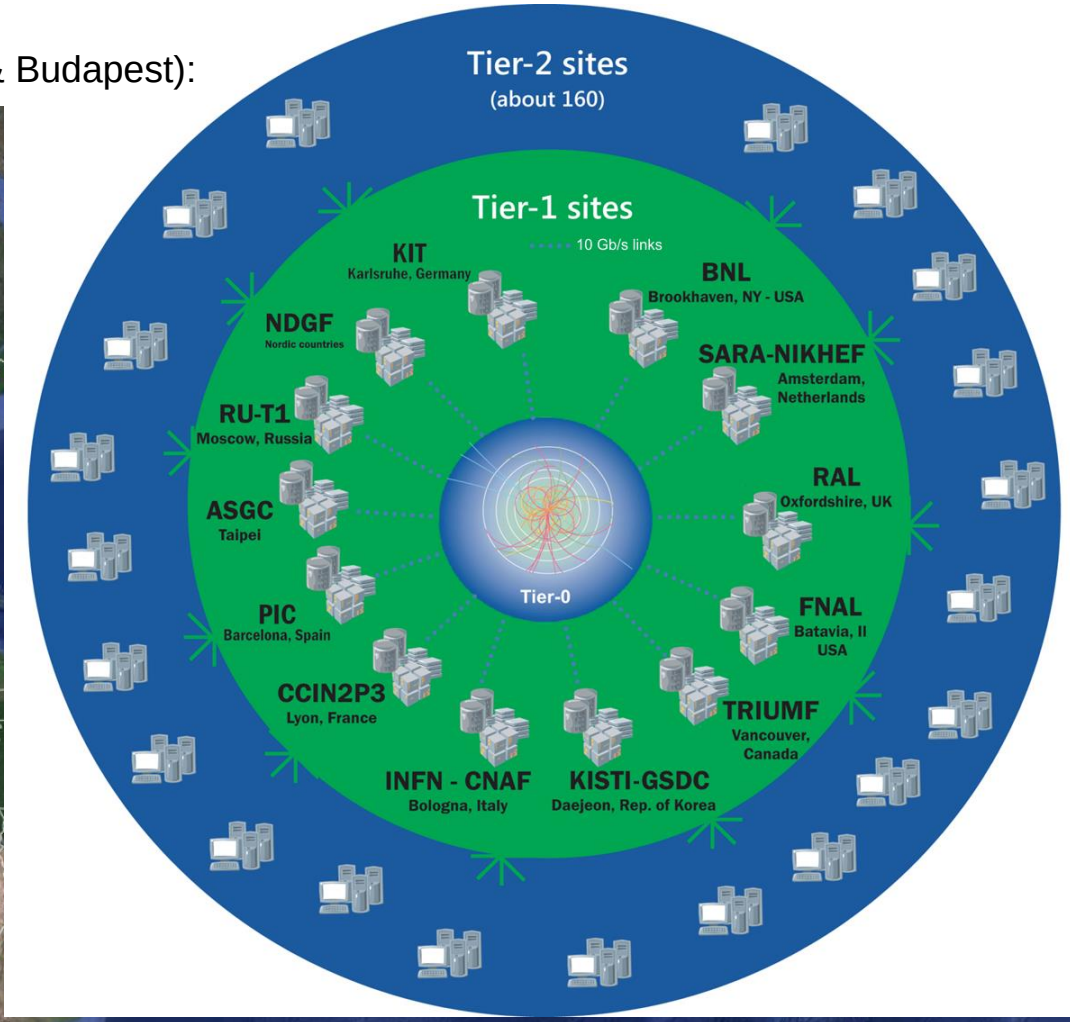
- Data recording
- Reconstruction
- Distribution

Tier-1:

- Permanent storage
- Reprocessing
- Analysis

Tier-2:

- Simulation
- End-user analysis



40 countries

~170 sites

~500k CPU cores

~500 PB of storage

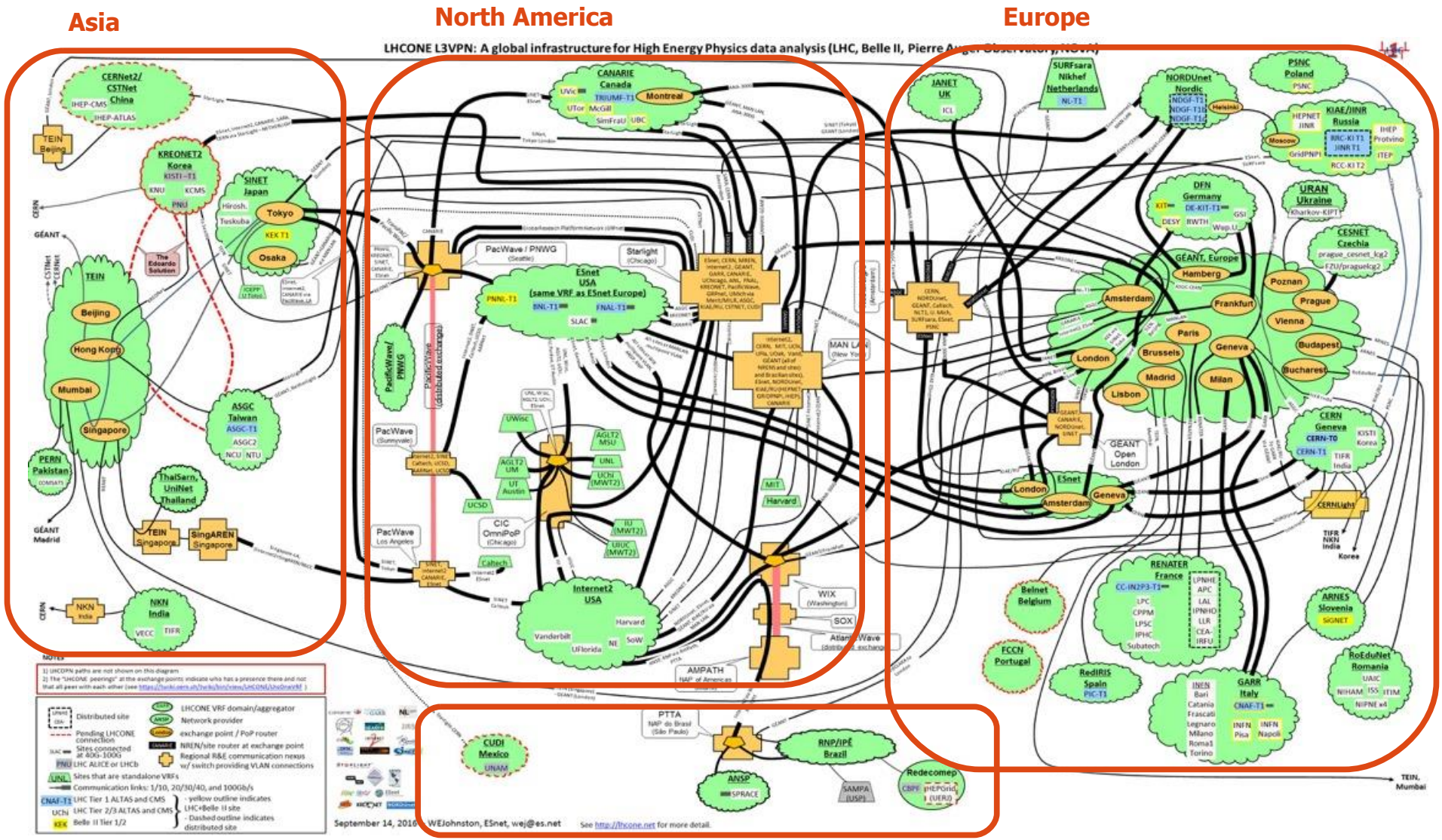
> 2 million jobs/day

10-100 Gb links

An international collaboration to distribute and analyse LHC data.

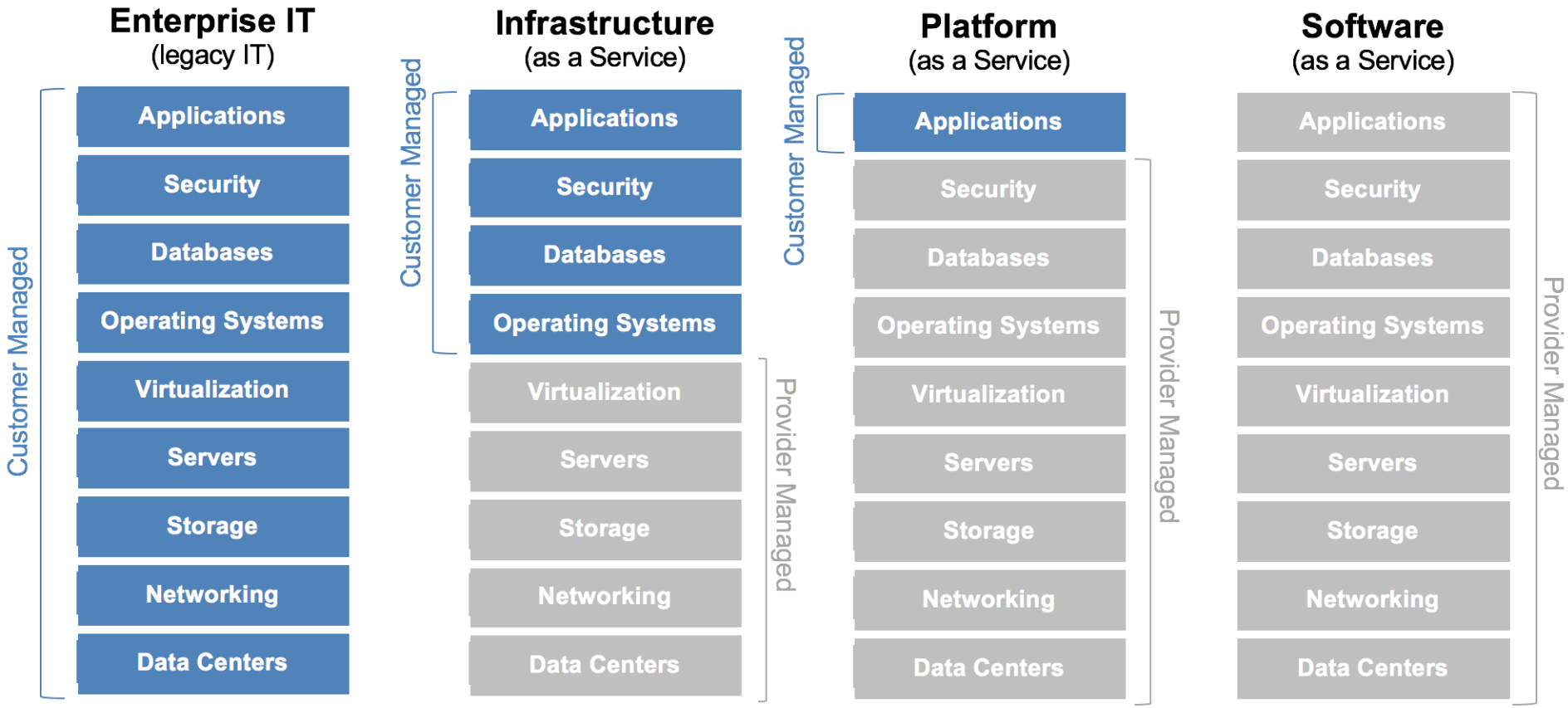
Integrates computer centres worldwide that provide computing and storage resource into a single infrastructure accessible by all LHC physicists.

Big Data on the Network



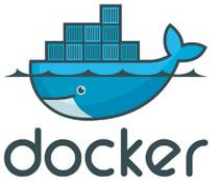
LHCOne: Overlay network that allows national network providers to manage HEP traffic on general purpose network.

Shift to (software) services



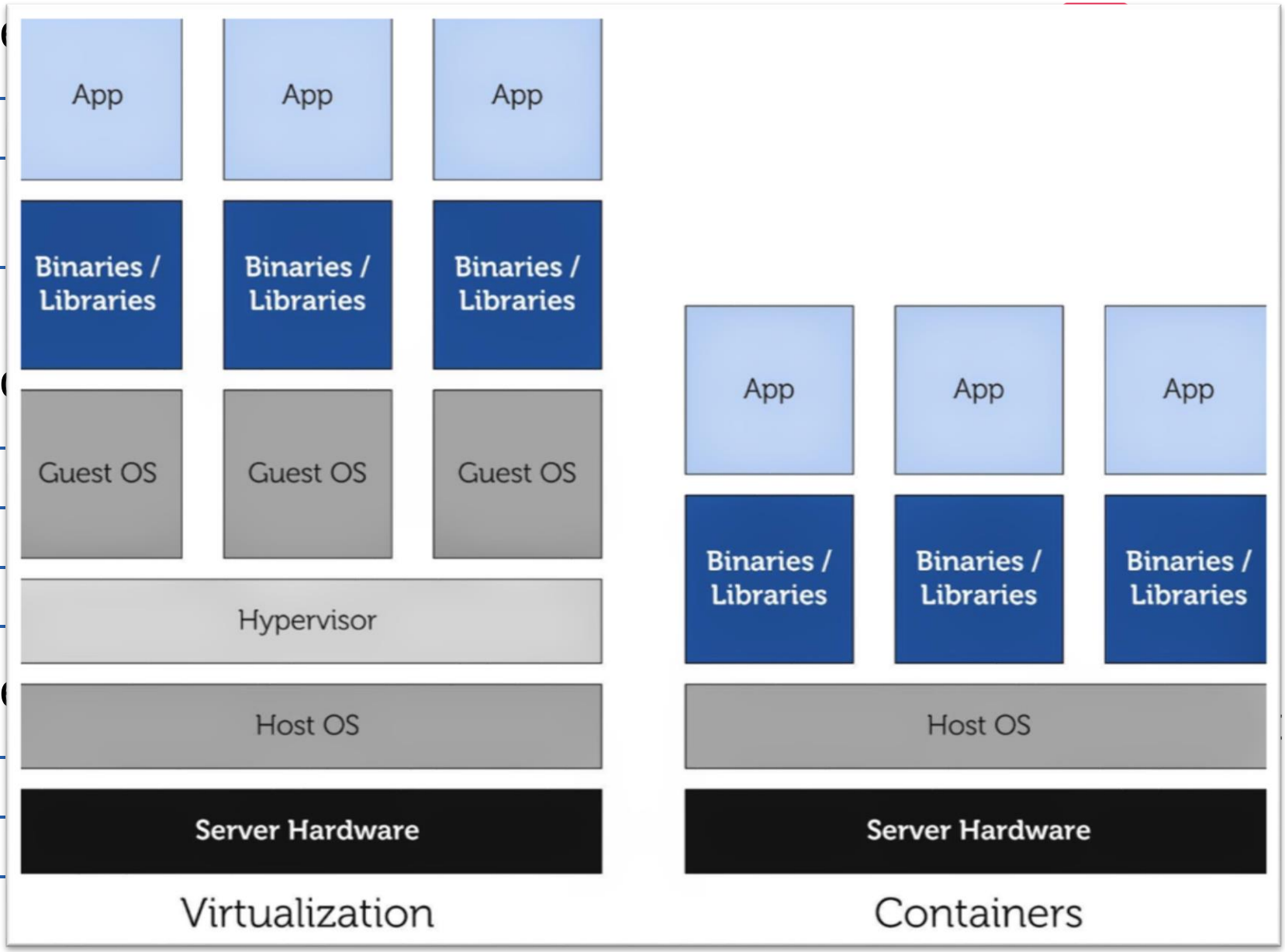
CERN IT follows the trends

- Reusing existing (open source) tools instead of developing everything in house.



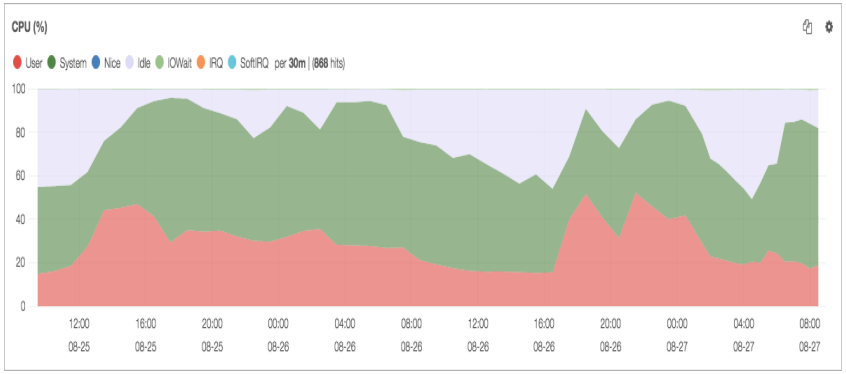
3 core (Agile) Infrastructure areas

- R
- C
- C

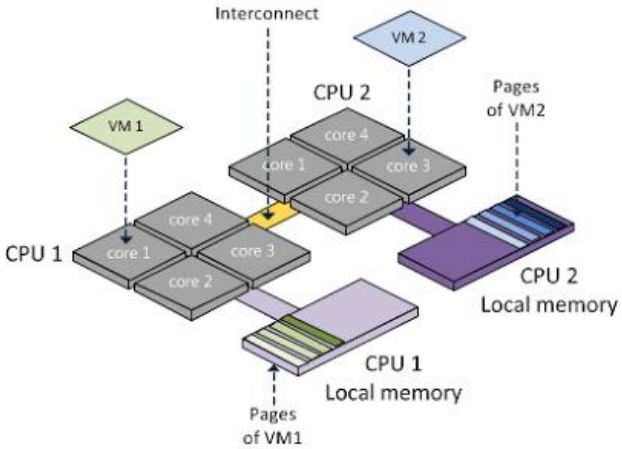


OpenStack CPU Performance: NUMA

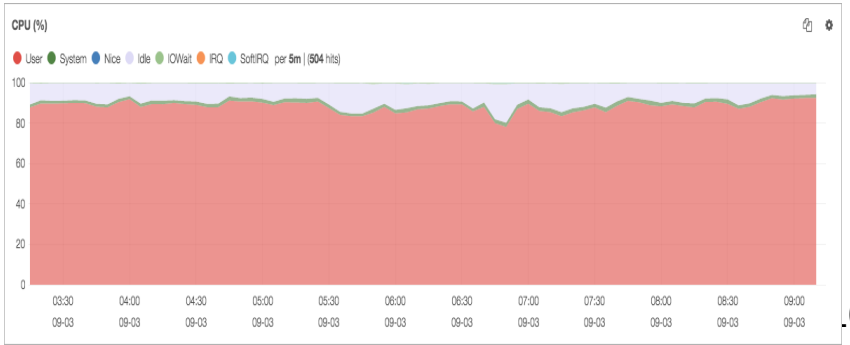
- The benchmarks on full-node VMs was about 20% lower than the one of the underlying host



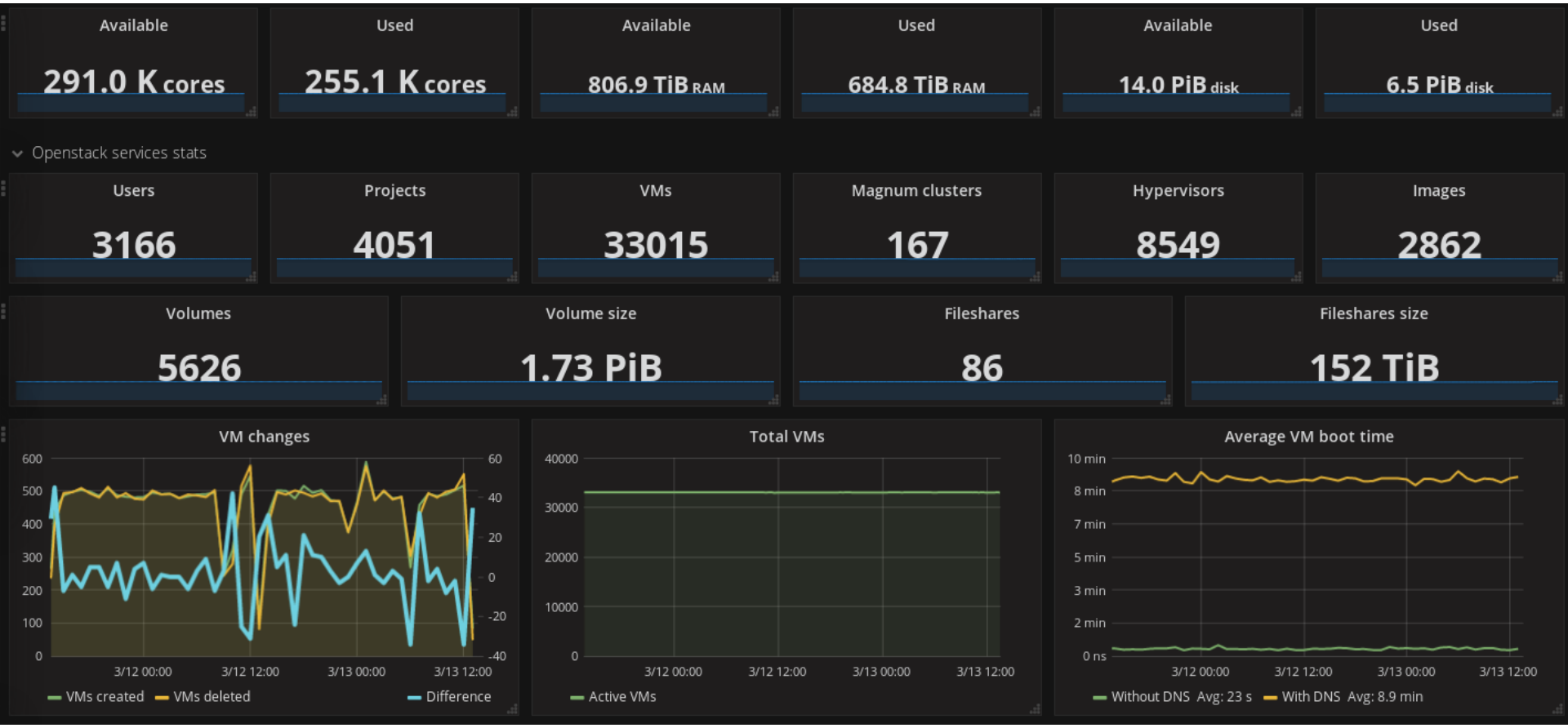
- Investigated various tuning options
- NUMA-awareness identified as most efficient setting



- Full node VMs have ~3% overhead in HEP-Spec06 benchmark



OpenStack in numbers



Configuration Management

- Client / Server architecture
  - ‘agents’ running on hosts plus horizontally scalable ‘masters’
- Desired state of hosts described in ‘manifests’
  - Simple, declarative language
  - ‘resource’ basic unit for system modeling, e.g. package or service
- ‘agent’ discovers system state using ‘facter’
  - Sends current system state to masters
- Master compiles data and manifests into ‘catalog’
  - Agent applies catalog on the host

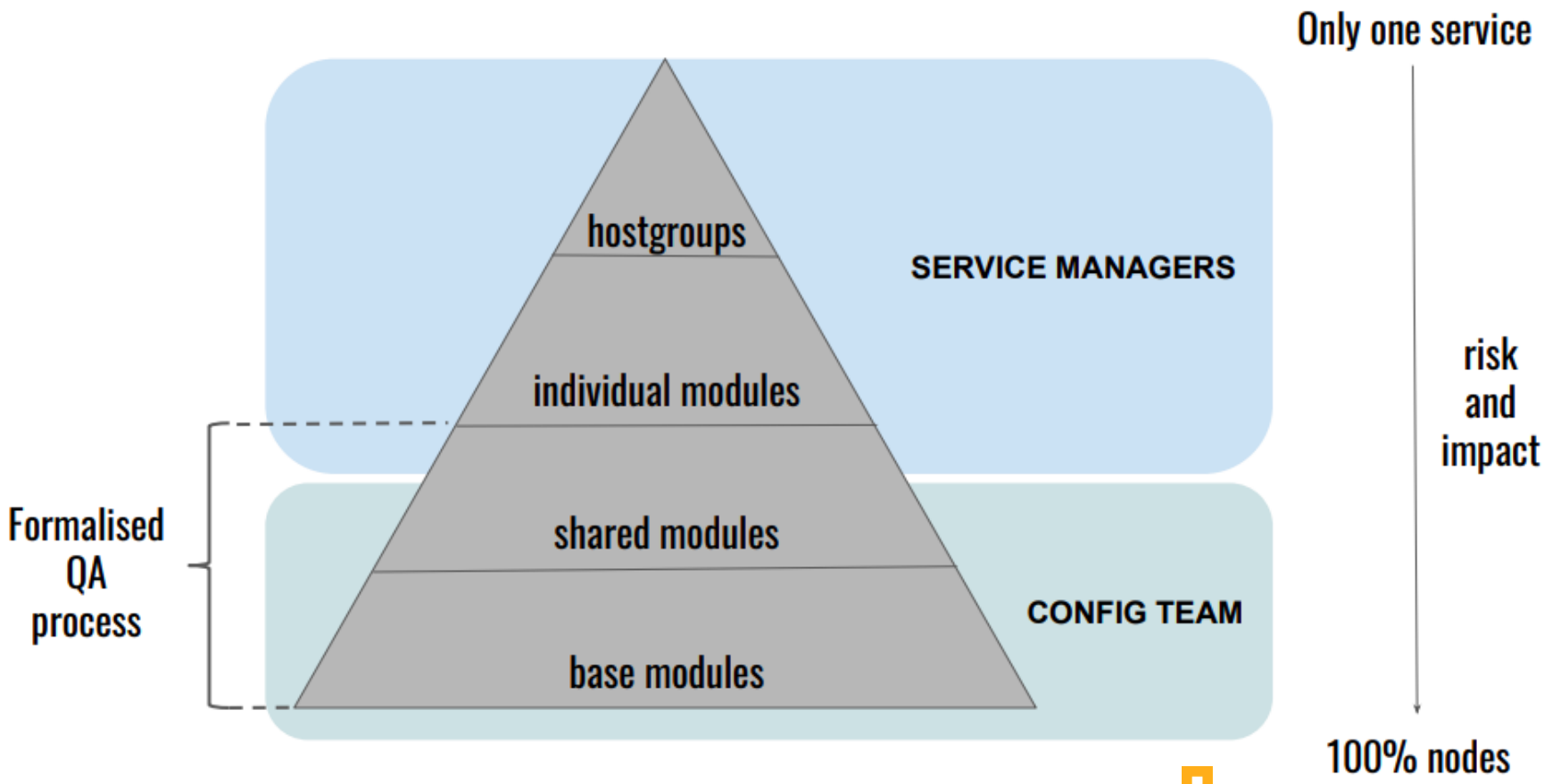
puppet

Status: Config' Management (1)

34k active nodes in PuppetDB	(virtual and physical, private and public cloud)
Base catalog contains **1.2k** resources	('base' is what every Puppet node gets)
350 catalogs compiled per minute	(compilations are spread out)
191 changes* per day	(this number includes dev changes)
350 puppeteers	(number Puppet code committers)

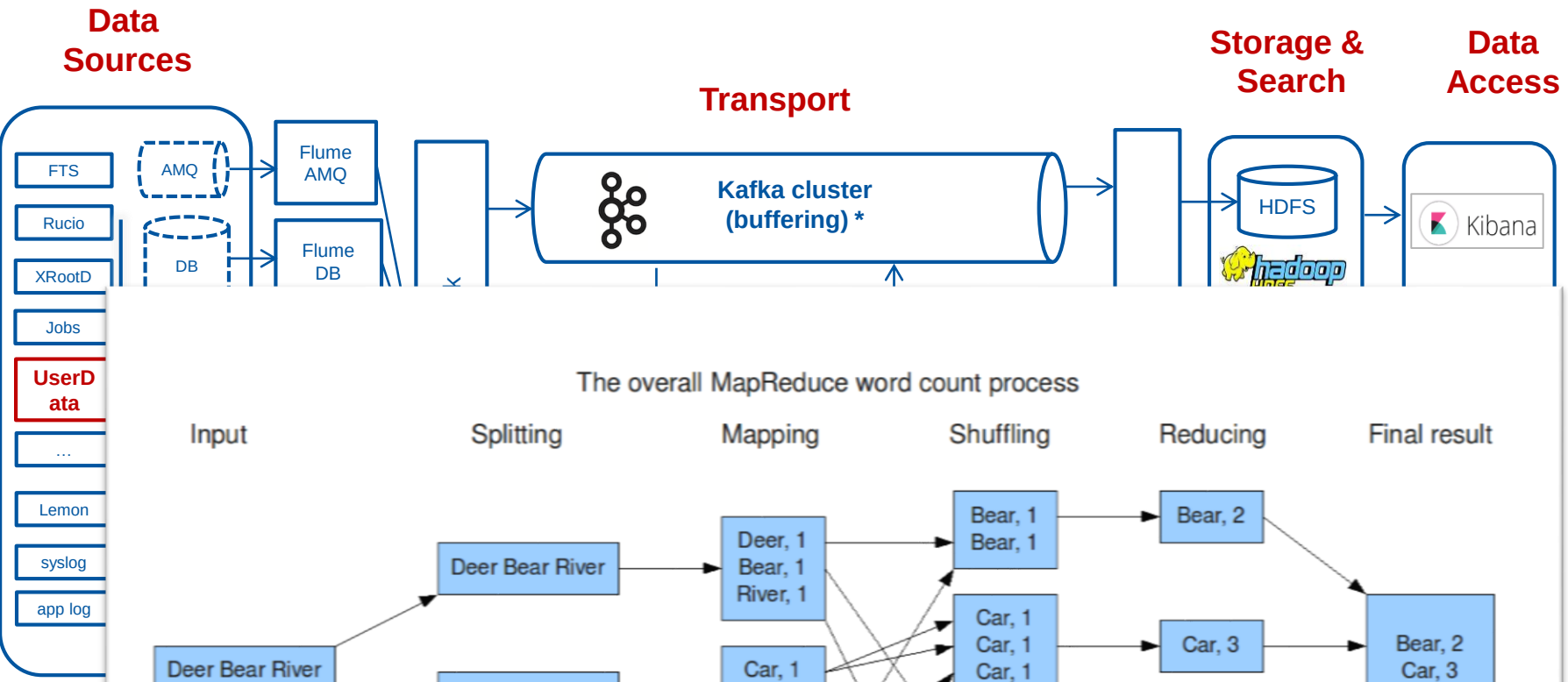
puppet

Status: Config' Management (2)



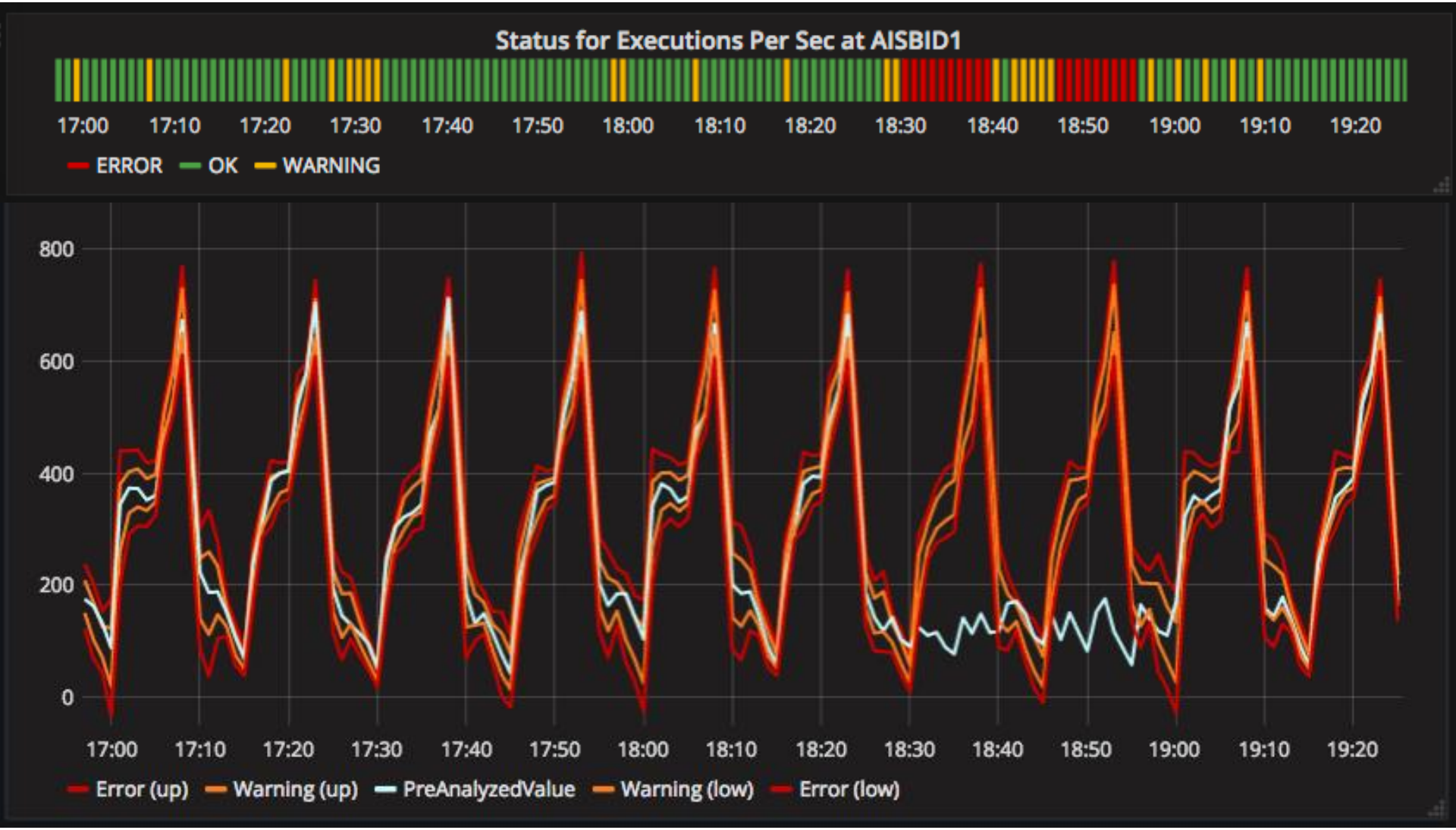
puppet

Monitoring infrastructure

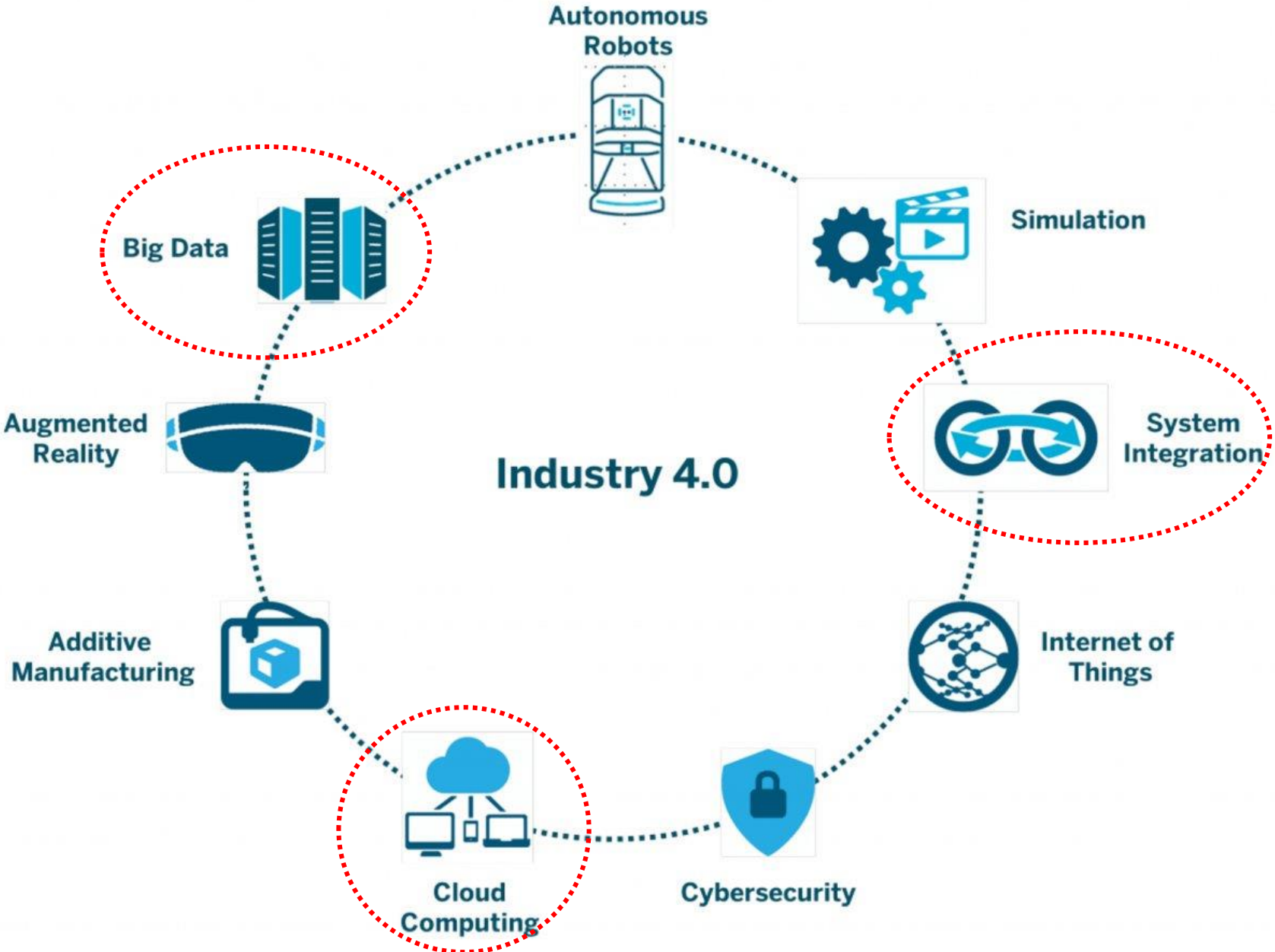


User Views

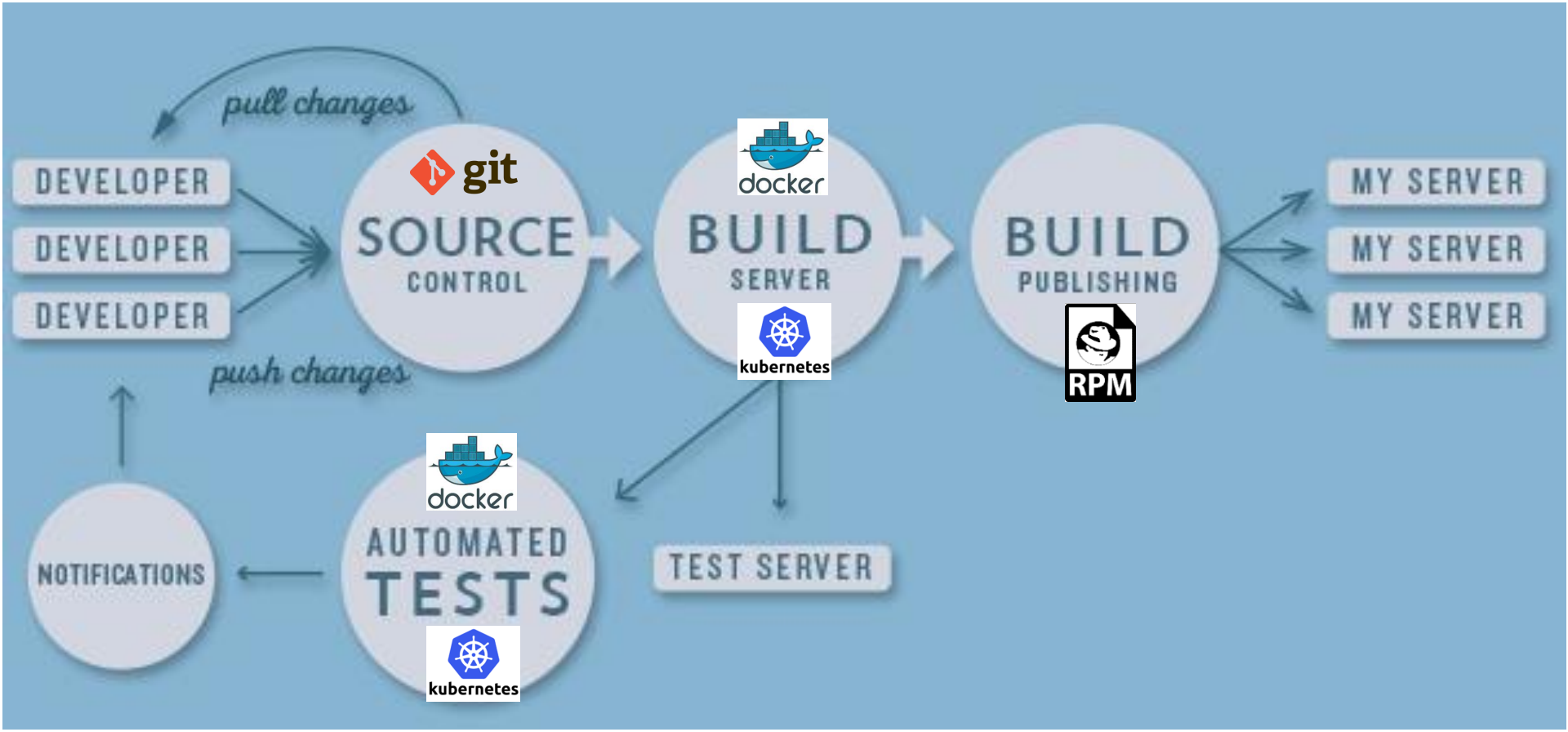
Metrics monitoring



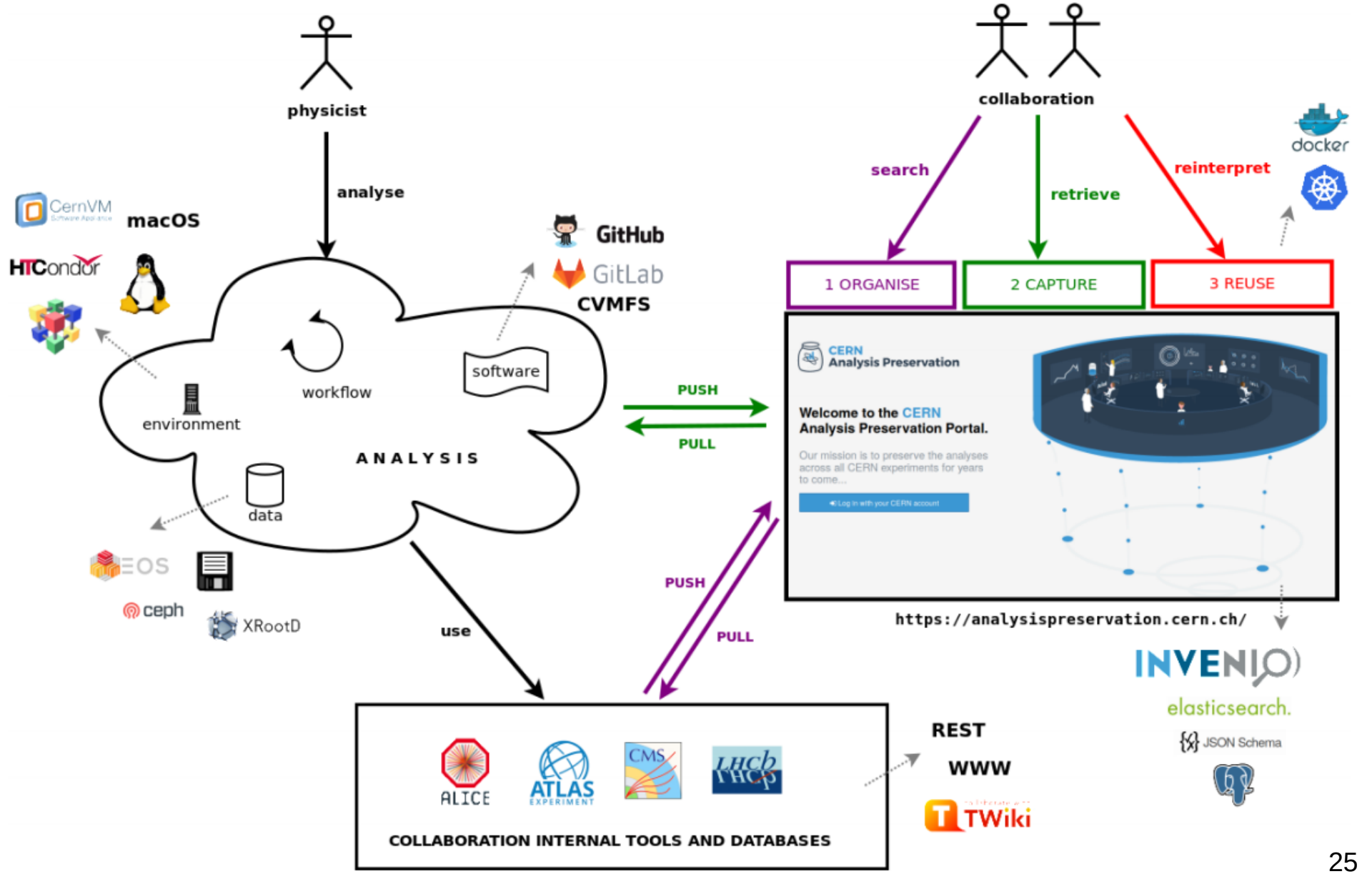
CERN IT & Industry 4.0



Continuous integration



Reusable Analysis



Summary

- Physics experiments at CERN produce large volumes of data
- CERN has performant IT infrastructure to:
  - Store large quantities of data on disk and tape
  - Analyze the data using private virtual cloud
  - Distribute the data over worldwide LHC computing grid
- Using open source software solutions wherever possible
  - We submit upstream all usefull changes we implemented
- Same building blocks can be used for data gathering and analysis by the Industry 4.0 processes