CERN Site Report

HEPiX Spring 2021

Andrei Dumitru
CERN IT Department
15th of March 2021

CERN

European Organisation for Nuclear Research
(**C**onseil **E**uropéen pour la **R**echerche **N**ucléaire)

Founded in 1954, 23 member states today

World's largest particle physics laboratory

Located at Franco-Swiss border near Geneva

$\approx$ 2500 staff members, $\approx$ 17000 users

CERN IT Department

Enabling the laboratory to fulfill its mission

Providing ICT services for the laboratory

Main data centre in Meyrin site

> 600 PB of data stored

COMPUTE

Servers
13.0 K

Cores
224.2 K

STORAGE

Disks
72.6 K

Tape Drives
93

NETWORK

Routers
319

Wifi Points
4.8 K

Batch Jobs Running

250 K
200 K
150 K
100 K
50 K

20:00 00:00 04:00 08:00 12:00 16:00

— alice — atlas — cms — lhcb — other
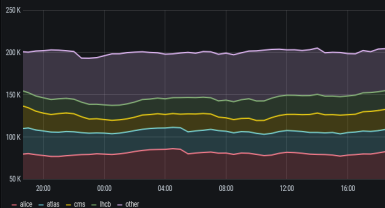
EOS Active Data Transfers

500 K
400 K
300 K
200 K
100 K
0

20:00 00:00 04:00 08:00 12:00 16:00

— alice — atlas — cms — lhcb — public
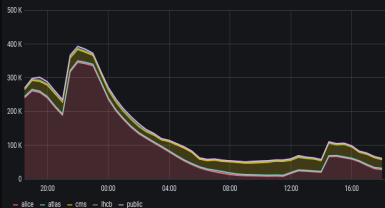
File Transfer Throughput

50 Gb/s
40 Gb/s
30 Gb/s
20 Gb/s
10 Gb/s
0 b/s

20:00 00:00 04:00 08:00 12:00 16:00

— lhcb — cms — atlas
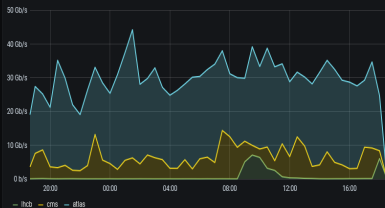
Cloud Virtual Machines Created

400
300
200
100

20:00 00:00 04:00 08:00 12:00 16:00

— vms

Databases Activity

7 K   200 K
6 K   175 K
5 K   150 K
                       125 K
4 K   100 K
3 K   75 K
2 K   50 K
1 K

20:00 00:00 04:00 08:00 12:00 16:00

— mysql (lps) — postgres) (lps)   — influxdb (pps)
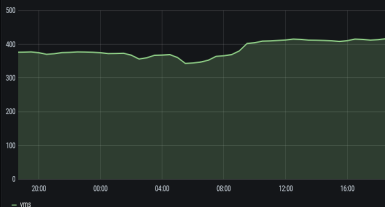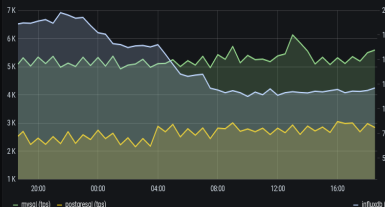
LHCOPN-LHCONE Total Traffic

300 Gb/s
200 Gb/s
100 Gb/s
0 b/s
-100 Gb/s
-200 Gb/s

20:00 00:00 04:00 08:00 12:00 16:00

— from cern — to cern
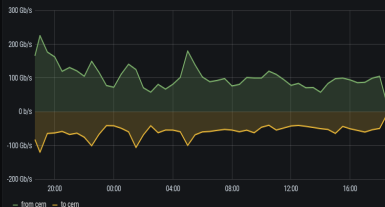
Computing Facilities

Tenders - prepared for adjudication in December 2020, incl. LHC experiments needs

- ▶ One for CPU servers for a capacity of 3.6MHS06
- ▶ One for external storage for a capacity of 300PB

CERN Prévessin Computer Centre (PCC)

- ▶ Tender completed, and approved at the December Finance Committee
- ▶ Final clarifications with the winning bidder before starting contract negotiations

Misc

- ▶ Benchmarking of new deliveries fully available in the Openstack/Ironic image
- ▶ OpenDCIM now in production - with DC monitoring
- ▶ Temperature monitoring deployed
- ▶ Ongoing refurbishment of 5 corridors (150 racks)
- ▶ Working on data privacy implementation in ServiceNow

This Wednesday

*HARRY: Aggregate hardware usage metrics too ptimise
procurement of computing resources
by Herve Rousseau*

Linux, Elasticsearch, Interactive Services

Linux

- Establishing support for ARM64 for system on a chip use cases
- Looking at Linux Futures in view of CentOS changes (more this week)

Elasticsearch

- ES5 and ES6 phaseout campaigns
- Evaluated OpenDistro and created Pilot instances
- Elastic License change as of 7.11 (SSPL 1.0) => Amazon/OpenDistro fork decisions
- Timelines of what we plan to provide to be presented by end March

Interactive Services

- LXPLUS8 facility for CentOS 8 interactive work
- LXTUNNEL facility (popular among teleworkers)
- Podman deployed in production
- Working on CentOS Stream 8 facility to be available in Q2-2021
- Access to GPUs for LXPLUS users: Evaluating options

Configuration, Cloud and Messaging

Configuration Management

- ▶ Migration to Puppet 6 concluded successfully
- ▶ PuppetDB v5 migration: Blocked on a performance issue (https://tickets.puppetlabs.com/browse/PDB-4830)
- ▶ Working on a Pilot Hashicorp VAULT for container secrets
- ▶ Support for CentOS Stream 8 and ARM64

Cloud and Containers

- ▶ Many service enhancements - details in dedicated talk this week

Messaging

- ▶ Usage continues to grow with now 46 brokers used in production
- ▶ Upgraded to the latest Open Source version of ActiveMQ 5.16.1

Monitoring and Experiment Test Framework

Monitoring

- ▶ Moved several accounting use cases (WLCG Transfers, ATLAS Job Monitoring, etc) to a new workflow based on Spark aggregation and Elasticsearch storage
- ▶ Completed the migration topology enrichment from VO Feed to CRIC for all WLCG and Experiments
- ▶ Introduced Service Levels Indicators (SLIs) and Service Level Objectives (SLOs) for the MONIT infrastructure and services

Experiment Test Framework

- ▶ ETF has been updated to Checkmk 1.6, details at https://indico.cern.ch/event/970604/
- ▶ New job submission and worker node testing frameworks are now in production
- ▶ DUNE and CMS (lightweight/opportunistic sites monitoring) are currently evaluating ETF

Compute Services

HTCondor farm is now over 280K cores

- ▶ New capacity will be installed as bare metal resources
- ▶ Tests running batch on Kubernetes continuing - details here
- ▶ Production Tests on Azure up to 7K cores including pre-emptible instances for lower cost

HPC SLURM now at 11K cores with Infiniband for Theory, HSE and Beams

- ▶ More capacity in 2021 and testing bursting to public clouds as budget allows

GPUs now available for various use cases

- ▶ Batch, notebooks, interactive, gitlab CI with some public cloud resources also available like TPUs

Validation of worker node on IPv6 mainly (details) with HEPiX IPv6 WG

- ▶ Some CERN services not yet dual stack, external services being checked

CentOS 8 services being added with FreeIPA to replace Active Directory and Tokens are being tested.

This Monday

Linux at CERN: current status and future
by Ben Morrice

This Tuesday

HEP Benchmarks: updates and demo
by Domenico Giordano

This Thursday

Anomaly Detection in the CERN Cloud Infrastructure
by Domenico Giordano

CERN Cloud Infrastructure status update
by Patrycja Gorniak

Database Services

Infrastructure and storage

- ▶ Major update on the infrastructure, servers and NetApp NAS storage.

Database on Demand

- ▶ High Availability: deployed a solution based on a master/replica setup with automatic switchover and ProxySQL for MySQL
- ▶ Automated testing of the recovery for databases, from existing backups in EOS

Oracle database service

- ▶ Upgrade and consolidation to have most services on version 18c and 19c

Java web service

- ▶ 90% of production deployments are now running in Apache Tomcat and Oracle WebLogic hosted in Kubernetes

Database Services

Reporting Pentaho service
- ► Improvements in instance deployment automation, scale-out clustering and access redundancy
- ► Integration with CERN SSO implemented

Hadoop, Spark, SWAN
- ► NXCALS (Hadoop based) is now the production for CERN Accelerator Logging. A lot of efforts on stabilising the platform (HBase especially)
- ► Created a Kubernetes implementation for SWAN (Jupyter notebook with the integration of CERN services), currently in beta
- ► Spark 3 deployed, both on Hadoop infrastructure and on Spark as a service, available from SWAN.

Streaming - Kafka
- ► Developed a multi-tenant solution to satisfy some of the use-cases
- ► CERN Kafka/streaming service is a building block of the LoRa/IoT (Internet of Things) used for example for the CERN proximeter

Storage - 2.5 Exabytes served by EOS Physics in 2020

EOS Physics - Monthly Data served

Legend:
- non-LHC
- LHCb
- CMS
- ATLAS
- ALICE
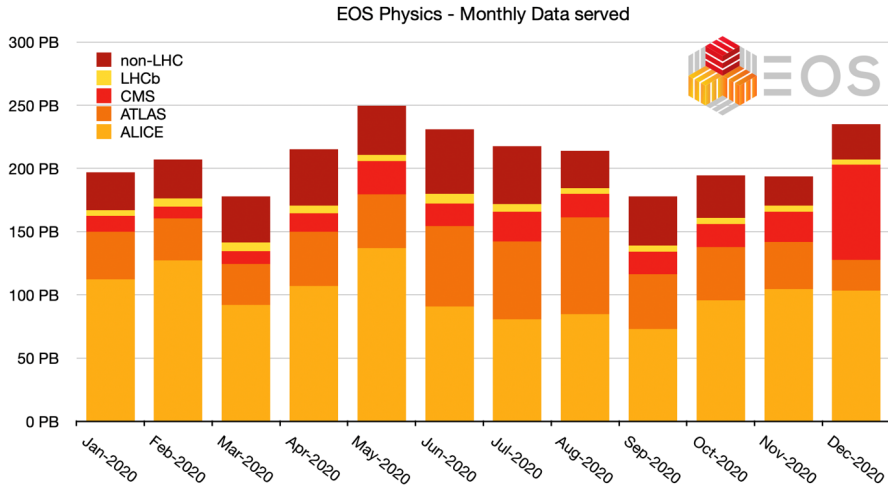
Storage - CEPH, CVMFS, CERNBox

Provision of new CEPH clusters and capacity

- ▶ for new high IOPS low-latency block storage with capacity-based IOPS (500 IOPS minumum + 5 IOPS/GB).

CVMFS prototype for efficient distribution of container images to worker nodes for data analysis

- ▶ Integration with Gitlab and Harbor services ongoing

CERNBox - simple, open document formats and apps

- ▶ Markdown editor integration under testing (CodiMD) - full integration soon
- ▶ Growing number of integrated applications - e.g. OnlyOffice, Collabora

Storage - SWAN, FTS, DTO, EOS

Data analysis facility with SWAN is under discussion with some T2s

- Jupyter fronted and full stack integration including storage, compute and software

FTS v3.10.0 released

- Almost all production instances upgraded to latest version

The first version of the Data Transfer Orchestrator (DTO) was released recently

- can be used by any VO to reliably transfer data from their DAQ to any storage

All EOS physics instances enabled with latest HTTP-TPC

Storage - ScienceMesh and conferences

ScienceMesh federation of sync and share services is growing

- ► Sites interested in joining the testbed early can get more information at cs3mesh4eosc.eu

Conferences and workshops

- ► 5th (virtual) EOS workshop (1-4 March) & CS3 2021 conference (25-28 January)
- ► Successfully tested the new interactive networking format with the Gather.Town platform
- ► Wider adoption of such tools may require to address possible data protection issues
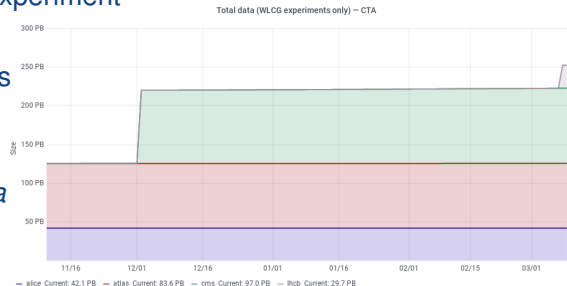
CERN Tape Archive (CTA) Deployment

Three new production CTA instances have been deployed

- ▶ CMS and LHCb instances have joined the existing Alice and Atlas services

A *CTA public* instance for smaller experiments and data preservation has been deployed

- ▶ The migration of CERNs smaller experiments from CASTOR to CTA started
- ▶ First data taking is underway - CAST experiment (data transferred from DAQ with FTS)
- ▶ Babar data has been imported as CTAs first data preservation use-case

CTA now manages over 250PB of tape data

Total data (WLCG experiments only) — CTA

— alice Current: 42.1 PB — atlas Current: 83.6 PB — cms Current: 97.0 PB — lhcb Current: 29.7 PB

CTA and Tape Infrastructure

Tape recall optimisation

- ▶ CTA demonstrated its support for a tape recall optimisation known as *RAO* (Recommended Access Order).
- ▶ Will improve efficiency in LTO drives which do not support this natively

Tape Infrastructure

- ▶ Planning of hardware for LHC Run 3 is advanced
- ▶ Procurements not yet launched, awaiting LTO-9 in particular
- ▶ Will run with minimum of 4 libraries and around 200 drives, split between Enterprise and LTO and accessed via SSD

Shared file systems strategy update

A working group was formed to have a forward look on the future of shared file systems at CERN

- The full report is now available: https://cds.cern.ch/record/2750122

Working Group recommendations were presented to IT management

- Currently there is no drop-in AFS replacement which can fulfill CERN's need at comparable service cost
- The OpenAFS community is healthy enough

The WG recommends

- To maintain the AFS service as Linux home directory for lxplus and lxbatch
- Continue reducing the dependency on AFS, CERN should move out of AFS as a long-term plan

Any new solution will require substantial effort from everyone in the community
Strategy to be reviewed towards the end of Run-3

This Wednesday

CTA production experience
by Julien Leduc

LHC Run 3 tape infrastructure plans
by Vladimir Bahyl

This Wednesday

XRootD5: what's in it for you?
by Michal Simon

Distribution of container images: From tiny deployments to
massive analysis on the grid
by Enrico Bocchi

Networking and Telephony services

Networking

- ▶ The additional 200Gbps to GEANT mentioned at the last HEPiX are now in service: we have 200Gbps for LHCONE + 200Gbps for Internet
- ▶ An additional 100Gbps LHCOPN link to DE-KIT has been deployed (for a total of 200Gbps)
- ▶ New Palo Alto firewalls connected to the Internet gate and being prepared for production (foreseen in April)
- ▶ Connections made to Azure datacentres in Amsterdam and Frankfurt

Telephony

- ▶ CERNphone mobile now considered in production
- ▶ Development of the CERNphone desktop client has restarted - pilot in April
- ▶ The *equivalent PLMN* pilot was a success
- ▶ Distribution of the new SIMs (which enable phones to switch to an Orange cell as easily as to another Swisscom one) is expected to start in April.

This Tuesday

CERN central DHCP service: Migration from ISC DHCP to Kea
by Maria Hrabosova

This Thursday

CERNphone update
by German Cancio

Collaboration services

Videoconference
- ▶ New production service based on Zoom
- ▶ Vidyo service retired

Email
- ▶ Ongoing pilot based on Dovecot/OpenXChange

Webhosting
- ▶ Moved ~2500 webeos sites out of SLC6 to modern container-based hosting environment (November 2020)
- ▶ Started the pilot of the new generation of Platform-as-a-Service web application hosting, based on OKD4 (February 2021)

Digital Repositories

InvenioRDM

- ► Research Data Management repository built with 20 world-wide partners (based on Zenodo success)
- ► Version v1.0 released which includes access control to resources, indexing, searching, deposition, APIs, among other features.

CERN Document Server

- ► CERN Library Catalogue (literature loans) re-developed using latest Invenio Technology (InvenioILS)
- ► Now entering deploying into production phase; see sandbox

CERN Open Data

- ► Datasets now easily downloadable via CLI using cernopendata-client
- ► CMS heavy-ion physics data release (210TB); support for ORCID and ROR identifiers

Digital Repositories

REANA (Reusable Analysis)

- ▶ New releases: 0.7.0, 0.7.1, 0.7.2
- ▶ Deploy via Helm charts; user resource quotas, hybrid workflows (HTCondor, Kubernetes, Slurm)
- ▶ New blog: https://blog.reana.io/

Digital Memory

- ▶ Working on long term perservation platform based on Archivematica
- ▶ Analogue content (6.000 videos, 420.000 images, 8.700 audio) digitised and partially uploaded to CDS

This Monday

OnlyOffice and Collabora Online Experience at CERN
by Maria Alandes Pradillo

Windows desktop service from computer to user centric IT
by Sebastien Dellabella

Building up and migrating to a FOSS-focused e-mail service at
CERN
by Vincent Brillault

This Thursday

CERN Authentication and Authorization
by Paolo Tedesco

The CERN-Solid code investigation project
by Jan Schill (IT University of Copenhagen)
and Maria Dimou (CERN)

Computer Security

This Tuesday

*Computer Security Update
by Liviu Vâlsan*

Talks from CERN this week

Talks from CERN this week

WLCG Authorization WG Update

- ► Hannah Short

XRootD5: what's in it for you?

- ► by Michal Simon

Anomaly Detection in the CERN Cloud Infrastructure

- ► by Domenico Giordano

LHC Run 3 tape infrastructure plans

- ► by Vladimir Bahyl

Linux at CERN: current status and future

- ► by Ben Morrice

The WLCG HEP-SCORE deployment task force

- ► by Helge Meinhard

Talks from CERN this week

IPv6-only on WLCG - update from the IPv6 working group

- by Andrea Sciabà

Computer Security Update

- by Liviu Vâlsan

CERN central DHCP service: Migration from ISCDHCP to Kea

- by Maria Hrabosova

OnlyOffice and Collabora Online Experience at CERN

- by Maria Alandes Pradillo

Distribution of container images: From tiny deployments to massive analysis on the grid

- by Enrico Bocchi

CERN Authentication and Authorization

- by Paolo Tedesco

Talks from CERN this week

HARRY: Aggregate hardware usage metrics too ptimise procurement of computing resources
- ▶ by Herve Rousseau

Windows desktop service from computer to user centric IT
- ▶ by Sebastien Dellabella

Building up and migrating to a FOSS-focused e-mail service at CERN
- ▶ by Vincent Brillault

CERNphone update
- ▶ by German Cancio

CERN Cloud Infrastructure status update
- ▶ by Patrycja Gorniak

CTA production experience
- ▶ by Julien Leduc

The CERN-Solid code investigation project
- ▶ by Jan Schill (IT University of Copenhagen) and Maria Dimou (CERN)

home.cern