CONTENTS

GROUPS REPORTS 2018
Collaborations, Devices & Applications (CDA) Group ................................................................. 6
Communication Systems (CS) Group ......................................................................................... 13
Compute & Monitoring (CM) Group ......................................................................................... 19
Computing Facilities (CF) Group .............................................................................................. 28
Databases (DB) Group .............................................................................................................. 31
Departmental Infrastructure (DI) Group ................................................................................ 36
Storage (ST) Group .................................................................................................................. 37

ACTIVITIES AND PROJECTS REPORTS 2018
CERN openlab ......................................................................................................................... 44
CERN School of Computing (CSC) ......................................................................................... 48
Computer Security .................................................................................................................. 51
Data Preservation ..................................................................................................................... 53
Externally Funded Projects ...................................................................................................... 54
Worldwide LHC Computing Grid (WLCG) ............................................................................. 55
CERN IT Department
Groups and Activities Report 2018

CERN IT department’s mission:

The IT Department provides the information technology required for the fulfilment of CERN’s mission in an efficient and effective manner.

This includes data processing and storage, networks and support for the LHC and non-LHC experimental programme, as well as services for the accelerator complex and for the whole laboratory and its users.

We also provide a ground for advanced research and development of new IT technologies, with partners from other research institutions and industry.

This report aims at summarising the key accomplishments performed by the seven CERN IT department groups in 2018, highlighting their contribution to the department’s mission. The report also highlights the contribution from the following projects and activities: CERN openlab, CERN School of Computing, Computer Security, Data Preservation, Externally Funded Projects, UNOSAT, Knowledge Transfer related activities, and Worldwide LHC Computing Grid (WLCG).
Groups Reports 2018
APPLICATIONS AND DEVICES

The contract for the provision and maintenance of multifunction copier machines and printer support was re-tendered, resulting in 20% savings. The new contract started on December 1. The contract change for Microsoft software license, including the Windows operating system and Microsoft Office, was in preparation throughout the year. At the same time, numerous applications were evaluated for common use-cases and resulted in the following new solutions:

- Draw.IO - for editing diagrams, integrated with CERNBox.
- Blue Griffon - for editing HTML web sites, available for Windows, macOS and Linux.
- Gantt-Chart plug-in - for project management, integrated with JIRA.
- DHTMLX Gantt-Chart plug-in - for sharing and viewing project files, integration in progress.

Migration of user home folders from DFS to CERNBox began, as part of the new strategy for devices, which will also bring increased focus on self-sustained devices (BYOD or CERN-owned). In this context, a study of which operating systems are used by CERN users for professional reasons was carried out, along with market research and tests of several MDM/MAM solutions.

Integrated Collaboration

A new version of the Vidyo client based on the webRTC technology, called VidyoConnect, was deployed and provides a simplified usage and new features like white boarding. The conference...
rooms service now manages 175 rooms and provides new wireless display options avoiding the use of cables and available on all platforms in selected rooms. A new web-based video editing tool was developed to prepare selected Academic Training recordings for publishing in YouTube.

Indico, the CERN collaboration hub, has seen two new releases, implementing role management interface, bulk import of data, and a redesign of the room booking module with a modern user interface and better workflows. The PSI Indico server is now running in the CERN data center, and new German, Portuguese, Japanese, Polish and Italian translation groups have been created. Thousands of CERN visitors have received their CERN access badges through Indico.

Three major projects have started in order to build more open and efficient collaboration services. Open source products have been evaluated to handle authentication, single sign on, resource directory service and federated identity, and the development of a new central authorization framework has started. A market assessment was carried out to select an open infrastructure for managing CERN emails, and a proof-of-concept was deployed and tested by real users. Finally, a new telephony service based on open source components and light local programming was designed and is being developed and tested in collaboration with the CS group.

**DIGITAL REPOSITORIES**

Invenio v3 has been officially released in June and introduced to the world with great success in the annual Open Repositories conference. Following this, many partners expressed desire to use Invenio for their future projects. We have on-going collaborations with Northwestern University, HZDR and Universität Hamburg; and we signed MoUs with the Japanese National Institute of Informatics (which will use Invenio to provide services to more than 500 Japanese universities) in order to enhance Invenio with NGR capabilities and with Data Futures to provide solutions for humanities research datasets.

CDS consolidated the Videos Platform, capturing 2 times more videos than in previous years, and in parallel, the team continued to migrate the current software stack to Invenio v3; substantial work has been invested on a new Library System, designed in collaboration with RERO, a Swiss foundation that will provide an Invenio-based library system to libraries across the country as well as the future
SONAR (Swiss Open Access Repository). Zenodo popularity keeps growing and it already accounts for 2.5 times more visitors than a year ago; during 2018 Zenodo has been awarded an Arcadia grant with partners on Biodiversity and has launched a new Usage Statistics feature according to industry standards, such as the COUNTER Code of Practice as well as the Code of Practice for Research Data Usage Metrics, that allow users to compare metrics from Zenodo with other compliant repositories.

![Figure 1 - Zenodo: Usage Statistics](image)

After the OPERA collaboration released more neutrino data on CERN Open Data, other HEP communities have shown much interest. At the same time, LHC collaborations released recommendations to preserve n-tuples and user code, which came hand in hand with the production beta release of CERN Analysis Preservation. REANA functionality has been expanded, deployed on CERN infrastructure and several successful pilot analysis examples created. A highlight for REANA is the declaration by the ATLAS Exotics Group that partial computational workflow preservation is mandatory for analysis approval.

On the Digital Memory front, the digitisation project of CERN’s photos, videos and audio material has already generated around 350TB of data; and many initiatives and collaboration are popping up for the “mise en valeur” of those historical documents. An example is the discovery of a desk with damaged 35mm slides that was esthetically so appealing that it was converted into an art collection. After an internal exhibition, the CERN VoMeur collection was on the wall of the Images de Marque Gallery in the old town of Geneva. It generated great interest across many media, including the RTS TV News and Le Monde newspaper.
WEB FRAMEWORKS

The web services portfolio was further consolidated in view of providing a consistent set of well-integrated services, facilitating the publishing of information, development of software and deploying services on the Web.

Continuing the evolution of the Platform as a Service for Web Applications based on OpenShift, several core applications were moved to it, in order to profit from the benefits of software container technology. The deep integration of OpenShift with CERN’s Software Development tools allows the deployment and maintenance of applications and documents stored and versioned in Gitlab, and using Continuous Integration and Continuous Deployment techniques to automate a wide range of operational tasks. A major move forward was the deployment of CERN’s Gitlab version control system to OpenShift applying DevOps principles and fully automating the software update process.

Early 2018, following a previous comprehensive evaluation and selection process, the Shared Authoring platform Overleaf was made available to the CERN user community. Within less than a year more than 2000 users have started to use the platform which provides advanced concurrent authoring features for scientific documents and has proven to be highly appreciated and productivity enhancing.

A new major version 8 of CERN’s main Content Management System Drupal has been deployed. In close collaboration with IR/ECO the Drupal 8 service was officially launched with the Release of new CERN Homepage on 5 November 2018.

Posters:

- H. Short/CERN et al, WISE Information Security for Collaborating e-Infrastructures, Computing in High Energy and Nuclear Physics, Sofia, Bulgaria, 09 July 2018
- P. Martin-Zamora, M. Kwiatek, V. Bippus, E. Cruz Elejalde, Increasing Windows security by hardening PC Configurations, Computing in High Energy and Nuclear Physics, Sofia, Bulgaria, 9-13 July 2018

Presentations:

- H. Short/CERN, FIM4R Presenting the 2nd Whitepaper, Nordugrid Conference (remote participation), Munich, Germany, 12 June 2018
- H. Short/CERN et al, FIM4R Presenting the 2nd Whitepaper, TNC18, Trondheim, Norway, 12 June 2018
- H. Short/CERN, WLCG Authorisation Pilot, AARC Plugfest, Milan, Italy, 25 September 2018


• A. Lossent, CERN IT/CDA/WF, “Container orchestration as key to the evolution of CERN Web Services”, HEPiX Autumn/Fall 2018 Workshop, Port d’Informació Científica (PIC), Barcelona Spain, 8-12 October 2018, [https://indico.cern.ch/event/730908/contributions/3153383/](https://indico.cern.ch/event/730908/contributions/3153383/)


• P. Ferreira/CERN, Indico - Event management in HEP, HEPiX Autumn/Fall 2018 Workshop, Case de Convalescència UAB, Barcelona, Spain, 9 October 2018, [https://indico.cern.ch/event/730908/contributions/3147873/](https://indico.cern.ch/event/730908/contributions/3147873/)

• H. Short/CERN, FIM4R Lightning Talk, DI4R, Lisbon, Portugal, 10 October 2018

• H. Short/CERN, Trust and Identity Plenary - What Research Communities want, TechEx28, Orlando, Florida, USA, 15 October 2018

• T. Šimko/CERN, "Reusable data, reproducible analyses: today, tomorrow, next decade", 14th eScience IEEE International Conference, Amsterdam, the Netherlands, 29 October 2018, [https://cds.cern.ch/record/265234](https://cds.cern.ch/record/265234)

• E. Cruz Elejalde, PC Hardening Project at CERN, 2nd Workshop on Cyber Security for HEP, IHEP, Beijing, China, 6-7 November 2018

• E. Cruz Elejalde, Securing a Windows Active Directory (AD) environment, 2nd Workshop on Cyber Security for HEP, IHEP, Beijing, China, 6-7 November 2018

• E. Cruz Elejalde, Securing the Windows Client, 2nd Workshop on Cyber Security for HEP, IHEP, Beijing, China, 6-7 November 2018

• E. Cruz Elejalde, Token-based authentication and authorization, 2nd Workshop on Cyber Security for HEP, IHEP, Beijing, China, 6-7 November 2018


**Publications:**

• P. Martin-Zamora, M. Kwiatek, V. Bippus, E. Cruz Elejalde, Increasing Windows security by hardening PC Configurations, Computing in High Energy and Nuclear Physics, Sofia, Bulgaria, 9-13 July 2018

Reports:

The Communication Systems (CS) Group is responsible for networking and telecommunications services for the Laboratory. For networking, we provide a campus network, including full Wi-Fi coverage, for general purpose connectivity, a technical network to support accelerator operations and critical laboratory infrastructure and, not least, a high-performance data centre network—including high-bandwidth connections to computing facilities around the world—to support physics computing. On the telephony side, in addition to fixed and mobile telephony services we support a TETRA digital radio services for the Fire and Rescue Services. The Group’s service offering expanded in 2018 to include support for LoRa.

The undoubted highlight for 2018 was the completion—on time and under budget—of the three-year project to improve Wi-Fi coverage at CERN. Full coverage is provided for the 200 office buildings at CERN and, in an extension to the project, outdoor coverage is well on the way to being provided for key areas, notably the routes taken as people walk around the sites. CERN staff and users now have a Wi-Fi service that is able to properly support their activities—a service worthy of CERN’s role as the world’s leading particle physics laboratory.
Growth in Wi-Fi client devices and Wi-Fi traffic since January 2018

Providing seamless Wi-Fi coverage, though, was only one aspect of the plan to modernise the Laboratory’s communication infrastructure to meet demands for mobility, flexibility and accessibility. Unfortunately, progress on the other two aspects of this vision—modernising the fixed telephony infrastructure and providing an effective mobile telephony service, has not been so successful during 2018.

On the fixed phone side, migration to a softphone environment, foreseen for the end of Run 2, was postponed until the end of LS2 given uncertainties about the licensing conditions for the favoured softphone client. Nevertheless, 2018 did see important progress on the “back end” developments to modernise our PABX infrastructure with the replacement of our ISDN connections to our upstream telephony operators by IP links.

Turning to mobile telephony, it became ever more clear in 2018 that the current arrangements—with coverage provided by an operator from each host state—cannot provide the level of service that CERN requires. Unfortunately, there are no obvious solutions given the regulatory and commercial arrangements in the host states. Starting in 2019, therefore, a new tripartite group will bring together CERN and host state representatives to review the situation and possible options to deliver an acceptable service. We also hope that the situation will improve in 2019, especially for staff and users based in France, as we migrate to Swisscom’s new 5G service offering.

As elsewhere at CERN, preparing for LS2 was a priority for the group during 2018. This long shutdown is a rare opportunity to upgrade the infrastructure of the Technical Network which provides connectivity to equipment essential for the operation of the accelerators. Following extensive evaluations, we selected a network router capable of providing the fine-grained access control necessary to protect this critical equipment. Deploying these routers will be a key challenge for 2019. The group also needs to replace large sections of the radiating cable that supports mobile telephony and digital radio services underground and 2018 also saw extensive preparations for this work, notably the testing of cable and components to assess their resistance to the higher levels of radiation foreseen for the HL-LHC.
Looking further ahead, the infrastructure supporting the campus network will need to be refreshed after LS2 and so the Group consulted key colleagues across CERN to assess the future needs for wired connectivity across the campus. We expect much traffic from users to migrate to the Wi-Fi service in future years but wired connectivity is still needed with the growing demand for connected infrastructure and smart buildings. Fibre-to-the-Office (FTTO) is an increasingly mature technology that offers the potential to reduce the amount of expensive equipment needed to support a campus network and, after an initial study, the Group purchased a small FTTO configuration for more extensive testing during 2019.

The demanding future needs of the LHC experiments were not forgotten either in 2018. In collaboration with LHCb, the Group successfully demonstrated operation of a remote High Level Trigger farm. Servers in the computer centre in Meyrin were operated in parallel with those at the LHCb pit for an extended period in the middle of the year. The remote nodes behaved correctly and performed exactly as the production nodes with the LHCb online team commenting that, “operationally and performance-wise, it seems to be the same as running a sub-farm at Point 8”.

Looking wider afield, all of the data produced by the LHC experiments at CERN naturally leads to a high demand for data export—and also for high volume data transfers between WLCG sites. For redundancy reasons there are usually multiple network links between sites but, usually, traffic flows over only one of these links at a time, limiting the available bandwidth. In 2018, together with our colleagues in the Storage team, the Group demonstrated the feasibility of a proposal to dynamically reconfigure the network to allow the use of parallel paths—Project NOTED, for Network Optimised Transfer of Experimental Data. During tests in October, an additional 20Gbps of bandwidth was added between CERN and Amsterdam and exploited transparently by FTS during transfers of ATLAS data to the Tier1 in Nikhef. As elsewhere, we expect this initial work to lead to more in-depth developments in 2019.
Even as we look to the future we don't overlook the needs of our users today, though! In 2018, the Group installed over 250 network switches and routers and created 14 new network star points. Additionally, nearly 2,300 new network outlets as we responded to over 500 user deployment requests. Returning to a point above, the promise of FTTO is not only that it could reduce the infrastructure costs of the campus network, it also has the promise to increase flexibility for our end users. Responsive as we are, our users would appreciate more being able to deploy new network outlets themselves! Whatever happens, the Group’s efforts during 2019 have laid the groundwork for much interesting work in 2019.

Presentations:

• S. Agosta/CERN, Radio Communications in a Hostile Environment, IEEE International Conference on Communications, Kansas City, USA, 21 May 2018


• E. Martelli/CERN, LHCb High Level Trigger in a remote IT datacenter, CHEP 2018, Sofia, Bulgaria, 12 July 2018, https://indico.cern.ch/event/587955/contributions/2937892/ (proceedings article in review)

• A. Shevrikuko/CERN, Network configuration management at CERN: status and outlook, HEPiX Fall 2018, PIC, Barcelona, Spain, 8 October 2018, https://indico.cern.ch/event/730908/contributions/3153336/

• C. Kishimoto/CERN, CERN campus network upgrade, HEPiX Fall 2018, PIC, Barcelona, Spain, 8 October 2018, https://indico.cern.ch/event/730908/contributions/3148844/


• T. Cass/CERN, WLCG and Related Networking Activities, LHCOPN-LHCOne Meeting, Fermilab, Batavia IL, USA, 31 October 2018, https://indico.cern.ch/event/725706/contributions/3118915/attachments/1744145/2823033/DataTransfer.pptx


**COMPUTE & MONITORING (CM) GROUP**

The Compute and Monitoring group is responsible for the service delivery and evolution of Compute, Monitoring and Infrastructure services for the CERN Tier-0 and WLCG.

In order to deliver these services, we work closely with other grid sites and open source communities to jointly develop and enhance the tools for the end users and service managers.

**LINUX**

With over 50,000 servers running one of CERN’s Linux distributions, careful planning of new releases and retirement of older ones is needed. Scientific Linux 6 remained the default version for physics during 2018 but CERN CentOS 7 usage has grown significantly throughout the year to approach the usage of the previous version. The plan to switch the default to CentOS 7 early in 2019 was agreed with the experiments.

The infrastructure for Linux software building and software delivery was redesigned to use the latest server provisioning tools such as OpenStack and CephFS.

Two major security incidents during the year (Spectre/Meltdown and L1TF) were handled in collaboration with the user community. Addressing these incidents promptly is required in order to ensure the CERN Linux systems remain protected.

**PUPPET CONFIGURATION MANAGEMENT**

The Puppet configuration management system allows for consistent management of servers, backed by quality assurance of changes through automatic testing. With over 41,000 servers now managed by Puppet, there was a need to tune the PuppetDB data warehouse to handle this expansion.

Further automation of operations for IT services is now enabled using notifications for events such as a computer centre power cut.
PUBLIC LOGIN SERVICES
The lxplus service provides a large scale, efficient, general purpose linux service for CERN users. The service has been enhanced to include further security checks using the Fail2ban tool to block repeated login failures. EOS storage access using the FUSE mounts have been added to allow transparent access to data stored in EOS.

CLOUD SERVICES
Over 90% of the compute resources in the data centre are provided through a private cloud based on OpenStack. With the growth of the computing needs of the CERN experiments and services, this has now reached more than 320,000 compute cores running in Meyrin and Budapest. New services such as Bare Metal management, workflows for regular activities, expiration of personal virtual machines, file shares, S3 object stores and support for GPUs have been added. During the year, there were further enhancements to the cloud container service such as integration with CERN storage solutions, with over 400 Kubernetes clusters now in use by the user communities.

Collaborations with Huawei through openlab and the Square Kilometer Array produced Cells V2 to enable clouds to scale further and this was deployed to production. The pre-emptible virtual machine solutions were agreed with the OpenStack community and will be deployed in 2019.
**Figure 2 - Cloud service resources and usage**

**BATCH**

CERN has used the Platform LSF batch system for over 20 years to deliver the compute resources to the experiments according to the pledges made. However, with the growth of compute capacity, it has reached the limits of what it can support. HTCondor, an open source batch system from the University of Wisconsin, has become widely adopted across high energy physics and has shown significant scalability improvements compared to LSF. The migration of the majority of the user communities to HTCondor was completed during 2018 and the public service will be stopped early in 2019. External clouds were also used to demonstrate the feasibility of running scientific workloads along with providing further capacity such as via the HNSciCloud or public providers such as Oracle through openlab.

Further resources have been made available using the processing power available in the CERN storage servers. Appropriate workloads can be routed to these servers and run concurrently with the data services that these machines provide. The workloads can be abruptly terminated in the event of a high volume of requests.
HIGH PERFORMANCE COMPUTING
While most of the computing done at CERN can be separated into individual batch jobs and then farmed across a large number of servers, some applications require multiple machines to be combined to deliver a high-performance single application. The Health and Safety, Theory and Beams departments are now using this facility and further consolation of engineering applications, for very large memory jobs has taken place during 2018.

VOLUNTEER COMPUTING
Volunteer computing for CERN experiments continued to provide significant capacity during 2018 and has continued to grow providing peaks of 400,000 jobs.
ELASTICSEARCH
Elasticsearch is a search and analytics store used widely to understand complex data stores. Using Solid State Drives, significant performance improvements were delivered during 2018. There are now nearly 50 clusters for use cases such as security, repository search and log analysis.

MONITORING
The unified monitoring infrastructure (MONIT) collects monitoring metrics and logs from the CERN Data Centres, IT Services and the WLCG grid. From the initial 200 GB/day of monitoring data, the volume is now of 3 TB/day with over 1,500 users. The new infrastructure monitoring is in place using the open source collected solution to replace the CERN Lemon tool used since 2003. Alarms can now be raised so Service Now tickets are automatically created when problems occur. Many WLCG dashboards are in production using Grafana rather than the dedicated instances developed in the past.

WLCG network and site monitoring continue to evolve with recent work into Software Defined Networking introducing some potential for future optimizations.
Figure 6 - Perfsonar world wide network monitoring

Posters:


Presentations:

- Jarka Schovancova, "Performance metrics and measurements in the Data Lake model", WLHC and HSF Workshop, Naples, Naples, Italy, 26 March 2018, Slides
- Ulrich Schwickerath, "ES at CERN update", IT Monitoring at CERN and FNAL, CERN, Geneva, Switzerland, 24 April 2018, Slides
- Marian Babik, "Introduction to SDN", HEPIX SDN working group, CERN, Geneva, Switzerland, 25 April 2018, Slides
- Ricardo Rocha, "CERN Experiences with Multicloud Federated Kubernetes", Kubecon Europe, Copenhagen, Denmark, 2 May 2018, Slides
- Asier Aguado Corman, "Monitoring Infrastructure for the CERN Data Centre", HEPIX Spring 2018 Workshop, Madison, USA, 14-18 May 2018, Slides
- Spyros Trigazis, "Baremetal provisioning in the CERN cloud", HEPIX Spring 2018 Workshop, Madison, USA, 14-18 May 2018, Slides
- Spyros Trigazis, "Status update of the CERN private cloud", HEPIX Spring 2018 Workshop, Madison, USA, 14-18 May 2018, Slides
- Helge Meinhard, "HTCondor at CERN", HTCondor Week, Madison, USA, 21-23 May 2018, Slides
- Ricardo Rocha, "Magnum Project Update", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Belmiro Moreira, "Moving from CellsV1 to CellsV2 at CERN", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Arne Wiebalck, "Defending the cloud", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Belmiro Moreira, "Containers on Baremetal and preemptible VMs at CERN and SKA", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Arne Wiebalck, "To Complete, not to Compete: Integrating Ironic into CERN's Private Cloud Service", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Arne Wiebalck, "(R)Evolution in CERN IT: Operational War Stories from 5 Years of Running OpenStack in Production", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Ricardo Rocha, "CERN experiences with Multi Cloud, Federated Kubernetes", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Dan van der Ster, Arne Wiebalck, "Ceph and the CERN HPC Infrastructure", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Slides
- Ricardo Rocha, "Containers on OpenStack keynote", OpenStack Summit Vancouver, Vancouver, Canada, 21-24 May 2018, Video 16m in
- Tim Bell, "A 5 year perspective on the CERN cloud", OpenStack Days Budapest 2018, Budapest, Hungary, 6 June 2018, Slides
- Paulo Canhão, "HNSciCloud overview", GridKa school of computing, Karlsruhe, Germany, 27-31 August 2018, Slides.
- Jan Van Eldik, "5 Years of OpenStack at CERN" (20 minutes), OpenStack days Italy, Roma, Italy, 21 September 2018, slides.
- Nacho Barrientos, "GitLab use at CERN: computer center configuration management", GitLab@CERN Day, CERN, Geneva, Switzerland, 3 October 2018, Notes.
- Ben Jones, "Too much BEER and the Kubernetes dance", HEPix Fall 2018 Workshop, Barcelona, Spain, 8-13 October 2018, Slides.
- Marian Babik, "Network Functions Virtualisation Working Group", HEPix Fall 2018 Workshop, Barcelona, Spain, 8-13 October 2018, Slides.
- Marian Babik, "WLCG Networking Update", HEPix Fall 2018 Workshop, Barcelona, Spain, 8-13 October 2018, Slides.
- Thomas Oulevey, "Update on Linux at CERN", HEPix Fall 2018 Workshop, Barcelona, Spain, 8-13 October 2018, Slides.
- Zhechka Toteva, "Workflow automation with Megabus", HEPix Fall 2018 Workshop, Barcelona, Spain, 8-13 October 2018, Slides.
- David Moreno, "The Hows and Whats of Service Monitoring at CERN", Puppetize Live Amsterdam, Amsterdam, Netherlands, 10 October 2018, Slides.
- Zhechka Toteva, "How do we control more than 40,000 machines in the CERN Computer Centre", Bulgarian Engineering Teacher Programme, CERN, Geneva, Switzerland, 16 October 2018, Slides.
• Spyros Trigazis, "Fedora Atomic Host at CERN", CentOS Dojo at CERN, CERN, Geneva, Switzerland, 19 October 2018, Slides
• Pablo Saiz, "Go implementation of DNS Load Balancing at CERN", GoLab - The International Conference on Go in Florence, Florence, Italy, 23 October 2018, Slides
• Marian Babik, "Introduction to HEPIX NFV Working Group and SDN/NFV for compute", LHCOPN-LHCONE meeting, Chicago, USA, 30-31 October 2018, Slides
• Marian Babik, "perfSONAR update", LHCOPN-LHCONE meeting, Chicago, USA, 30-31 October 2018, Slides
• Nacho Barrientos, "Puppet and Foreman", University of Bonn Visit, CERN, Geneva, Switzerland, 7 November 2018
• Pablo Llopis, "CephFS and HPC at CERN", Supercomputing 2018, Dallas, TX, USA, 12-15 November 2018, Slides
• Belmiro Moreira, "Pre-emptible instances and bare metal containers", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Belmiro Moreira, "Scaling Nova with Cells V2", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Daniel Abad, "Our relationship is complicated, Nova and Ironic", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Robert Vasek, "Cephfs shares on Kubernetes", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Spyros Trigazis, "Providing turnkey container clusters on OpenStack", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Jose Castro Leon, "Towards a fully automated cloud", OpenStack summit Berlin, Berlin, Germany, 13-15 November 2018, Slides
• Jose Castro Leon, "SDN and the CERN cloud", Tungsten Fabric User Meeting, Berlin, Germany, 14 November 2018, Slides
• Tim Bell, "A 5 year perspective on clouds at CERN", UCC 2018, Zurich, Switzerland, 19 December 2018, Slides

Publications:

The IT-CF group contributes in the following areas: Data Centre (DC) Operations and System Administration, Service Management Support and Hardware Evaluation and Procurement. In terms of DC Operations, the day-to-day work has been to deploy new servers and data storage (some 485 servers and 40PB of storage) at the main DC in Meyrin, as well as the retirement of old equipment.

The vast majority of the retired equipment was prepared for the DG’s program of donating IT equipment to other institutes around the world. 7 pallets of equipment were donated to one institute in Nepal and a further 20 pallets of equipment to LHCb with a further three in preparation (Uganda, Lebanon and Palestine). In addition, the Operations Team continues to look at improving the efficiency of the Data Centre both in terms of energy as well as operational efficiency. The campaign to re-organise the layout of the main computer rooms, as well as upgrading the electrical distribution, has continued, as equipment is retired and replaced, and now stands at around 85% complete and a further cold aisle with new larger racks has been successfully installed.

*Donation to Nepal made by the Director for Computing and Research, Eckhard Elsen.*

The DC Operations contract was successfully re-tendered and the transition was handled smoothly. The new contract has resulted in an optimization of the service. Finally, for the DC Operations, efforts have continued to introduce more and more automation of manpower intensive workflows, one example of this being the automation of many of the steps for server hardware retirement.

For the Service Management team this year, as well as providing the day-to-day Service Management support the team, a major effort has been made in support of introducing Data Privacy Notices for all
services. This included producing configurable template as well as instructions and support for their completion.

A selection of these are listed below:

- Continual consolidation of new 2nd Line Computing Support, including monthly reports on user feedback.
- Sustained collaboration with SMB to improve the Service Desk delivery for IT tickets.
- Follow-up and reporting on negative feedback on all IT tickets.
- Regular newsletter to all Snow users.
- Continual training of IT 2nd line.
- Automatic detection and follow-up of Major IT Incidents.
- Privacy notices for our Services.
- Analysis and implementation of Record Producers.
- OLAs for various IT contracts and groups.
- Central IT CMDB project.
- Part of Push Notifications project.
- Hosted event “ServiceNow user group Suisse Romande”.
- Moderation of IT SSB entries.
- Management of S220 Operations contract.

In terms of procurement, the team has conducted one major tender amounting to 4.2 MCHF (purchases for BE/CO and the LHC experiments included). These purchases included not only equipment for CERN’s main Data Centre at Meyrin, but also a significant amount for BE and the LHC experiments. As the number of configuration required was much higher than in previous recent tenders, a significant amount of time and effort was required for testing of the deliveries. In addition, the hardware monitoring sensors are gradually being ported to a new infrastructure based on collected and this should be completed by the end of this year. Also in the area of the support infrastructure, good progress has been made on the integration of our existing workflows (registration, burn-in, commissioning, repurposing and decommissioning) in Openstack Ironic to benefit from existing IT services as well as to improve the efficiency of these processes.

A final activity in the group is the running of hardware labs, CERN openlab and Techlab, which evaluate new hardware technologies for suitability for later use in the CERN DCs or for the experiments. As well as performing hardware evaluations and benchmarking, the team also evaluates code optimisation on these hardware platforms. A new company has joined openlab during the year, E4, to work with IT-CF and to contribute in the domain of GPU computing, which will provide valuable input for future tenders requiring GPU accelerated servers. For Techlab, the team
specified, procured and commissioned a Quad chassis with four dual socket Marwell ThunderX2, which are the most powerful 64bit ARM processors available on the market. Together with IT-CM group the team also successfully tested the provisioning GPGPU resources via virtual machines in Openstack.

Posters:


Presentations:

- D. Clavo/CERN, “Challenges, solutions and lessons learnt in 7 years of Service Management at CERN”, HEPiX Autumn/Fall 2018 Workshop, Barcelona, Spain, 8-12 October 2018, https://indico.cern.ch/event/730908/contributions/3147838/
Database On Demand (DBoD)

The DBoD service has an incredible acceptance by CERN community, supporting many applications and critical services. During this year, we have reached more than 700 instances deployed in all departments and added features and enhancements based on the CERN community continuous feedback and needs. The DBoD backup infrastructure has been improved using EOS as a second backup solution for extended binary log archiving and full database backups (as a possible replacement of TSM in the future). The monitoring of these backups was also improved to guarantee that a minimum number of backups is always available. In addition, the new DBoD API has been finished, being the interface between the data and other components like the DBoD web interface, Rundeck, and the DBoD core framework. The new web interface based on Angular is being tested and will soon be available to the users. InfluxDB, officially launched in 2017, was upgraded in several occasions and has reached today a stable release for CERN use cases.

Oracle Databases

The focus during 2018 has been the preparation for the major upgrades and changes to come during LS2. Still some individual components, such as the Oracle Application Express installations have been upgraded. The Database monitoring infrastructure has moved forward to the integration with the IT solutions and the internal, home-made database monitoring system called RACMon has been decommissioned. In addition, Database Service access control has also been enforced implementing Oracle Connection Manager for administrative and some experiment databases (that requires external access).

The CERN Accelerator Logging Oracle database has reached 1PB during the year 2018
Hadoop and Spark

During this year, we renewed ~25% of the hardware infrastructure and worked into the standardization and consolidation of the different Hadoop clusters into the Apache native distribution. This has allowed us to have newer versions of the different components in use and easier to plug our custom integrations (e.g: E-group support for authorization and resource ownership). At the same time, we also consolidated the client access based on RPM distribution of both the software needed to access our services and the configuration for each cluster. We also performed the roll-out of the production server infrastructure and services for NXCals project (next generation accelerator logging service, a critical service for LHC operations). In terms of the Hadoop ecosystem, we further improved our hadoop xrootd connector, that allows to seamlessly access xrootd filesystems with industry standard big data frameworks (e.g.: Spark and Hadoop).

For easing the access to the data analytics tools (e.g. Spark) provided by IT-DB-SAS, we have carried out the necessary integrations of the notebooks service (SWAN) with Hadoop and Spark Service which allows users to access and offload computations to Hadoop and Spark cluster from notebooks. During the course of the year we have enhanced the integration by automating deployment of NXCals software to CVMFS, adding password-less access to Hadoop and Spark clusters, browsing of HDFS filesystem from notebooks and prototype support for Kubernetes clusters (running on OpenStack Magnum) for spark workloads.

In the context of the Openlab collaboration we have worked together with the users community to develop use cases for HEP physics data processing and machine learning.

Streaming

This new service based on Apache Kafka is now in production mode, our offering consists on managed dedicated Kafka clusters (intensively used for instance by the CERN IT Security Team), and also a newly built beta service for shared Kafka clusters, where multiple users can use the same infrastructure with a self-service portal to manage topics and Access Control Lists (including integration with E-groups and support for both Kerberos and certificates). This type of cluster covers not only small use cases but also as basis for other shared platforms that need a streaming component. During this year we partnered with IT-CS into building an IoT data plane (LoRa) prototype where this architecture is used.

InfluxDB ingest per second on the CERN Database on Demand platform, over 70 000 data points in total are acquired per second, essentially for monitoring services.
Posters:

- Luis Rodriguez Fernandez, Antonio Nappi Weblogic on Kubernetes, CERN Openlab Technical Workshop, Geneva, Switzerland, 11 January 2018
- Oracle Weblogic on Kubernetes, Borja Aparicio Cotarelo, CHEP 2018, Sofia, Bulgaria, 9-13 July 2018

Presentations:

- Zbigniew Baranowski/CERN, Big data at CERN, Seminar at Padova University, Padova, Italy, 19 January 2018
- Rebekka Alsvåg/CERN, Oracle JET Experiences at CERN, Oracle User Group Norway, Oslo, Norway, 8 March 2018
- Vaggelis Motesnitsalis/CERN, Hadoop and Spark Services at CERN, DataWorks Summit Berlin, Berlin, Germany, 18-19 April 2018
- Vaggelis Motesnitsalis/CERN, From Collision to Discovery: Physics Analysis with Apache Spark, CERN Spring Campus, Riga, Latvia, 23-26 April 2018
- Vaggelis Motesnitsalis/CERN, Big Data Technologies at CERN, CERN Spring Campus, Riga, Latvia, 23-26 April 2018
- Zbigniew Baranowski/CERN, Evolution of the Hadoop and Spark platform for HEP, HEPIX Spring, Wisconsin-Madison, USA, 14-18 May 2018
- Vaggelis Motesnitsalis/CERN, Physics Analysis with Apache Spark in the CERN Hadoop Service and DEEP-EST Environment, IT Technical Forum, CERN, Geneva, Switzerland, 25 May 2018
- Prasanth Kothuri/CERN, Open Source Big Data Tools Accelerating Physics Research at CERN, OpenExpo Europe Conference, Madrid, Spain, 06-07 June 2018
- Rebekka Alsvåg/CERN, Oracle Technologies at CERN - Experiences from a Summer Student Project, Oracle Code, Berlin, Germany, 12 June 2018
- Manuel Martin Marquez/CERN, Boosting Complex IoT Analysis with Oracle Autonomous Data Warehouse Cloud, Oracle Global Leaders Meeting – EMEA, Budapest, Hungary, 19-20 June 2018
- Prasanth Kothuri/CERN, Hadoop and Spark at CERN, JP Morgan Innovation Week, Geneva, Switzerland, 25 June 2018
- Manuel Martin Marquez/CERN, Franco Amalfi (Oracle), Accelerating Particles, the Role of Cloud and Analytics, Keynote, The 11th PErvasive Technologies Related to Assistive Environments (PETRA) Conference, 25-29 June 2018
- Prasanth Kothuri/CERN, Apache Spark usage and deployment models for scientific computing, CHEP2018 conference, Sofia, Bulgaria, 9-13 July 2018
- Zbigniew Baranowski/CERN, Evolution of the Hadoop platform for HEP, CHEP2018 conference, Sofia, Bulgaria, 9-13 July 2018
- Vaggelis Motesnitsalis/CERN, CERN IT Lectures (CERN openlab Summer Student Programme), From Collision to Discovery: Physics Analysis with Apache Spark, CERN, Geneva, Switzerland, 7 August 2018
- Franck Pachot/CERN, Join Methods: Nested Loop, Hash, Sort, Merge, Adaptive, POU, Poland, 7 September 2018
- Ludovico Caldara/CERN, Effective Oracle Home Management with the New Release Model, POU 2019, Sopot, Poland, 07 September 2018
• Luca Canali/CERN, Apache Spark for RDBMS Practitioners, Spark Summit Europe 2018, London, UK, 4 October 2018
• Prasanth Kothuri/CERN, Piotr Mrowczynski/CERN, Experience of Running Spark on Kubernetes on OpenStack for High Energy Physics Workloads, Spark Summit EU conference, London, United Kingdom, 04-05 October 2018
• Ludovico Caldara/CERN, LHC Long Shutdown 2 and database changes, Hepix Autumn 2018, Barcelona, Spain, 11 October 2018
• Franck Pachot/CERN, Microservices: Get Rid of Your DBA and Send the DB into Burnout, Oracle Code One - San Francisco, CA, USA, 22 October 2018
• Franck Pachot/CERN, Join Methods: NL, HASH, MERGE, & Adaptive in Slow-Motion, OakTable World 2018 - San Francisco, CA, USA, 23 October 2018
• Franck Pachot/CERN, Multitenant Security Features Clarify DBA Role in DevOps Cloud, Oracle Open World 2018 - San Francisco, CA, USA, 23 October 2018
• Monica Riccelli/Oracle, David Cabelus/Oracle, Antonio Nappi/CERN, Running a Modern Java EE Server in Containers Inside Kubernetes, Oracle Openworld 2018, San Francisco, USA, 23 October 2018
• Luis Rodriguez Fernandez/CERN, A Thousand Things You Always Wanted to Know About SSO but Never Dared Ask, Oracle CodeOne – San Francisco, CA, USA, 24 October 2018
• Franck Pachot/CERN, Join Methods: Nested Loop, Hash, Sort, Merge, Adaptive, DOAG Conference, Germany, 20 November 2018
• Ludovico Caldara/CERN, Effective Oracle Home Management with the New Release Model, DOAG 2019, Nuremberg, Germany, 21 November 2018
• Franck Pachot/CERN, CBO Panel, OUG Tech18, Liverpool, UK, 4 December 2018
• Franck Pachot/CERN, Oracle Database on Docker: Lessons Learned, UKOUG Tech18, Liverpool, UK, 4 December 2018
• Ludovico Caldara/CERN, Effective Oracle Home Management with the New Release Model, UKOUG Tech 2019, Liverpool, GBR, 5 December 2018
• Eric Grancher/CERN, Manuel Martin Marquez /CERN, Sebastien Masson/CERN, Managing one of the largest IoT Systems in the world, Oracle Global Leaders Meeting – EMEA, Sevilla, Spain, 11-12 December 2018
• Vaggelis Motesnitsalis/CERN, Big Data at CERN, Second International PhD School on Open Science Cloud, Perugia, Italy

Publications:

• Rebekka Alvsvåg/CERN, A Dream Come True: Experiences from a Summer Intern at CERN openlab, OraWorld Magazine, October 2018
• Franck Pachot/CERN, Why you must run on the latest release update, UKOUG OracleScene Magazine, December 2018
Reports:

DEPARTMENTAL INFRASTRUCTURE (DI) GROUP

The Department Infrastructure (DI) group provides and maintains the administrative and infrastructure support for the IT Department, i.e. the planning of departmental resources (Budget and Personnel) and the general services such as the financial administration (follow-up of invoices, contracts, requests for funds, inventory, etc.), and the IT Car Pool.

As well as these administrative activities, the group hosts the following activities:

- CERN openlab
- Computer security office
- Externally funded projects
- IT secretariat
- UNOSAT
- Worldwide LHC Computing Grid (WLCG) project

Most of these activities are detailed in the ‘Activities and Projects Reports 2018’ section of this report.
The ST group is responsible for storage services for the CERN physics program (notably EOS, CASTOR, FTS, CVMFS, AFS, DPM) and infrastructure services (CERNBox, Ceph, Backup and FILER).

During 2018, over 115PB of new data have been written to tape, out of which 100PB were persistently stored. In addition, more than 100PB of data have been migrated to newer-generation media. The total volume in the CERN data archive now exceeds 330PB, therefore more than tripling the amount of data since the beginning of Run-2 (100PB in 2015). In November alone, a record-breaking 15.8PB was written to tape – corresponding to a year’s worth of data during LHC Run-1.

The development of the new CERN Tape Archive (CTA) software is nearing completion. The main goal of CTA is to make more efficient use of the tape drives, to handle the data rate explosion anticipated for Run–3 and Run–4. CTA will be deployed during LS2, replacing CASTOR.

The IT tape infrastructure is being refreshed. After thorough testing, LTO (Linear Tape Open) technology has been adopted as a complement to enterprise tape. LTO offers a competitive pricing structure but lacks some enterprise features, in particular drive-assisted seeking, resulting in comparatively poor performance when retrieving multiple files from tape. However, our investigations have shown that it is possible to implement this optimisation outside the drive and it will be provided as part of CTA.
In 2018, ST coordinated the heavy-ion readiness tests (IT and experiments). This was needed due to the increased requirement from the experiments (as high as 8~GB/s in the case of ALICE). In September, IT (mainly ST and CS) prepared and demonstrated the readiness of the IT infrastructure together with the 4 experiments operating at higher rate during an LHC technical stop.

The actual data taking with ions (3 last week or run 2) were smooth thanks to the preparation and instant rates were often exceeding the design values. EOS could ingest all data without interference between the different experiments and CASTOR sustained the rates without significant backlog induced by the higher activity.

No interference between the total EOS activity was observed (LHC + PUBLIC + USER/HOME). Read activity routinely exceeded 50 GB/s during data taking with no interference with the write activity dominated by the characteristic LHC-related duty cycle in the data taking part (exceeding 20 GB/s). In the following plots, the data of November 14 and 15 are displayed.

The EOS infrastructure continued to evolve in 2018. Main changes were the demonstration of the new catalogue (disk based). This has been tested up to the scale required to replace the current AFS (over 4 Billions entry). We introduced the new deployment model which allows scale-out instances: this is in use as replacement of the old (monolithic) EOSUSER. EOSUSER data have been migrated to the infrastructure called EOSHOME (5 instances so far) which removes the scale-out concerns and reduces the potential impact of incidents. 14k users (700Mfiles) have been moved.

This change moved all the data of both CERNBox users and users using EOSUSER for file access (POSIX via FUSE and XROOT). Now we are preparing the project space to be migrated to the new
infrastructure. Together with IT-CDA we are preparing to take on board also user DFS space. The COMTRADE OpenLab activity entered a new phase and we had a 2-week visit (3 engineers) to kick start their new contribution (full-stack-container deployment of EOS including CERNBox and SWAN: ScienceBox).

During 2018, the Ceph infrastructure continued its stable growth from 14 to 18 PB JBOD capacity across several instances. The largest use-case by volume continues to be block storage for our OpenStack cloud, which now hosts more than 6000 volumes attached to virtual machines. During an Openlab project with Rackspace, we injected some development effort in order to better understand and optimize block storage for high-IOPS use-cases; we expect that work to enable new database-type applications in 2019. On the CephFS front, we continued stable operations and performance tuning for the CephFS HPC scratch areas, highlighted by our first entry into the IO-500 list presented at SuperComputing 2018 (which was the first CephFS entry in the list).

We also saw active early CephFS adoption by IT infrastructure applications managed by OpenStack Manila, and the majority of our previous NFS Filer users have migrated to CephFS/Manila, with the remaining planned for 2019 after we improve the snapshotting and backup capabilities of the service. S3 was announced in early 2018 as a production service for infrastructure, archival, and other use-cases which can profit from this Amazon AWS-compatible storage protocol.

The largest users by volume are the ATLAS Event Service and LHC@HOME, but several services now depend on S3 for daily operations, including Gitlab, Indico, and others. One service moving in the S3 direction is CVMFS, which remained stable during 2018 while we prepared for large architectural changes to come in 2019. Those include a new "gateway" feature to allow modifications to a CVMFS repository to happen in parallel (e.g. by CI pipelines), and indeed with a new storage backend based on S3.

SWAN is getting larger user acceptance: almost 1000 userids connected to the system (30% with ATLAS affiliation, ~20% with CMS). The breakdown by department is 60% with EP, with IT and BE usage at the 15% level each.

SWAN (as a part of ScienceBox) and HelixNebula resources were offered to TOTEM. They successfully used the hybrid cloud resources, benefiting the parallel-processing capability of SWAN (Spark adapter): it is an interesting usage of an innovative stack to enable physicists to perform important data analysis while collecting essential operational experience and user feedback.

ST organised the 4th CS3 conference. This time the event was hosted in Krakow (Cyfronet) on January 29-31 2018. The event was successful with more than 120 registered participants, with a significant participation from companies (about 10) and new communities (notably HPC centres as Pawsey - Australia, CSCS Switzerland and other representatives from this community). In total 41 institutes from 27 countries were represented and the feedback was excellent: a new event is being organised (January 2019 - Rome).
The second EOS workshop organised by ST group took place early 2018, with more than 80 participants from the growing international user community. During the year, the EOS development team implemented a new scale-out architecture based on a micro-service approach to minimise the impact of system upgrades and to create a distributed, more scalable namespace. The new system that has been demonstrated to support even very large deployment of more than 4 billion files.

The DPM team in the group has organised a DPM community workshop in Prague and now supports the DPM deployment teams at over 100 sites in their transition to the consolidated and modernised DOME code-base.

The File Transfer Service (FTS) project, which implements the data-exchange back-bone of the worldwide LHC grid, has transferred over 1 billion files with a total volume of some 0.8 exabyte and is now used by over 20 experiments at CERN and in other data intensive sciences. In 2018 the FTS team has been very active in supporting the technology R&D with regard to more user-friendly authentication and delegation methods that took place between the experiments and WLCG sites in the DOMA activity.

As part of the group’s collaboration in community-wide projects, we improved the security and credential delegation of the widely used XROOT framework and further increased its robustness at very large deployment scales. The group further made multiple contributions in data-management area in the XDC project and the WLCG DOMA activity.

Presentations:

- H. Gonzalez Labrador, "Future architectures for sync&share based on microservices" and "CERNBox service", CS3 January 2018, Krakow, Poland
- H. Gonzalez Labrador, "CERNBox: the CERN cloud storage driven by EOS" and "A microservice architecture for CERNBox", EOS workshop, CERN, Geneva, Switzerland, February 2018,
- D. van der Ster, "Ceph for Big Science", Cephalocon, Beijing, China, March 2018
• J. Moscicki, "Storage at CERN" and "CERNBox and SWAN", Hepix May 2018, Madison, USA
• D. van der Ster, "Ceph and the CERN HPC Infrastructure", Openstack summit May 2018, Vancouver, Canada
• H. Gonzalez Labrador, "CERNBox: the CERN cloud storage hub" and "Cloud storage for data-intensive sciences in science and industry", CHEP July 2018, Sofia, Bulgaria
• H. Rousseau, “Storage at CERN”, CHEP July 2018, Sofia, Bulgaria
• Michael Davis et al, CERN Tape Archive – from development to production deployment, CHEP July 2018, Sofia, Bulgaria
• CHEP 2018, 9-13 July 2018, Sofia, Bulgaria:
  o Plenary - EOS Open Storage - evolution of an ecosystem for scientific data repositories, Andreas Joachim Peters (CERN), [https://indico.cern.ch/event/587955/contributions/3012718/](https://indico.cern.ch/event/587955/contributions/3012718/)
  o Ready to Go Data Transfers: supporting the long tail of science, [https://indico.cern.ch/event/587955/contributions/2936900/](https://indico.cern.ch/event/587955/contributions/2936900/)
  o Building a global file system for data access using Large Scale CVMFS and DynaFed, [https://indico.cern.ch/event/587955/contributions/2936825/](https://indico.cern.ch/event/587955/contributions/2936825/)
  o An http data-federation eco-system with caching functionality using DPM and Dynafed, [https://indico.cern.ch/event/587955/contributions/2936833/](https://indico.cern.ch/event/587955/contributions/2936833/)
  o Advancements in data management services for distributed e-infrastructures: the eXtreme-DataCloud project, [https://indico.cern.ch/event/587955/contributions/2936860/](https://indico.cern.ch/event/587955/contributions/2936860/)
  o Using a dynamic data federation for running Belle-II simulation applications in a distributed cloud environment, [https://indico.cern.ch/event/587955/contributions/2936834/](https://indico.cern.ch/event/587955/contributions/2936834/)
  o Testing of complex, large-scale distributed storage systems: a CERN disk storage case study, Andrea Manzi (CERN), [https://indico.cern.ch/event/587955/contributions/2938243/](https://indico.cern.ch/event/587955/contributions/2938243/)
  o Scaling the EOS namespace, A. Manzi, [https://indico.cern.ch/event/587955/contributions/2936873/](https://indico.cern.ch/event/587955/contributions/2936873/)
  o Disk failures in the EOS setup at CERN - A first systematic look at 1 year of collected data, D. Duellmann, [https://indico.cern.ch/event/587955/contributions/2936875/](https://indico.cern.ch/event/587955/contributions/2936875/)
• G. Lo Presti, “Collaborative tools for CERNBox”, Nextcloud conference, Berlin, Germany, August 2018
• M. Lamanna “Data Storage at CERN”, Switch user meeting, Bern, Switzerland, September 2018,
• H. Gonzalez Labrador, "What is it special in CERNBox" and "From Application to Service development", OwnCloud conference, Nuremberg, Germany, September 2018
• “File Transfer Service at Exabyte Scale”, GRID 2018, Dubna, Russia, 10-14 September 2018, [https://indico.jinr.ru/contributionDisplay.py?sessionId=10&contribId=297&confId=447](https://indico.jinr.ru/contributionDisplay.py?sessionId=10&contribId=297&confId=447)
• “Latest developments of the CERN Data Management Tools”, HEPIX Fall 2018, Barcelona, Spain, 8-12 October 2018, [https://indico.cern.ch/event/730908/contributions/3153262/](https://indico.cern.ch/event/730908/contributions/3153262/)
• D. Castro, UP2University General Assembly (October, Germany)
• C. Contescu, "Storage at CERN", Hepix October 2018, Barcelona, Spain
• J. Moscicki “CERNBox architecture changes and OCM”, SIG-CISS Berlin, Berlin, Germany, October 2018,
• L. Mascetti, “CERN Storage”, IHEP, Beijing, China, October 2018, (and hand-on EOS training)
• M. Lamanna, “Collaborative analysis at a global scale with SWAN” at the eScience Infrastructure Ecosystem, eScience Conference, Amsterdam, Netherlands, October 2018
• D. van der Ster, "Mastering Ceph Operations: Upmap and the Mgr Balancer", Ceph day November 2018, Berlin, Germany
• Vladimír Bahyl, Benefits of using SSDs when repacking ~100 PB of data on tape @CERN, Fujifilm Summit, Chicago, USA
• Vladimír Bahyl and Oliver Keeble, WLCG Archival Storage group report, HEPiX Spring 2018, U. of Wisconsin, Madison, USA
• Daniel Lanza, ExDeMon: a new scalable monitoring tool for the growing CERN infrastructure, HEPiX Spring 2018, U. of Wisconsin, Madison, USA
• Germán Cancio, LTO experiences at CERN, HEPiX Fall 2018, U. Autonoma de Barcelona, Spain
• Julien Leduc, CERN Tape Archive initial deployments, HEPiX Fall 2018, U. Autonoma de Barcelona, Spain
• Aurélien Gounon, Backup Infrastructure at CERN, HEPiX Fall 2018, U. Autonoma de Barcelona, Spain
• SWAN: CERN’s Jupyter-based interactive data analysis service, Jupytercon 2018, New York, USA
• 2nd EOS Workshop 2018 at CERN with 15 oral presentations by development team members, https://indico.cern.ch/event/656157/timetable/#20180205
• XDC Allhands Meeting, EOS Smart Caching, Mihai Patrascoiu, https://indico.desy.de/indico/event/21131/session/1/material/1/8.pptx
• XDC Allhands Meeting, EOS – Adopting External Storage, Mihai Patrascoiu, https://indico.desy.de/indico/event/21131/session/1/material/1/9.pptx
• International symposium for GRIDs and CLOUDs, EOS Open Storage - the CERN storage ecosystem for scientific data repositories, http://indico4.twgrid.org/indico/event/4/session/15/contribution/50

Publications and Reports:

• M. Davis et al. CERN Tape Archive – from development to production deployment. CHEP 2018, Sofia, Bulgaria (to be published)
• H. Labrador, "CERNBox: the CERN cloud storage hub", CHEP proceedings (to be published)
• H. Rousseau at al. "Storage at CERN, CHEP proceedings (to be published)
• CS3 Book of Abstract (5 contributions and ST acting as editor): https://zenodo.org/communities/cs3
• C. Contescu, G. LoPresti, H. Rousseau, “Heavy-Ion run 2018”, internal report
Activities and Projects Reports 2018
2018 marked the start of a new three-year phase for CERN openlab. The foundations for this new phase were put in place in late 2017, with the publication of CERN openlab’s whitepaper on ‘future ICT challenges in scientific research’. This document — the result of an in-depth consultation process with a variety of stakeholders — identified a number of key challenges to be tackled in CERN openlab’s sixth phase (2018-2020).

New members join for new phase

IBM, E4, Micron, and Google all joined CERN openlab in 2018.

The first work to begin with IBM involves the evaluation of Power Architecture for potential machine-learning applications. Meanwhile, E4 are working with CERN on projects related to deep learning and artificial intelligence.

As part of their work with CERN, Micron will develop and introduce a specially designed memory solution that will be tested by researchers at CERN for use in rapidly combing through the vast amount of data generated by experiments.

The collaboration with Google, however, is still at an earlier stage. A number of exciting options for joint R&D projects are currently being explored, including possibilities related to cloud computing and machine learning.

As of the end of 2018, the list of members was as follows:

**Partners:** Intel, Oracle, Siemens, Huawei, Google, Micron

**Contributors:** Extreme Networks, Rackspace, IBM, E4

**Associates:** Comtrade, Open Systems

**Research members:** Eindhoven University of Technology, the European Bioinformatics Institute (EMBL-EBI), the European Society of Preventive Medicine, Fermilab, the Innovation Value Institute, the Italian National Institute for Nuclear Physics (INFN), King’s College London, Newcastle University, the SCimPULSE Foundation.

Major events

CERN openlab’s outreach activities have played a key role in helping to support this growth in membership. In March, CERN openlab launched a new website, bringing together information on around 20 joint R&D projects. In June, the CERN openlab CTO gave the opening keynote presentation at the ISC High Performance conference in Frankfurt, Germany — one of the most important annual events for the computing community. The event also saw a CERN openlab poster — presenting work with Intel to explore the feasibility of using deep-learning algorithms for the simulation of particle transport in the LHC experiments — selected for an award.
Maria Girone, CERN openlab CTO, gives the opening keynote presentation at ISC High Performance (Image: Roberta Maggi).

Another important event in 2018 was the first-of-its-kind workshop on quantum computing that CERN openlab organised in November. More than 400 people followed the workshop, which provided an overview of the current state of quantum-computing technologies. The event also served as a forum to discuss which activities within the HEP community may be amenable to the application of quantum-computing technologies. Following this event, CERN openlab established a new collaboration with IBM related to quantum computing. Discussions are also currently ongoing with Google regarding future potential work in this area.

Eckhard Elsen, Director for Research and Computing at CERN, addresses the CERN openlab workshop on quantum computing (Image: Andrew Purcell).
**Summer-student success**

The CERN openlab summer-student programme continued to go from strength to strength this year. In 2018, a record 1820 students applied to the programme. The 41 students selected took part in a series of lectures, visited sites at CERN and beyond, and worked on hands-on projects with the latest ICT solutions.

Find out more about CERN openlab’s summer-student programme, its joint R&D projects, and other interesting initiatives on the collaboration’ website. The website also features information on each of the companies and research institutes collaborating with CERN through CERN openlab.

**Presentations**

- A.L Krajewski/CERN, Extreme Networks project highlights, CERN openlab Technical Workshop, Geneva, Switzerland, 11 January 2018
- L. Breitwieser/CERN, BioDynaMo, University Hospital of Geneva Café de l’Innovation, Geneva, Switzerland, 01 February 2018
- L. Breitwieser/CERN, The Anticipated Challenges of Running Biomedical Simulations in the Cloud, Early Career Researchers in Medical Applications, CERN, Geneva, Switzerland, 12 February 2018
- F. Tilaro/CERN, F. Varela/CERN, Industrial IoT in CERN Control Systems, Siemens IoT Conference, Nuremberg, Germany, 21 February 2018
- B. Moreira/CERN, Moving from CellsV1 to CellsV2 at CERN, OpenStack Summit, Vancouver, Canada, 21 May 2018
- B. Moreira/CERN, Containers on Baremetal and preemtible VMs at CERN and SKA, OpenStack Summit, Vancouver, Canada, 24 May 2018
- D. H. Campora Perez/CERN Millions of circles per second. RICH at LHCb at CERN, University of Seville Seminar, Seville, Spain, 07 June 2018
- A.L Krajewski/CERN, Bro optimisations and network topologies, WLCG Security Operations Center WG Workshop/Hackathon, Geneva, Switzerland, 27 June 2018
- B. Moreira/CERN, Optimisations OpenStack Nova for Scientific Workloads, CHEP 2018, Sofia, Bulgaria, 10 July 2018
- G. Jereczek/CERN, The design of a distributed key-value store for petascale hot storage in data acquisition systems, CHEP 2018, Sofia, Bulgaria, 12 July 2018
- N. Nguyen/CERN, Distributed BioDynaMo, CERN openlab summer student ‘lightning talks’, Geneva, Switzerland, 16 August 2018
- F. Tilaro/CERN, F. Varela/CERN, Optimising CERN control systems through Anomaly Detection & Machine Learning, AI workshop for Future Production Systems, Lund, Sweden, 29 August 2018
- A. Hesam/CERN, Faster than the Speed of Life: Accelerating Developmental Biology Simulations with GPUs and FPGAs, Delft, Netherlands, 31 August 2018
• L. Breitwieser/CERN, The BioDynaMo Project: towards a platform for large-scale biological simulation, DIANA, Geneva, Switzerland, 17 September 2018

• F. Carminati/CERN, G. Khattak/CERN, S. Vallecorsa/CERN, Three-dimensional energy parametrized adversarial networks for electromagnetic shower simulation, 2018 IEEE International Conference on Image Processing, Athens, Greece, 07 October 2018

• T. Aliyev/CERN, Smart Data Analytics Platform for Science, i2b2 tranSMART Academic Users Group Meeting, Geneva, Switzerland, 31 October 2018


• F. Tilaro/CERN, F. Varela/CERN, Online Data Processing for CERN industrial systems, Siemens Analytics Workshop, Munich, Germany, 12 November 2018

• B. Moreira/CERN, S. Seetharaman/CERN, Scaling Nova with Cells V2, Presented at OpenStack Summit, Berlin, Germany, 13 November 2018

• D. Van Der Ster/CERN, Mastering Ceph Operations: Upmap and the Mgr Balancer, Ceph Day Berlin, Berlin, Germany, 13 November 2018

• T. Tsioutsias/CERN, Pre-emptible instances and bare metal containers, OpenStack Summit, Berlin, Germany, 15 November 2018

• J. Collet/CERN, Ceph@CERN, JTech Ceph day, Paris, France, 28 November 2018

• T. Aliyev/ CERN, AI in Science and Healthcare: Known Unknowns and potential in Azerbaijan, Bakutel Azerbaijan Tech Talks, Baku, Azerbaijan, 3 December 2018

Publications

• F. Carminati/CERN, G. Khattak/CERN, S. Vallecorsa/CERN, 3D convolutional GAN for fast simulation, CHEP 2018 Proceedings (submitted for publication)


Reports

• All reports from CERN openlab summer students listed here: https://zenodo.org/communities/cernopenlab/search?page=2&size=20# (reports uploaded starting from 28 September 2018)

Additional reports, publications, posters, and presentations related to each of the 18 CERN openlab R&D projects active in 2018 can be found on the CERN openlab website: https://openlab.cern/.
The CERN School of Computing (CSC) fosters the dissemination of knowledge and learning in the field of scientific computing. Its mission is to create a common culture in scientific computing among scientists and engineers involved in particle physics and other sciences. The CSC, along with the CERN Schools of Physics and the CERN Accelerator School, are the three schools that CERN has set up to help train the next generation of researchers across the laboratory’s main scientific and technical domains. Since the first CSC in Italy in 1970, the school has visited 22 countries and been attended by more than 2700 students from 5 continents and 80 nationalities. Participants come from many different backgrounds, but all share a passion for computing and science.

The CSC is composed of up to three schools per year, each one characterized by its own flavour and specific goals: the Main School, the Thematic School, and the Inverted School.

The Main School takes two weeks in late summer, and usually welcomes 50 to 80 students. It is a veritable “Summer University”, which offers both formal lectures and hands-on exercises. The practical part, where students work in pairs on common projects, is an important component of the learning process. Since 2002, the school offers a CSC diploma upon successful completion of an optional exam. Since 2008, the university hosting the CSC audits its academic programme and awards successful students five or six ECTS (European Credit Transfer System) credit points that are recognised across Europe for any doctoral and master programme. Additionally, various social and sport activities are proposed to students to socialise and establish lifelong links that will be useful throughout their careers. In particular, a daily optional sports programme helps maintaining a healthy work-life balance, while offering additional opportunities for interactions between students, lecturers and organisers.

Students of the CERN School of Computing 2018. (Photo: N. Kasioumis/CERN)
The 2018 CERN School of Computing (http://indico.cern.ch/e/CSC-2018) took place in Tel Aviv, Israel on October 1-14, and was organised together with Tel Aviv University (TAU). The school has welcomed 71 students (10 of which were female) from 41 different universities and institutes based on 21 countries worldwide. Representing 25 nationalities, these students were selected from a record number of 112 applicants. This year, the usual intensive academic programme (53 hours of lectures and exercises covering base technologies, physics computing and data technologies) was complemented by a number of optional activities, such as scientific visits at Tel Aviv University, guest lectures by Israeli speakers (including the inventor of USB memory sticks), and special evening lectures (History of Internet; Future of Humanity and of the Universe). In addition, a rich social programme was offered – an excursion to Jerusalem and the Dead Sea was clearly its highlight. At the end of the school, 65 students passed the optional exam – 15 of them with distinctions! - and were awarded 4 credit points by TAU.

The Thematic School (tCSC), organized since 2013, is a smaller school (one week, 20-30 students) focused on a specific, more advanced topic. In 2018, this topic was "High Throughput Distributed Processing of Future HEP Data". The 2018 Thematic CSC (http://indico.cern.ch/e/tCSC-2018) was held on June 3-9 in Split, Croatia. 28 students from 18 different institutes have followed a rich programme, consisting of 26 hours of lectures and exercises.

Finally, the Inverted School (iCSC), is a place where "students turn into lecturers": former CSC and tCSC attendees are invited to become teachers themselves and prepare series of lectures on advanced topics within their respective domains of expertise. This event, held in winter at CERN, usually lasts two to three days. Attendance to this event is open to anyone at CERN, and the lectures are also webcast. The 2018 Inverted CSC (http://indico.cern.ch/e/iCSC-2018) took place at CERN on March 5-8. A record of 194 people (20% female) registered for the event. The event was very well attended, with peaks up to 100 (in person + via webcast). 8 lecturers delivered 19 hours of lectures and exercises, on various topics including Blockchain and Decentralized Consensus, Backend
Systems, Complexity and Data Structures, Data Analysis, Identity Federation, Medical Imaging, Parallel Programming, and more.

The practical part, where students work in pairs on common projects, is an important component of the learning process. (Photo: S. Lopienski/CERN)

In 2019, three CSC schools are planned:

- the main CERN School of Computing 2019 (September 15-28 in Cluj-Napoca, Romania), organized jointly with Babeş-Bolyai University (UBB) together with Politehnica University of Bucharest (UPB);
- Thematic CSC 2019 (May 12-18 in Split, Croatia), organized together with the University of Split; and
- Inverted CSC 2018 (March 4-6 at CERN).

More information at [http://cern.ch/csc](http://cern.ch/csc)
The mandate of the Computer Security Team is to “protect the operations and reputation of CERN against cyber-threats”. As the years before, the Team was deeply involved in and usually leading a series of computer security investigations, including forensic analyses, helping CERN’s internal audit and legal service, or understanding computer break-ins at CERN or within the community (e.g. the WLCG, EGI, AARC, other universities). Mayor consequences were averted such that no major (externally visible) impact been recorded for CERN so far this year.

In order to prevent incidents (in particular so-called “APTs”), the Team has established a large network of peers within the community as well as with law enforcement agencies and industry. Special successes were dedicated training sessions for our peers in China. Intelligence gathered through this network will be channelled back to the Team’s detection infrastructure which is supposed to analyze online about one terabyte of logs stemming from different sources (e.g. network/firewall, DNS, computing cluster, web servers, SSO). This infrastructure is based on standard CERN IT provisionings and is thoroughly discussed for re-usage within the community. As a start, CERN is already acting as a dissemination point for sharing this intelligence within the community.

On the protective side, the Computer Security Team has led the analysis of a world-wide Linux malware in collaboration with colleagues from the EGI and WLCG and coordinated the mitigation of a large number of security vulnerabilities (e.g. acron, HP iLO, Singularity, Smash, SMBv1, Spectre/Meltdown/L1TF, SSL configurations) within the Organization. Furthermore, the Team ran a series of severe security investigations linked to CERN-hosted office PCs and laptops as well as to CERN central computer center services, and assisted in numerous fraud & theft investigations.

As a continuous process, the Computer Security Team has assisted to further secure different computing services within the CERN data centre, in the office domain of CERN, as well as of systems used for controlling accelerators or experiments at CERN. Numerous vulnerability warnings as well as corresponding mitigation recommendations were disseminated to the CERN staff and users as well as to the community. More proactively, the Team has deployed in collaboration with CERN’s Windows desktop support team more than 250 hardened PCs used in the administrative and financial sectors. Compared to previous years, the infection rate on those PCs has significantly dropped. Furthermore, the Team has conducted several security audits on computer systems used in all areas at CERN. Computer security trainings and awareness sessions were held through-out CERN including a full scale “Clicking” campaign intended to sensitize people of the risk of clicking on arbitrary links. The Computer Security Team’s “WhiteHat Challenge” has now taken full speed in educating and training CERN’s staff and users in order to become penetration testers for their software.

Presentations:

- "Building and operating a large scale Security Operations Centre", SWITCH Security Working Group Meeting, Bern, Switzerland, 31 January 2018
- "Building and operating a large scale Security Operations Centre", CH CERTs Forum Meeting, Canton Vaud, Switzerland, 2 February 2018
"IoT within the CERN accelerator complex", SWITCH IoT Workshop, Zurich, Switzerland, 21 February 2018
"CERN's Computer Security Operations Centre", CERN IT Technical Forum, CERN, Geneva, Switzerland, 9 March 2018
"Finding the Balance", open.ch, Zurich, Switzerland, 15 March 2018
"Building a large scale Intrusion Detection System using Big Data technologies", International Symposium on Grids & Clouds 2018, Taipei, Taiwan, 22 March 2018
"Finding the Balance", Swiss Federal Department of Defence, Civil Protection and Sport, Bern, Switzerland, 19 April 2018
"CERN Computer Security", Lectures at FH Nordwest Schweiz, Windisch, Switzerland, 24 April 2018
"Situational Awareness: Security (& Privacy)", HEPix Spring 2018 workshop, Madison WI/USA, 14-18 May 2018
"Why (Control System) Cyber-Security sucks", IoT week, Bilbao, Spain, 4-8 June 2018
"The Ever Emerging Threats", IoT week, Bilbao, Spain, 4-8 June 2018
"Phishing/clicking campaign and user-facing SIEM interface", SWITCH joint Security-Network WG, Fribourg, Switzerland, 5-6 June 2018
"Finding the Balance", Swiss Federal Department of Defence, Civil Protection and Sport, Bern, Switzerland, 20 June 2018
"Honeypot Resurrection", CERN Summer Student Poster Session, Geneva, Switzerland, 1 August 2018
"Introduction to Security" & "Practical Security and Cryptography" Lectures, CODATA RDA Research Data Science Summer School, Trieste, Italy, 13 August 2018
"Bridging the gap between ICS/IoT and corporate IT security", CRITIS 2018, Kaunas, Lithuania, 24-26 September 2018
"Why (Control System) Cyber-Security Sucks", ShareIT, Valmiera, Latvia, 5 October 2018
"Computer Security Update", HEPix Autumn 2018 workshop, Barcelona, Spain, 8-12 October 2018
"Finding the Balance", IKT Sicherheitskonferenz, Alpbach, Austria, 16-17 October 2018
"Finding the Balance", Airbus Security Symposium, Toulouse, France, 18 October 2018
see also https://cern.ch/security/reports/en/presentations.shtml

Publications:

- Melchior Thambipillai: "Anomaly-based Intrusion Detection System for the CERN Security Operations Center", MSc, Ecole Polytechnique Fédérale de Lausanne, Switzerland, 8/2018
- Fabiola Buschendorf: "Honeypot Resurrection - Redesign of CERN's Security Honeypots"
- Shivam Kapoor: "Malware Analysis Management"
During 2018, a major concern regarding worldwide HEP data preservation was raised by the BaBar experiment at SLAC. The SLAC directorate have decided that the BaBar data can no longer be stored (and hence no longer preserved at SLAC) for more than the next 1-2 years. This was reported to the ICFA meeting in Cambridge in March (DPHEP continues to operate as an ICFA panel: its scope is much wider than either CERN IT or CERN itself.

Fortunately, the CERN directorate have agreed that this data could be stored at CERN and plans are being prepared to transfer the data during the Long Shutdown 2 of the LHC.

Following on from the work by the HEP Software Foundation on the Community White Paper, a DPHEP session was held in conjunction with the WLCG / HSF workshop in Naples in March. The agenda and presentations can be found at https://indico.cern.ch/event/658060/sessions/266382/#20180327.

In early 2018, a self-assessment of CERN’s various data preservation activities against the metrics in ISO 16363 (in order to obtain certification as a Trustworthy Digital Repository) was submitted to an external auditor. Feedback was received in August 2018 and an action plan developed to address the concerns raised.

Of particular importance is the status of the proposed Operational Circular on Scientific Data Preservation that is considered a “show-stopper” to obtain certification. Given the delays in approval of an OC, it has been suggested that a Policy Document be prepared instead and/or as an initial step. This will need to be pursued with some urgency in 2019.

Following the successful approval of the H2020 ARCHIVER project (described elsewhere), it was agreed that CERN would join the Digital Preservation Coalition for an initial period of 3 years (covering the lifetime of the ARCHIVER project). This should help address a variety of open questions, including training on long-term preservation issues as well as guidance in procurement of LTDP services – an area where the DPC has a particular focus.

Presentations:

- See https://indico.cern.ch/category/4458/.

Publications:

- The HEP Community White Paper can be found at https://arxiv.org/pdf/1804.03983.pdf
- The Data Preservation chapter has been published separately: https://cds.cern.ch/record/2643937?ln=en
- Collaborative Long-Term Data Preservation: From Hundreds of PB to Tens of EB (https://cds.cern.ch/record/2642194?ln=en) - presented at the PV 2018 conference
EXTERNALLY FUNDED PROJECTS

The Externally Funded Projects (EFP) section of DI group hosts the IT EC Project Office that performs oversight and support functions necessary to coordinate IT department’s engagement in European Union projects. The objective of CERN’s participation in the Horizon 2020 work programme is to develop policies, technologies and services that can support the Organization’s scientific programme, promote open science and expand the impact of fundamental research on society and the economy. IT department is actively engaged in 11 Horizon 2020 projects.

The most recently started projects include:

- ARCHIVER: Archiving and Preservation for Research Environments
- ESCAPE: European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures
- EOSCsecretariat.eu: European Open Science Cloud secretariat
- OCRE: Open Cloud for Research Environments

All of these projects are contributing to establishing the European Open Science Cloud (EOSC). The EOSC initiative has been proposed in 2016 by the European Commission as part of the European Cloud Initiative to build a competitive data and knowledge economy in Europe. The EOSC will offer 1.7 million European researchers and 70 million professionals in science, technology, the humanities and social sciences a virtual environment with open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines by federating existing scientific data infrastructures, currently dispersed across disciplines and the EU Member States.

Figure 1: The European Open Science Cloud
In 2018, the LHC performance again reached new levels, with a total integrated luminosity in pp running of \( \sim 65 \text{ fb}^{-1} \) delivered to ATLAS and CMS, exceeding the target for the year. This was followed by a significant heavy ion run at the end of the year. In total about 88 PB of LHC data were written to tape at CERN, with 17 PB of that produced during the heavy ion run. This was a significant increase over the previous year’s acquisition of 50 PB. Figure 3 shows the monthly data ingest to the Tier 0 archive system, showing new peak levels in 2018. As in previous years, the global WLCG infrastructure has performed stably and well, and has managed the significantly increased needs of the experiments, supporting the timely delivery of high-quality physics results.

The Tier 0 continued to provide the essential processing and data quality monitoring during data taking, without any of the backlogs of work seen during previous years. Globally delivered compute time across the WLCG collaboration has continued to grow (Figure 4). Data delivery to the global collaboration is a key aspect of the WLCG service, and as shown in Figure 5 the transfer rates have also reached new levels of around 60 GB/s with very smooth performance, reliably delivering data to the 170 participating WLCG computer centres worldwide.

Looking forward to Run 3 where the integrated luminosity is anticipated to equal that taken in Run 1 and Run 2 combined, the estimates of needed computing resources still fit within the constrained budgets anticipated, largely due to performance and efficiency improvements in the overall WLCG service, in the experiments’ core software, and through evolution of their computing models.

Looking further forward to HL-LHC, the estimates of the experiments for the computing needs have become better quantified over the last year, and while there is still a significant challenge, the various strategies of the WLCG collaboration and the experiments give an optimistic view of being able to manage the challenge. These strategies have been documented for the LHCC\(^1\) and form the basis for a number of R&D projects that have been started during 2018 and that will continue over the coming years. These cover many aspects from data management, technology evolution, and networking to key software challenges that must be collaboratively addressed. In particular, the use of new computer architectures is the subject of a number of different initiatives together with the HEP Software Foundation\(^2\).

The WLCG team has also worked closely with other large-data sciences. In particular, the collaboration with SKA is important, and has continued. Together with SKA and other major HEP and astronomy ESFRI projects we have successfully proposed a project to prototype a European data infrastructure for Exabyte-scale sciences, fulfilling the needs of open science. The ESCAPE\(^3\) project has been funded and will commence in early 2019, with the CERN WLCG team leading the infrastructure work package. This is fully in line with the stated strategy required for HL-LHC data management.

---


\(^2\) [https://hepsoftwarefoundation.org](https://hepsoftwarefoundation.org)

\(^3\) ESCAPE project: [https://escape2020.eu](https://escape2020.eu)
Figure 3: Data Acquired at the Tier 0, showing new records in 2018, up to 14 PB/month

Figure 4: Global CPU time delivered, showing continual increase; this corresponds to ~800k cores in 2018

Figure 5: Global data transfers, also reaching new levels up to 60 GB/s
Posters:


Presentations:

- I. Bird/CERN, Challenges of Big Data in Science, Invited Seminar, Kavli Institute for Astronomy, University of Cambridge, UK, 1 February 2018
- A. Valassi/CERN, "Fisher information metrics for binary classifier evaluation and training", 2nd Inter-Experimental LHC Machine Learning Working Group (IML) Workshop, CERN,
• X. Espinal/CERN-IT, Datalake R+D project at CERN, 2nd GEANT SIG-CISS, Amsterddam, Netherlands, 11 April 2018, https://wiki.geant.org/display/CISS/2nd+SIG-CISS+meeting
• A. Valassi/CERN, "Binary classifier metrics for event selection optimization in HEP", 23rd International Conference on Computing in High-Energy and Nuclear Physics (CHEP2018),


D. Smith; S. Muralidharan, Behind the scenes perspective: Into the abyss of profiling for performance, GridKa 2018, KIT, Karlsruhe, Germany, 27-31 August 2018, https://cds.cern.ch/record/2637493


I. Bird/CERN, WLCG Experience, Talk at AENEAS all-hands meeting, Bologna, Italy, 10 October 2018, https://indico.astron.nl/conferenceTimeTable.py?confId=172#20181010

• A. Sciaba'/CERN, et al., Cost and system performance modelling in WLCG and HSF: an update, HEPiX Autumn/Fall 2018 Workshop, Casa de Convalescència UAB, Barcelona, Spain, 11 October 2018, https://indico.cern.ch/event/730908/contributions/3148853/
• M. Sharma/IT-DI-LCG, MOOC-CA: Choosing the right person for the job!, CERN IT Lightning Talks #17, Geneva, Switzerland, 9 November 2018, https://cds.cern.ch/record/2646606
• M. Sharma/IT-DI-LCG, M. Litmaath/IT-DI-LCG. The SIMPLE Framework - simplifying complex container clusters, PyParis 2018, Paris, France, 15 November 2018, https://speakerdeck.com/maany/the-simple-framework-deploy-complex-clusters-with-ease (video: https://www.youtube.com/watch?v=iNDx0r9Zk&list=PLzjFI0G5nSry3cm_k1tPoi95RaAXsZAt&index=14)
• I. Bird/CERN, Distributed Computing for LHC, Invited talk at Celebration of 10 years South Africa-CERN Collaboration, iThemba LABS, Cape Town, South Africa, 19 November 2018, https://indico.tlabs.ac.za/event/73/

Publications:
• A. Valassi/CERN, "Binary classifier metrics for optimizing HEP event selection", Submitted to EPJ Web of Conferences, Proceedings CHEP2018, Sofia, Bulgaria
• D. Smith et al., "Sharing server nodes for storage and compute", Submitted to EPJ Web of Conferences, Proceedings CHEP2018, Sofia, Bulgaria
• J. Elmsheuser et al., "ATLAS Grid Workflow Performance Optimization", Submitted to EPJ Web of Conferences, Proceedings CHEP2018, Sofia, Bulgaria
  http://cds.cern.ch/record/2645396
• C. Biscarat et al., "System performance and cost modelling in LHC computing", Submitted to EPJ Web of Conferences, Proceedings CHEP2018, Sofia, Bulgaria, https://hal.archives-ouvertes.fr/hal-01835622/

Reports: