CERN site report
HEPiX Autumn 2023

16-20 October, Victoria, Canada

Jarek Polok <jaroslaw.polok@cern.ch>
On behalf of CERN IT

CERN Data Centre(s)

Prévessin Data Centre (PDC):

- **The project is on time (and budget). Civil engineering works completed.**
- **Commissioning testing ongoing since few months.**
- **Integrated Systems Tests of full phase 1 (4 MW) capacity using loadbanks (industrial heaters) in progress.**
- **Final trial operation with servers in one pod (0.5 MW) planned for December - concluding the acceptance tests: PDC will enter operation phase after.**
- **10 years maintenance and (infrastructure) operation contract starts immediately after completing the acceptance.**

Hardware deliveries:

- **Physics batch processing (PDC) received and ready for deployment: 280x quads → 1'120 compute nodes → 143'350 SMT cores (AMD Milan) → 2'380MHS06**
- **Data storage: 46x quads → 184 compute nodes → 11'776 SMT cores (AMD Milan)**
- **360x JBODs mostly for EOS→ 8'640 hard drives → 155'520 TB**
- **Business Continuity and Disaster Recovery (PDC): ordered - to be deployed in PDC**

Other:

- **Reduction of 24x7 data center operator coverage (night shift suppression) implemented.**

October 2023
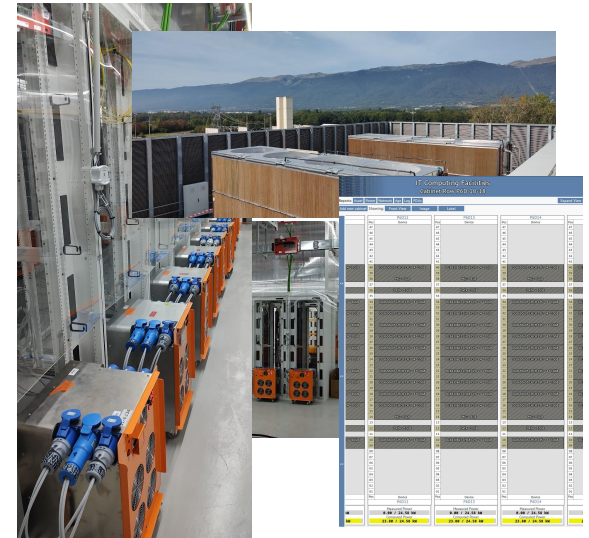
Telephony

Final stages of the Alcatel PABX decommissioning: Two issues to address

- **Easy: phone line migration**
  - Less than 500 active phone lines - progressive migration of users.
  - 4500 inactive phone lines decommissioned 1st of July 2023.

DONE

- **Hard(er): replacing the Call Centre functionality**
  - Pilot in-house development in operation for reception and CERN guards.
  - On track to migrate Service Desk, CERN Control Center and the Fire Brigade.

END OCTOBER

- **Still expect to complete Alcatel decommissioning by the end of the year.**

The end of a nearly 10-year project to migrate to SIP, Asterisk and softphones!
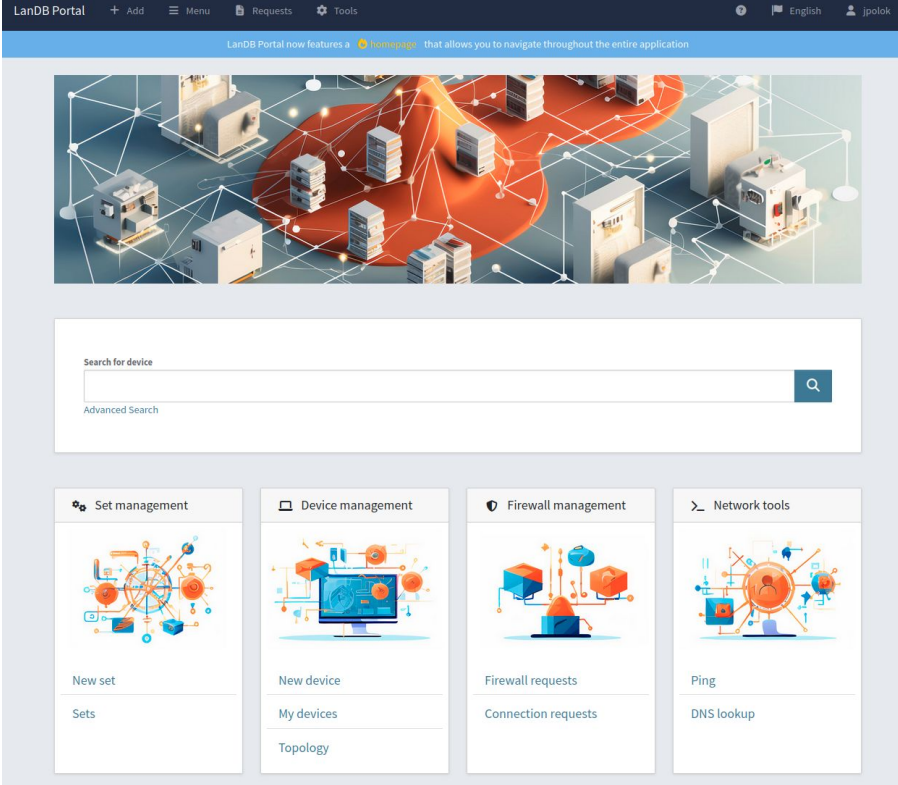
Networking

Changes to the CERN Wi-Fi SSID

- **Now restricted to people with CERN accounts**
- **Visitors must use the CERN-Visitor SSID**
- **The option for requesting access to the CERN SSID has been removed.**

A pilot for WPA3 userid/password authentication instead of today's MAC-based authentication will start in IT before the end of the year.

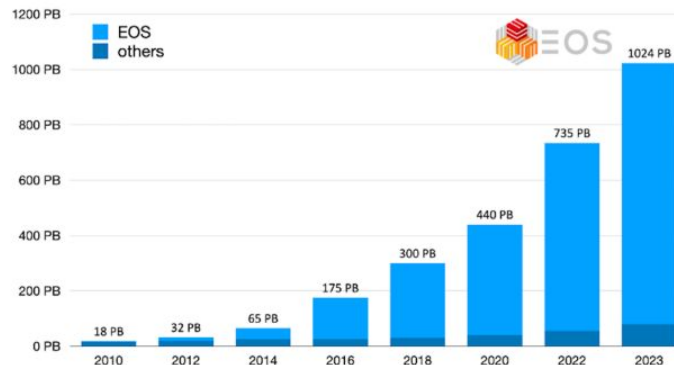- **If successful, it will likely be rolled out across CERN during 2024.**

Storage Technologies

An Exabyte of disk storage at CERN:

    This data capacity is provided using 111 000 devices, predominantly hard disks along with an increasing fraction of flash drives.
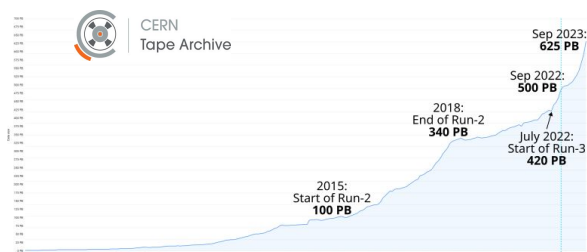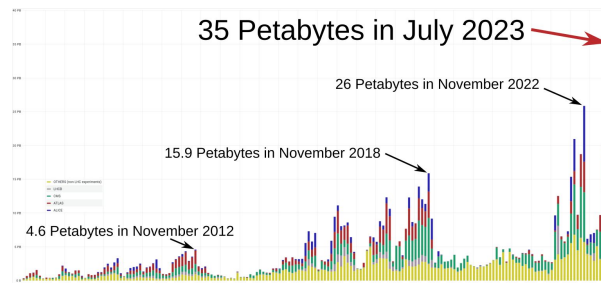
    Major part of the capacity is managed by EOS.

    https://home.cern/news/news/computing/exabyte-disk-storage-cern

Tape Storage:

    New CERN Record - 35 PB (July 2023)

New CERN Record
Data archived to tape storage each month since 2008

35 Petabytes in July 2023

26 Petabytes in November 2022

15.9 Petabytes in November 2018

4.6 Petabytes in November 2012

IT Department | Storage and Data Management Group

Database and Analytics

OpenSearch deployment

- **Completed the migration of ElasticSearch v6 and v7 clusters to OpenSearch.**
- **Still missing the migration of OpenDistro clusters (different ES distribution) to OpenSearch.**
- **Migrated all OpenSearch and OpenDistro clusters to AlmaLinux 9.**
- **Enabled the k-NN plugin for users interested in VectorDB search use cases.**

Hadoop disaster recovery

- **Running for several years daily incremental HDFS backups stored on tapes in CTA.**
- **Missing DR for HBase and Zookeeper. Added hourly snapshots for these tools.**
- **Snapshots copied to external HDFS clusters for further protection.**
- **Implementation based on native Hadoop and HBase tools.**

Evolution of Monitoring infrastructure at CERN

- **Pilot of Prometheus-based metrics service has been open.**
- **Update to Grafana 10.**

Email service

Migration to Exchange Online
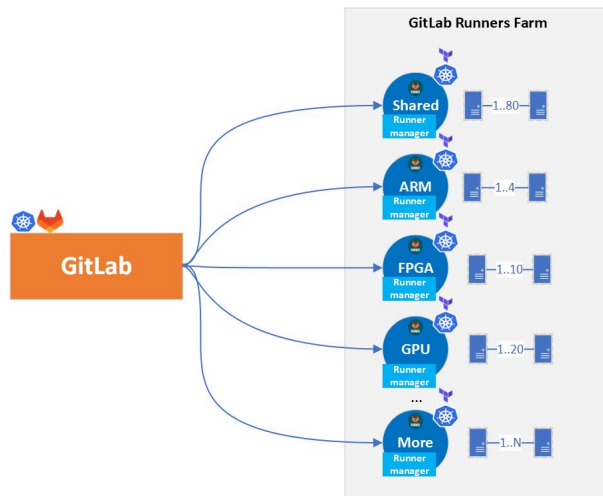
- **completed as planned end of June 2023**

xorlab

xorlab ActiveGuard gradually deployed as a complementary protection measure to Microsoft Exchange Online Protection and Microsoft Defender for Office.

GitLab - New GitLab Runners Architecture

- **New GitLab Runners deployment in place**
- **Docker+machine executor runners cleaning campaign about to finish (ETA October 2023):**
  - Deprecation of the current GitLab Runners based on Docker+machine executor in favor of Kubernetes+executor
  - Taking the opportunity for users to modernize their workflows.
    - Encourage users to use non privileged approaches.
    - Enforce users to use more secured workflows.
    - Security Policies ready to be used.
- **New Architecture for GitLab Runners**
  - GitLab Runners clusters decoupled from the GitLab installation cluster.
  - Several benefits:
    - **Scalability**: GitLab runners can be scaled up independently of the GitLab installation.
    - **Reliability**: If one of the GitLab runners goes down, there are several replicas to get into scene, not impacting the availability of the GitLab installation cluster.
    - **Security**: GitLab runners are isolated from the GitLab installation cluster, and only accessible from the intranet, making them less of a target for attackers.
    - **Flexibility**: Each GitLab Runner serves a specific purpose, and it can be interchanged and scaled up/down upon needs.
- **Automated deployment using Terraform and GitLab**
- **Road to a major GitLab Runners ecosystem**
  - CVMFS, ARM, Spark, TN, etc.

GitLab Runners Farm

GitLab

Shared Runner manager 1..80

ARM Runner manager 1..4

FPGA Runner manager 1..10

GPU Runner manager 1..20

More Runner manager 1..N

Keycloak - Upgrade and new Architecture

- **Upgrade to version 20.0.5**
  - Sensible improvements in performance
- **Switched from VMs Architecture to Kubernetes one**
  - Why Kubernetes:
    - **Red Hat direction**: Provides a Kubernetes Operator to simplify deployment and replaced Jboss with Quarkus, Java framework designed to run on Kubernetes
    - **Portable**: The only thing you need is a Kubernetes cluster
    - **Reproducible and Immutable:** Speeds up operations, reducing team effort
      - Increased automation and traceability thanks to GitOps approach
    - **Easier to maintain and deploy in long term**
      - Vibrant community supporting Kubernetes
      - Single maintainer for the Keycloak Puppet module

  - Separation of Infinispan cache from Keycloak servers benefits
    - **Monitoring**: Possibility to monitor Keycloak and Infinispan separately
    - **Scalability**: Keycloak and Infinispan scale differently
    - **Flexible deployment model**: keep VMs for Infinispan (but planning to move to K8s as well)
    - **Operations**: User sessions are persisted across Keycloak restarts, helping transparency of future interventions
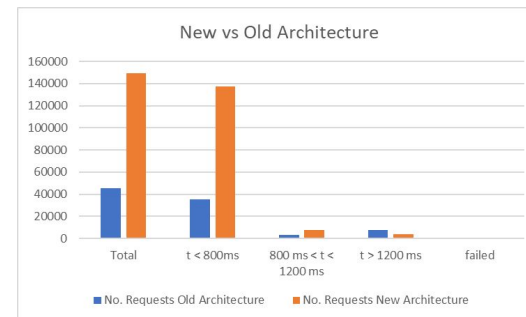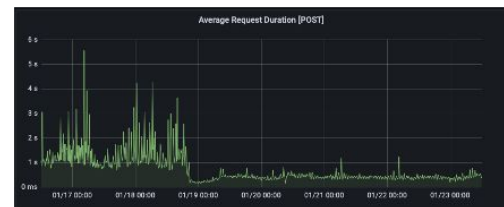
Kubernetes Service

- **Latest releases: 1.26 and 1.27**
  - SeccompDefault now stable, in-place update of Pod resources, provisioning of cross-namespace snapshots, group snapshots
- **Service Highlights**
  - Registry: support for artifact signing with cosign, early support for generation and storage of SBOMs (Software Bill Of Materials)
  - Runtime: support for ARM nodes, including mixed deployments. Very extensive series of blog posts on GPU sharing (see also talk at kubecon)
  - Security: new oauth2-refresh-controller for auto refresh of OAuth2 credentials in a cluster - useful for long lived jobs, EOS access
  - Storage: csi-cvmfs with automount support, csi-eosxd as a replacement of the old eosxd daemonset (deprecated) and including automount and support for OAuth2 based authentication
- **Recent Webinars:** https://indico.cern.ch/category/16678/
  - CVMFS and EOS updates, Full Automation with ArgoCD
- **Major Effort: Kubernetes Service in the TN, Air-Gapped**
  - Collaboration with Accelerator Sector, aiming production Q1 2024
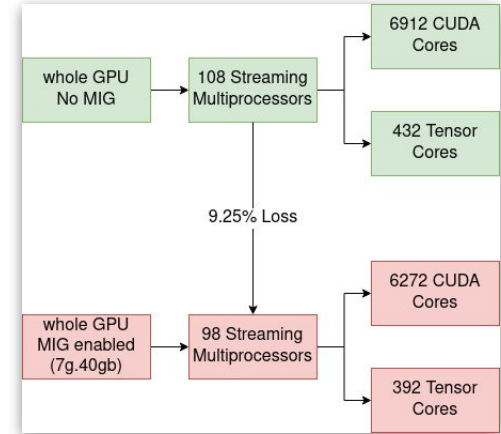
CERN Open Data

- **New open data release**
  - CMS completes Run1 heavy ion open data collection. New release contains about 560 TB of 2013 and 2015 proton-proton reference data, proton-lead collision data and simulations. See the [release announcement](#).
- **LHCb Open Data Ntupling Service**
  - Ongoing integration of LHCb Ntuple Wizard to allow theorists/phenomenologists to ask for custom open data production. Alpha testing phase expected in Q4 2023.
- **"Continuous Reuse"**
  - A new system to periodically execute open data usage examples to ensure future reusability. Testing file presence, log presence, runtime durations. See the [PV2023 talk](#).
- **Forthcoming disk/tape infrastructure updates**
  - Started investigation of a mixed disk/tape storage backend solution to save storage costs for the 4 PB of open data.
    - E.g. 10% data files on disk, 90% on tapes.

REANA reusable analyses

- **REANA 0.9.1 was released on September 28th 2023.**
  - User-oriented features
    - Visualise PDF and ROOT files in the web interface.
    - Purge temporary files after workflow execution.
    - Improved Snakemake between-run caching of results.
  - Admin-oriented features
    - Support any Keycloak IAM. Used by ESCAPE, PUNCH4NFDI.
    - Support Kubernetes clusters from 1.21 to 1.28.
    - Configurable site gallery of launch examples.
    - Auto-closure of inactive interactive notebook sessions.
    - Numerous performance and stability improvements.
  - … and much more. See the [full release notes](#).
- **First results from ATLAS Run2 pMSSM searches using REANA.**
  - Running O(10k) workflows to cover enough pMSSM model points.
- **Forthcoming**
  - Sharing of workflow runs within research teams.
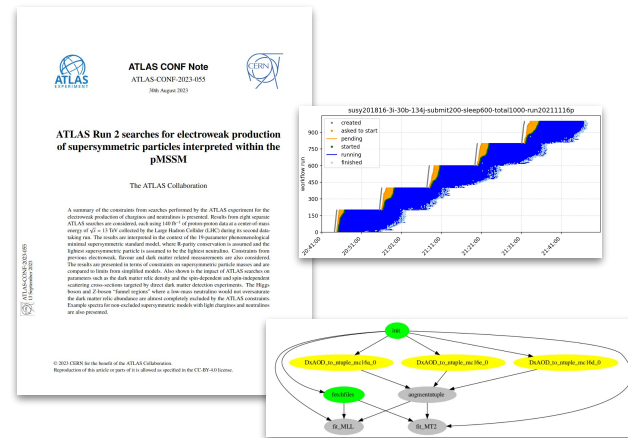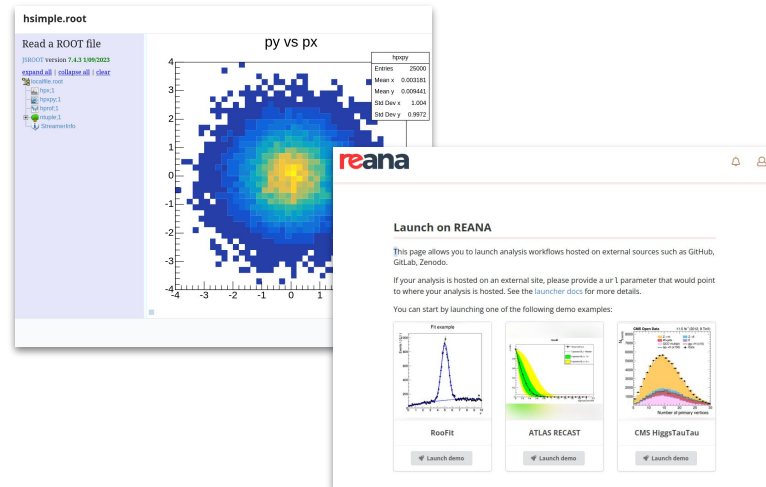
Tomcat / Web Java Application Hosting

- **Replace jobs and cronjobs infrastructure based on VMs**
  - New solution based on Argo Workflow
  - Deployed in Kubernetes
  - Users can track/manage jobs definitions in git repositories
    - Possibility to use their own Docker images
- **Finalized migration to new SSO based on Keycloak**
- **Upgrading Kubernetes clusters to 1.25**
  - Replacing deprecate ingress controller with new one
  - Upgrade all internal dependencies (Prometheus, ArgoCD, ArgoWF) to latest version
  - Manage internal infrastructure via GitOps
  - Deploying few new applications completely via GitOps paradigm
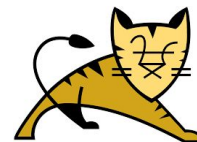
Atlassian products

- **CERN obtained a Community License for Jira and Confluence and the required plugins**

- **Preparation for the consolidation of all instances of JIRA (5) and Confluence (4) at CERN into single, centrally provided Data Centre instances**

Web Hosting Infrastructure

- **Preparing for CC7 end of life in June 2024**
  - CC7 still used for a majority of web sites with static or CGI content
  - Central web hosting services will transition directly to Alma9
  - Notable differences between the CC7 and Alma9 environments:
    - PHP 5 replaced by PHP 8
    - Python 2 not available
- **Started migration campaign for web sites on CC7**
  - 4000 web sites with source files on EOS
    - Provided "previews" of the same content served by Alma9
    - Site owners invited to check previews work properly
    - One-click UI for site owners to switch live sites between CC7/Alma9
  - 2300 web sites with source files on AFS
    - Will not update AFS web servers
    - Move site content to EOS instead
    - Special cases will be followed up with the site owners

InvenioRDM and Zenodo

Zenodo migrated to InvenioRDM platform on October 13th

- Major release focused on making collaboration easy for researcher including support for community members and curation, reviews of upload and new possibilities for sharing records with e.g. peer-reviewers

- https://help.zenodo.org/docs/about/whats-new/

Digital Memory project

A Preserve platform with a web front-end has been implemented and rolled out: http://preserve.web.cern.ch

A Digital preservation strategy for IT and for CERN has been drafted, to help setting up a new Operational Circular ruling digital archival practices

home.cern