Annual Report
2016

CERNopenlab

Contents

A Word from the DG

2016 marked 15 years of CERN openlab. Founded in 2001 to help develop the innovative systems needed to cope with the unprecedented ICT challenges of the Large Hadron Collider (LHC), CERN openlab unites science and industry at the cutting edge of research and innovation.

Collaboration is central in enabling CERN to fulfil its mission; CERN openlab is a prime example of this. For 15 years, this unique public-private partnership has worked to ensure that members of CERN's scientific community have access to the very latest ICT solutions to help them carry out their challenging physics research. I would like to thank each of the companies collaborating in CERN openlab — as well as, most importantly, the people themselves — for their terrific efforts in supporting CERN's work.

Throughout its existence, both education and training have been central to CERN openlab's mission. The young researchers hired by CERN and funded by the collaborating companies continue to play a key role in CERN openlab's work; they are the catalyst for many creative ideas. To date, the CERN openlab Summer Student Programme alone has seen over 270 students — selected through a highly competitive process — come to the laboratory.

2016 was also an important year for looking forward. CERN openlab is continuing its work to support our research community, with a particular focus on the upgrades to the LHC and the experiments that will be carried out during the next two 'Long Shutdown' periods, known as 'LS2' and 'LS3'. With the data rates from the experiments set to increase significantly, efforts have been focused on supporting the work to overhaul and modernise their data-acquisition systems. Work was also invested in ensuring that the maximum benefits are gained from the available hardware by contributing to making sure the software running on it has been fully optimised.

Work has now begun to identify the ICT challenges that will need to be tackled in CERN openlab's sixth phase, which will run from 2018 to 2020. I am sure CERN's scientific community will continue to greatly benefit from the solutions this unique public-private partnership produces.

The LHC produced 7,000,000,000,000,000 proton-proton collisions in 2016.

The Context

Records smashed

The LHC's performance exceeds expectations, producing a deluge of data.

In April 2015, the Large Hadron Collider (LHC) restarted after a two-year shutdown period, during which major upgrade work was carried out. Throughout the remainder of the year, records were broken as the LHC was ramped up to higher energy and luminosity.

2016 continued to see records fall as the LHC's performance soared to new heights. The rigorous consolidation and upgrade work carried out over recent years — for both machines and processes — enabled the LHC operators to surpass operational availability targets for the accelerator complex. This meant that more physics data could be taken for longer. The number of particle collisions recorded by the ATLAS and CMS experiments during 2016's seven-month proton run was 60% higher than anticipated. That equates to more data than was collected in the previous three proton runs combined. Overall, the LHC experiments observed around 7 quadrillion collisions, at an energy of 13 TeV.

In technical terms, the integrated luminosity received by the ATLAS and CMS experiments reached 40 inverse femtobarns (fb^{-1}), compared with the $25fb^{-1}$ originally planned. Luminosity is an essential indicator of the performance of an accelerator, measuring the potential number of collisions that can occur in a given amount of time, and integrated luminosity is the accumulated number of potential collisions. At its peak, the LHC's proton-proton collision rate reaches about 1 billion collisions per second, giving a chance that even the rarest processes at the highest energy could occur. An estimated 16,000 Higgs bosons have been produced by collisions in the LHC to date, compared to around just 400 at the time of the announcement of the discovery of the particle back in 2012.

The CERN Data Centre received a staggering 50 petabytes of data from the LHC experiments in 2016. Having nearly doubled during 2015 to reach 150,000 compute cores, the size of CERN's private cloud was increased further in 2016 to handle the computing needs of the CERN experiments and services. By the end of 2016, it provided over 190,000 compute cores running across the data centres in Meyrin, Switzerland, and Budapest, Hungary.

Much effort was also needed to handle the data volumes sent from CERN to the sites in the Worldwide LHC Computing Grid (WLCG), a global computing infrastructure used to store, distribute, and analyse the data. Data transfer rates around the globe reached new peaks: up to 50-gigabytes-per-second continuous rates. This is double the rate that was typical during the LHC's first run (2009-2013).

Delegates at the 2016 meeting of the CERN openlab Collaboration Board.

The Concept

Catalysing collaboration

CERN openlab is a unique public-private partnership that accelerates the development of cutting-edge solutions for the worldwide LHC community and wider scientific research. Through CERN openlab, CERN collaborates with leading ICT companies and research institutes.

Within the CERN openlab framework, CERN provides access to its advanced ICT infrastructure and its engineering experience — in some cases even extended to collaborating institutes worldwide. Testing in CERN's demanding environment provides the collaborating companies with valuable feedback on their products, while enabling CERN to assess the merits of new technologies in their early stages of development for possible future use. This

framework also offers a neutral ground for carrying out advanced R&D with more than one company.

Collaboration can be at the associate, contributor, or partner level. Each status represents a different level of investment, with projects lasting typically between one and three years. The collaborating companies engage a combination of cash and in-kind contributions, with the cash being used to hire young ICT specialists dedicated to the projects. The associate status formalises a collaboration based on independent and autonomous projects that do not require a presence on the CERN site. The contributor status is a collaboration based on tactical projects, which includes a contribution to hire a young ICT specialist supervised by CERN staff to work on the common project, in addition to the hardware and software products needed by the projects. The partners commit to a longer-term, strategic programme of work and provide three kinds of resources: salaries for young researchers, products and services, and engineering capacity.

CERN openlab was established in 2001, and has been organised into successive three-year phases. In the first phase (2003–2005), the focus was on the development of an advanced computing-cluster prototype called the 'opencluster'. The second phase (2006–2008) addressed a wider range of domains. The combined knowledge and dedication of the engineers from CERN and the collaborating companies produced exceptional results, leading to significant innovation in areas such as energy-efficient computing, grid interoperability, and network security. CERN openlab's third phase (2009-2011) capitalised and extended upon the successful work carried out in the second phase. New projects were added focusing on virtualisation of industrial-control systems and investigation of the then-emerging 64-bit computing architectures. The fourth phase (2012-2014) addressed new topics crucial to the CERN scientific programme, such as cloud computing, business analytics, next-generation hardware, and security for the ever-growing number of networked devices.

This annual report covers the second year of CERN openlab's fifth phase (2015-2017). This phase is tackling ambitious challenges covering the most critical needs of ICT infrastructures in domains such as data acquisition, computing platforms, data storage architectures, compute provisioning and management, networks and communication, and data analytics. The work carried out in

Members of the CERN openlab management team.

2016 built on strong foundations established in these domains during the first year of CERN openlab's current phase. However, it also expanded upon these through its knowledge-sharing projects with communities beyond high-energy physics.

CERN openlab has grown to include more collaborating companies, thus enabling a wider range of challenges to be addressed. New research institutes have also joined CERN openlab in its endeavour to accelerate the development of ICT solutions that support the research community.

Each CERN openlab team is supervised by a project coordinator, who liaises with the collaborating company. At the annual technical workshops, representatives of the collaborating companies and research institutes meet with the teams, who present their latest results, and consider possible synergies. At the annual collaboration board meetings, the board receives information and exchanges views on the progress and medium-term plans of CERN openlab. In addition, CERN openlab celebrated its 15th anniversary in 2016 with a public 'open day' event held in June.

The CERN openlab team consists of three complementary groups of people: young engineers hired by CERN and funded by the partners, technical experts from partner companies involved in the projects, and CERN management and technical experts working partly or fully on the joint activities. A list of the people across CERN most closely involved in the CERN openlab activities is given on page 11, while the positioning of CERN openlab activities within CERN is detailed on pages 12 and 13.

The distributed nature of the CERN openlab team enables close collaboration with computing experts in the LHC experiments, as well as with engineers and scientists from CERN openlab collaborators, who contribute significantly to our activities. Valuable contributions are also made by the students participating in the highly selective CERN openlab Summer Student Programme, either directly to CERN openlab activities or more widely to the WLCG and other CERN activities in the IT Department.

CERN openlab CRO Fons Rademakers presenting at the 2016 Collaboration Board.

CERN openlab Management

Fabiola Gianotti	CERN Director-General, Chair of CERN openlab Collaboration Board
Frédéric Hemmer	Head of CERN IT Department
Alberto Di Meglio	Head of CERN openlab
Maria Girone	Chief Technology Officer
Fons Rademakers	Chief Research Officer
Andrew Purcell	Communications Officer
Maria-Athanasia Pachou	Junior Communications Officer
Kristina Gunne	Administrative and Finance Officer
Sotirios Pavlou	Junior Administrative Officer
Marco Manca	Senior Medical Fellow

CERN openlab personnel (Collaborator Indicated)

Omar Awile	Intel (Staff)
Lukas Breitwieser	Intel (Technical Student)
Placido Fernandez Declara	Intel (Doctoral Student)
Christian Faerber	Intel (Fellow)
Andrei George Gheata	Intel (Staff)
Karel Ha	Intel (Technical Student)
Christina Quast	Intel (Technical Student)
Sebastien Jean Valat	Intel (Fellow)
Sofia Vallecorsa	Intel (Visiting Scientist)
Luis Rodriguez Fernandez	Oracle (Staff)
Liana Delia Lupsa	Oracle (Fellow)
Antonio Romero Marin	Oracle (Fellow)
Manuel Martin Marquez	Oracle (Staff)
Antonio Nappi	Oracle (Fellow)
Flavia Castro Alves	Siemens (Technical Student)
Carlos Garcia Calatrava	Siemens (Technical Student)
Alejandro Herrero De Campos	Siemens (Technical Student)
Jakub Guzik	Siemens (Fellow)
Piotr Jan Seweryn	Siemens (Technical Student)
Filippo Tilaro	Siemens (Staff)
Adam Lukasz Krajewski	Brocade (Fellow)
Ioannis Charalampidis	Cisco (Fellow)
Lazaros Lazaridis	Cisco (Fellow)
Simaolhoda Baymani	IDT (Fellow)
Marek Kamil Denis	Rackspace (Fellow)
Spyridon Trigazis	Rackspace (Fellow)
Paul Lensing	Seagate (Cooperation Associate)

CERN contributing personnel (Collaborator Indicated)

Luca Atzori	Intel
Daniel Campora	Intel
Paolo Durante	Intel
Flavio Pisani	Intel
Rainer Schwemmer	Intel
Balazs Voneki	Intel
Zbigniew Baranowski	Oracle
David Collados	Oracle
Andrei Dumitru	Oracle
Katarzyna Maria Dziedziniewicz-Wojcik	Oracle
Lorena Lobato	Oracle
Nicolas Marescaux	Oracle
Emil Pilecki	Oracle
Giacomo Tenaglia	Oracle
Pavel Fiala	Siemens
Piotr Golonka	Siemens
John Bradford Schofield	Siemens
Jean-Charles Tournier	Siemens
Axel Voitier	Siemens
Edoardo Martelli	Brocade
Stefan Stancu	Brocade
Konstantinos Alexopoulos	IDT

ICE-DIP Fellows

Grzegorz Jereczek
Przemyslaw Karpinski
Aram Santogidis
Srikanth Sridharan
Marcel Zeiler

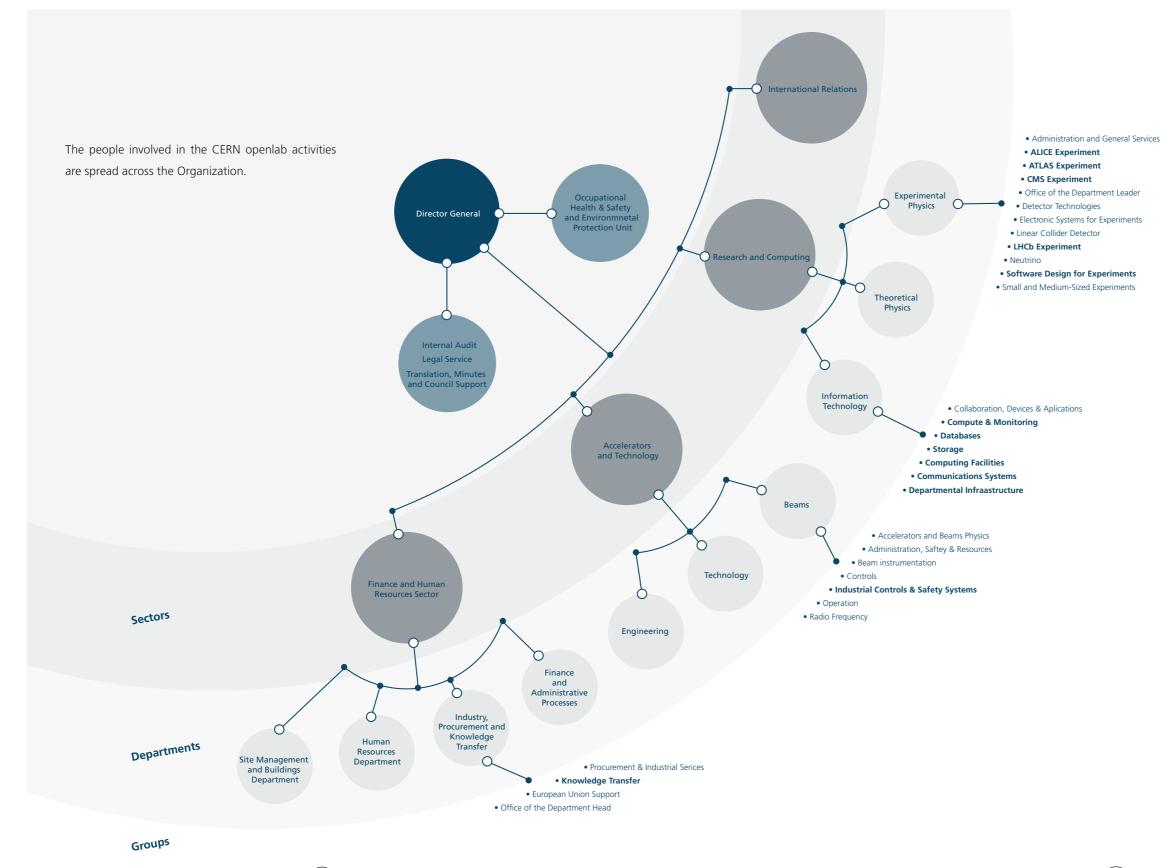## CERN openlab project coordinators (Collaborator indicated)

Olof Bärring	Intel and IDT
Federico Carminati	Intel
Niko Neufeld	Intel
Luca Canali	Oracle
Eric Grancher	Oracle
Eva Dafonte Pérez	Oracle
Artur Wiecek	Oracle
Manuel Gonzalez Berges	Siemens
Tony Cass	Brocade
Predrag Buncic	Cisco
Tim Bell	Rackspace
Dirk Duellmann	Seagate and Huawei
Xavier Espinal Curull	Comtrade
Stefan Roiser	Yandex

Collaborator liaisons with CERN

Claudio Bellini	Intel
Laurent Duhem	Intel
Jonathan Machen	Intel
Klaus-Dieter Oertel	Intel
Marie-Christine Sawley	Intel
Karl Solchenbach	Intel
Jean-Philippe Breysse	Oracle
Andrew Bulloch	Oracle
Vincent Burguburu	Oracle
Joel Cherkis	Oracle
Dmitrij Dolgušin	Oracle
David Ebert	Oracle
Pauline Gillet-Mahrer	Oracle
Rachael Hartley	Oracle
Mark Hornick	Oracle
Cris Pedregal	Oracle
Elisabeth Bakany	Siemens/ETM
Thomas Hahn	Siemens
Mikhail Roschin	Siemens
Ewal Sperrer	Siemens/ETM
Mohammad Hanif	Brocade
Pierre Hardoin	Brocade
Artur Barczyk	Cisco
Frank Van Lingen	Cisco
Trevor Hiatt	IDT
Devashish Paul	IDT
Keith Bray	Rackspace
Jason Cannavale	Rackspace
Giri Fox	Rackspace
Adrian Otto	Rackspace
Paul Voccio	Rackspace
Paul Kusbel	Seagate
Goran Garevski	Comtrade
Gregor Molan	Comtrade
Rade Radumilo	Comtrade
Vedran Vujinovic	Comtrade
Thomas Schoenemeyer	Huawei
Andrey Ustyuzhanin	Yandex
Steven Newhouse	EMBL-EBI
Mohammad Al-Turany	GSI
Gaetano Maron	INFN
Manuel Mazzara	Innopolis University
Max Talanov	Kazan Federal University
Mario Falchi	King's College London
Roman Bauer	Newcastle University

Positioning CERN openlab activities at CERN

The people involved in the CERN openlab activities are spread across the Organization.

- Director General
- Occupational Health & Safety and Environmnetal Protection Unit
- Internal Audit
- Legal Service
- Translation, Minutes and Council Support
- International Relations
- Research and Computing
  - Experimental Physics
    - Administration and General Services
    - **ALICE Experiment**
    - **ATLAS Experiment**
    - **CMS Experiment**
    - Office of the Department Leader
    - Detector Technologies
    - Electronic Systems for Experiments
    - Linear Collider Detector
    - **LHCb Experiment**
    - Neutrino
    - **Software Design for Experiments**
    - Small and Medium-Sized Experiments
  - Theoretical Physics
  - Information Technology
    - Collaboration, Devices & Aplications
    - **Compute & Monitoring**
    - **Databases**
    - **Storage**
    - **Computing Facilities**
    - **Communications Systems**
    - **Departmental Infraastructure**
- Accelerators and Technology
  - Beams
    - Accelerators and Beams Physics
    - Administration, Saftey & Resources
    - Beam instrumentation
    - Controls
    - **Industrial Controls & Safety Systems**
    - Operation
    - Radio Frequency
  - Technology
  - Engineering
- Finance and Human Resources Sector
  - Finance and Administrative Processes
  - Industry, Procurement and Knowledge Transfer
    - Procurement & Industrial Serices
    - **Knowledge Transfer**
    - European Union Support
    - Office of the Department Head
  - Human Resources Department
  - Site Management and Buildings Department

Sectors

Departments

Groups

CERN openlab CTO Maria Girone presents results at the main annual technical workshop.

The Results

The results of the 16 CERN openlab projects active in 2016 are organised by technical challenge.

Data Acquisition (online)

Existing and emerging large-scale research projects are producing increasingly high amounts of data at ever-faster rates. A prime example of this comes from CERN's LHC, which produces millions of particle collisions every second in each of its detectors, thus generating approximately 1 PB of data per second.

Intel High-Throughput Computing Collaboration

In 2016, the High-Throughput Computing Collaboration (HTCC) continued its studies of three new Intel data-centre technologies: Intel® Omni-Path, Intel® Xeon™/FPGA, and Intel® Xeon Phi™.

Intel Omni-Path is used for data-exchange at 100 gigabits per second (Gb/s) between servers for data-acquisition systems. We performed a first large-scale test at the Marconi supercomputer hosted by the Italian inter-university computing consortium 'Cineca'.

Intel Xeon/FPGA integration was used to decompress and reformat packed binary data from the detectors and to accelerate key kernels of the event-filtering algorithm base. To demonstrate the utility of FPGA acceleration, a floating-point heavy, complex algorithm was used to perform photon reconstruction in a large detector. To enable better comparison, the algorithm was implemented in both Verilog hardware-description language and in Intel OpenCL for FPGAs. In both cases, very competitive acceleration results (up to a factor of 35) over the reference were achieved. The HTCC was also able to demonstrate that FPGAs are extremely power-efficient compared to other programmable technologies (up to a factor of four). In addition, the tight Xeon integration of Intel's new product effectively removes the 'PCIe-bottle-neck', thus overcoming one of the main challenges in using FPGAs as accelerators: the need to 'feed' them with sufficient amounts of data.

In 2016, Intel released the next generation of its Xeon Phi processors, which has the code name 'Knights Landing'. Some of our most computationally expensive algorithms have been tested on these processors, with significant speedups being observed in certain cases. Work is now being carried out to further optimise our algorithms for running on this platform. Many of the improvements we have obtained during this process are also applicable to running on other Intel Xeon processors — this is particularly true in the case of very-wide vector registers (AVX and AVX2).

Promising results from first scaling test for data-acquisition event-building with Intel Omni-Path, up to 512 nodes. The results were obtained using two benchmarks: DAQPIPE and LSEB (which is simpler). DAQPIPE was used in a range of modes: one can use a barrel-shifting or random strategy in DAQPIPE for the communication scheduling. Finally, we also tried a non-full-duplex communication pattern. In 2017, work will be undertaken to further improve performance through better understanding the possible effects of the network configuration.

IDT RapidIO for Data Acquisition

RapidIO is an open-standard system-level interconnect, which is today used in all 4G/LTE base stations worldwide. RapidIO is often used in embedded systems that require high reliability, low latency, low energy consumption, and scalability in a heterogeneous environment. The collaboration with IDT, a major provider of RapidIO technology, started in June 2015. The project, which is investigating the potential of the interconnect in new domains, took several steps forward in 2016.

In terms of hardware installations, our machine pool saw extensions and improvements: a 32-port top-of-rack switch was installed in the CERN Data Centre, as well as 16 server nodes equipped with RapidIO-PCIe network interface controllers.

In terms of software installations, the IDT team made regular updates to our Linux modules and libraries. Two of these updates were especially relevant: (i) the introduction of reserved memory, which enables applications to allocate any amount of memory for RapidIO operations; (ii) the introduction of the riosocket drivers, which expose a standard TCP/IP interface running on top of RapidIO, thus enabling us to use any TCP/IP compatible application (such as iperf or Hadoop) without any porting work at all.

CERN openlab fellow Simaolhoda Baymani presents results from the project with IDT.

The work of the project this year has been focused on the use cases of data analytics and data acquisition. In addition, an initial investigation into multicast was carried out by a 2016 CERN openlab summer student.

Within the data analytics use case, the porting of ROOT to RapidIO has undergone big changes. We have gone from using a simple transfer protocol at the beginning of the year to a more efficient implementation that makes use of circular buffers and RapidIO doorbells, meaning quick notifications can be sent between nodes. With the introduction of riosockets and the performance improvement they provide, the Hadoop installation was also able to move from the test cluster to the server cluster in the CERN Data Centre.

Significant strides were made for the data acquisition use case, too. 2016 saw the main work start for the LHCb experiment's benchmark DAQPIPE, which is an interconnect-agnostic application that emulates the behaviour in the LHCb data-acquisition network. We ported DAQPIPE to RapidIO and are conducting benchmark runs to evaluate observations related to the RapidIO technology's core.

The project results and experiences have been shared through presentations and publications, including at the 20th IEEE Real Time Conference and at the 22nd International Conference on Computing in High Energy and Nuclear Physics.

Networks and Connectivity

Today, the ever-increasing external data traffic (more than 100 Gb/s) is putting pressure on the CERN firewalls. CERN and other large organisations require robust, scalable network solutions that provide high bandwidth for data transfers while maintaining appropriate levels of security. Networking also plays a vital role in data acquisition from the experiments, and is paramount to the success of the WLCG.

Brocade Flow Optimizer

The Brocade Flow Optimizer (BFO) project aims to enhance and generalise Brocade's BFO application, and to use it to build an intelligent steering system for network traffic. The system will be capable of supporting use cases that are not easily handled using traditional networking approaches. These use cases include intrusion detection system (IDS) mirroring, a firewall load-balancer, and an advanced policy-based routing engine. To achieve this, Brocade is funding an engineer at CERN, who is fully integrated in the BFO development team. His contributions, which extend beyond the IDS use case, have been successfully merged into three official BFO releases.

BFO is a software-defined networking (SDN) application designed to improve visibility of network traffic and to enable flow-steering by using the OpenFlow protocol to 'program' the fast, specialised application-specific integrated circuits (ASICs), which serve as the hardware-forwarding engines in network devices. BFO can be

There are tens of thousands of devices connected to the CERN network.

deployed to control a set of OpenFlow-enabled devices and thus enhance network flexibility and programmability. BFO features a user interface and a REST API, which can be used for consuming its services in an automated manner.

In 2016, significant progress was made on the IDS mirroring use case, which is illustrated below. The designed system receives traffic intercepted at the CERN firewall system and load-balances it across a pool of IDS servers, each running the open-source Bro Network Security Monitor system. When a security threat is detected, the malicious traffic is temporarily mirrored to a dedicated set of servers for storage and later analysis. The use of BFO is key to achieving the desired intelligence and automation in the system.

The prototype setup features a Brocade MLXe router, controlled by a BFO instance that has been extended with a specially developed Bro plugin. This plugin enables the IDS software to call the BFO REST API directly and thus achieve the desired automation. The prototype has now been deployed in the CERN Data Centre for testing.

The prototype setup for the IDS mirroring use case.

Data Storage Architectures

Every year, the four large-scale LHC experiments create tens of petabytes of data, which need to be reliably stored for analysis in the CERN Data Centre and many partner sites in the WLCG. As the user demands are increasing in terms of both data volume and aggregated speed of data access, CERN and its partner institutes are continuously investigating new technological solutions to provide their user communities with more scalable and efficient storage solutions.
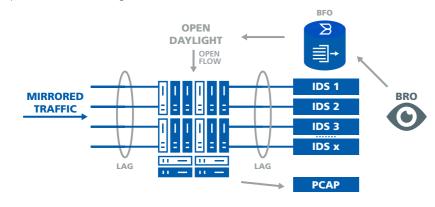
Oracle Database Technology and Monitoring

Throughout 2016, we continued to investigate Oracle Database In-Memory, with the key focus placed on performance. We evaluated the query response time under different database configurations for business-intelligence applications. In parallel, we tested the new 'Active Data Guard' features for Oracle Database In-Memory, and gave a number of presentations on this at the Oracle OpenWorld event, which took place in September in San Francisco, US.

In addition, we upgraded our Oracle Enterprise Manager monitoring systems — both those used for testing purposes and those used in production — to the latest version, 13c. This work included configuration and integration of CERN's active directory authentication system in the upgraded version of Oracle Enterprise Manager, and required configuration of custom certificates.

Finally, we took part in several beta programmes related to Oracle Database 12.2 in 2016. The tests were extensive, with special focus being placed on components for Oracle Clusterware and In-Memory. Throughout the year, we provided in-depth feedback to Oracle, as well as suggesting possible improvements to the current technologies.

Seagate Alternative Storage Architecture: 'Kinetic'

The collaboration with Seagate made considerable progress in 2016. Building on the software foundation laid in 2015 with the Kinetic I/O module (open-source library available on GitHub), the project was able to show first usage in a prototype service. Using the existing 1-PB Kinetic cluster (with first generation 'Lombard' drives), a storage pool was created for a related CERN openlab project with the European Bioinformatics Institute. This project focuses on the application of the ROOT analysis system to genomic data, with around 400 TB of genomic data having been served to a group of researchers throughout 2016.

The Seagate Kinetic project hosted a CERN openlab summer student for nine weeks in 2016. The student developed a configuration management and monitoring toolkit for the Kinetic drive cluster. A web interface was built using AngularJS, enabling efficient configuration and management of the Kinetic cluster and

Building on the software foundation laid in 2015 with the Kinetic I/O module, the Seagate project was able to show first usage in a prototype service in 2016.

The goal of the project with Comtrade is to simplify the usage, installation, and maintenance of the EOS large-scale storage system, as well as adding support for new platforms.

the monitoring of individual drive parameters. A REST API has also been added to all administrator and user commands provided by the EOS management server, so as to enable web front ends. The EOS scheduling software has been adapted to allow multi-path access to the Kinetic cluster via dynamically selected proxy nodes, thus providing high-availability. In addition, the I/O interface of the EOS storage server has been re-factored to better facilitate the requirements of having both locally attached and remote disks.

In parallel to the ongoing software developments, work was also carried out to prepare for the installation of a new cluster with second-generation Kinetic drives, which will complement the existing installation. This work also included pre-production testing and firmware optimisation — in particular for non-sequential access patterns, which are typical of some analysis uses cases at CERN. This work, which was carried out by a Seagate engineer, has resulted in a substantial increase in the performance of the firmware in the final product. We plan to demonstrate this in a production setup at CERN in 2017.

Comtrade Software EOS Productisation

The scope of this project — undertaken in collaboration with CERN openlab associate member Comtrade Software — is the evolution of the EOS large-scale storage system. The goal is to simplify the usage, installation, and maintenance of the system, as well as adding support for new platforms.

The main target of the project's initial phase was to provide a

robust installation kit to enable rapid installation of EOS both for evaluation purposes and for fast deployment in production. This has now been accomplished. The kit includes the necessary installation instructions and tools for operations (admin guide) and for users (user guide), as well as a first version of the EOS whitepaper.

In addition to its use at CERN, EOS is now in production at both Australia's Academic and Research Network (AARNet) and the European Commission's Joint Research Centre in Ispra, Italy. It is also being evaluated for use at the Institute of High Energy Physics in China; the National Autonomous University of Mexico; Moscow Engineering Physics Institute in Russia; Subatech in France; and Lawrence Berkeley National Laboratory, Oak Ridge National Laboratory, and Fermi National Accelerator Laboratory, all in the US.

The next phase of the project is focused on several items. The first goal is the integration of Comtrade Software engineers into the development and operations team at CERN to gain experience and autonomy in operating, maintaining, and developing EOS. The second goal is to continue the evolution of the installation framework and the documentation, in order to provide all EOS functions at installation time, including 'Sync&Share' capabilities, erasure-coding, and support for geographically distributed multi-site instances.

An additional milestone for this next phase is to develop a simultaneous testing framework to run after the build of every release. This will be used to certify each EOS version.
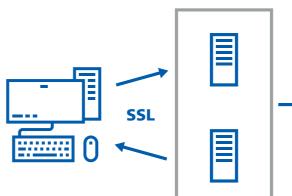
Compute Management and Provisioning

CERN, as infrastructure and service provider for the high-energy physics community, has been very actively involved in grid and cloud computing since the early days. As the use of virtualisation has become an increasingly viable solution for instantiating computing nodes, the concept of 'the cloud' or cloud computing has gradually established itself as an efficient and cost-effective solution for scientific computing.

Oracle Java EE

2015 saw significant work carried out to consolidate our Java EE platform-as-a-service, known as 'the Middleware on Demand' (MWOD). All sites were migrated from our previous platform, 'the J2EE Public Service'. During 2016, we focused on developing new functions for the MWOD, such as support for WebSockets and fine-grained authorisation for application paths.

Given that the developer community is demanding ever more agile and dynamic environments for developing and testing their applications, we also started to assess technologies such as Docker, HAproxy, and Kubernetes in 2016. We created a proof-of-concept platform based on these technologies for provisioning Apache Tomcat application servers.

We had much success in running our applications on Oracle WebLogic: we now have more than 250 clusters and a vibrant community of developers. We studied a number of container technologies in 2016, as part of our drive to continuously improve our platform. We created custom WebLogic-Docker images and integrated our CERN management and monitoring tools. Finally, we also began evaluation of several different aspects of the platform:
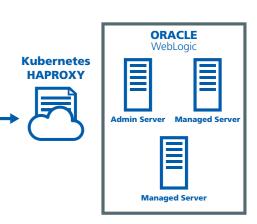• Load balancing: we tested different architectures based on Apache httpd, HAProxy, and Oracle Traffic Director.
• Security: we tested keystores, SSL configurations, and other setups.
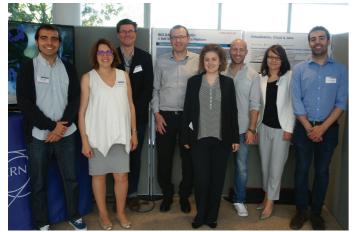• Clustering: we tested features including session replication and dynamic clusters.

Oracle Database Cloud

In order to find the best ways of overcoming the computing challenges the high-energy physics community is set to face in the coming years, many projects and experiments at CERN are looking at various cloud solutions from a range of vendors. Oracle's cloud portfolio has grown dramatically in recent years; the company now offers a wide range of products in this area.

WebLogic-Docker load-balancing architecture.

Apache Servers (Load Balancer)

Kubernetes Cluster

Kubernetes HAPROXY

SSL

ORACLE WebLogic

Admin Server   Managed Server

Managed Server

Results from each of the four projects with Oracle were presented at the CERN openlab open day.

Previous collaboration with Rackspace in CERN openlab's fifth phase has focused on cloud federation.

In late 2016, the Database Services Group in the CERN IT Department started some preliminary tests with Oracle Cloud, focusing primarily on aspects related to security (single sign-on integration) and backup management. In December, we started an evaluation of Oracle Cloud as a disaster-recovery solution, focusing on the feasibility of its use in the CERN environment, as well as the effort needed in case of future migration. We will test the complete stack of CERN applications from the storage layer up to application servers. It is in this context that we began evaluating services such as 'Database as a Service', 'Database Schema Service', and 'Database Backup'. This testing work will continue in 2017.

Rackspace Containers at Scale

Since its first releases, OpenStack has supported virtual machines to dynamically provision self-service resources for end users. New application frameworks such as Jupyter and Kubernetes rely on containers, a lightweight abstraction with lower overhead than virtual machines. OpenStack's support for containers comes through Magnum, a project within OpenStack. This is an area of major interest for industry and research, with over 78% of deployments looking into this function.

A CERN openlab fellow has been working with Rackspace and the open-source communities to enhance OpenStack's container support for use in high-energy physics. 72 patches have been submitted, supporting new storage technologies, documentation, local container repositories, and the usability of the command-line tools. These patches have been made available to the open-source community.

The service is now in production for the physicists and engineers at CERN. There are over 40 CERN projects underway, supporting new approaches to analysis and application architectures. An example of such a project is Jupyter SWAN, which is a web application that enables users to create and share documents containing live code, equations, visualisations, and explanatory text. With the dynamic provisioning of containers from OpenStack (and integration of ROOT), these notebooks can make it possible for researchers to interactively collaborate, publish, and preserve their results.

With container usage growing rapidly, it is important to be sure that the service can scale to meet the needs of the experiments at CERN. A 1000-node scale test was able to handle over 7 million requests/second. The detailed results from this test were presented at the OpenStack Summit in Barcelona in October. A number of areas for potential improvement have subsequently been planned with the community.
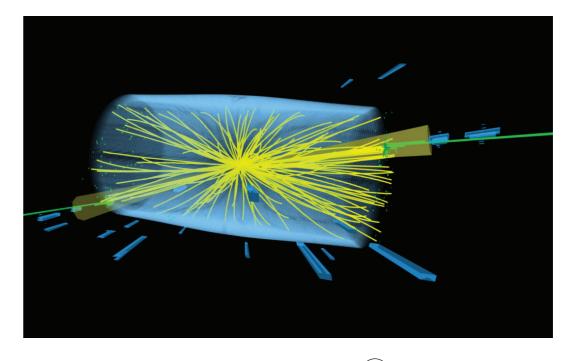
Computing Platforms (Offline)

The success of existing and future scientific and experimental programmes depends — among other factors — on efficient exploitation of the recent and future advances in computing technology. Existing software needs to be revised, optimised, or completely redesigned to fully exploit the performance gains provided by newer multi-core platforms, fast co-processors, and graphical processors.

Intel Code Modernisation

Across research fields, code optimisation is of paramount importance in ensuring that available hardware is used as efficiently as possible. The increased computing requirements of the LHC Run 2 mean it is more important than ever to optimise high-energy physics codes for new computing architectures. As part of this project, Intel experts have delivered a number of workshops at CERN addressing the latest Intel software tools and providing training on code-vectorisation technologies.

In addition to the specific projects described below, optimisation work is to be carried out on several injector simulation codes used by the CERN Beams Department. An agreement has been signed with Italy's National Institute for Nuclear Physics (INFN) that will see work undertaken in this area in 2017.

The GeantV work aims to develop the next-generation simulation software used for describing the passage of particles through matter. It started as a research-and-development effort aiming to develop an alternative path to detector simulation software, using a multi-event approach and multi-particle 'basket' parallel processing to achieve higher efficiency with simulation-specific computations on modern hardware.

This was achieved by expressing algorithms in a type-agnostic manner using either scalar or vector interfaces. In turn, this was enabled by the implementation of a new library called 'VecCore', using dedicated backends to support different SIMD architectures. It demonstrated excellent results on both Intel Xeon and Xeon Phi (Knights Corner). A new AVX-512-aware backend called 'UME::SIMD' enabled GeantV to be deployed 'out of the box' and to demonstrate SIMD gains on the Intel Knights Landing architecture.

One of the key focus areas of the GeantV development in 2016 was the adaptation of the core architecture to offer improved scalability on many-core platforms. We introduced and tested a multi-propagator model where the work of transporting particle baskets in the detector is split among several managers, each taking charge of a limited number of threads and having weak communication with one another for workload balancing. This approach has shown good scalability on the Intel Knights Landing

A collision event with two high-energy jets recorded in the CMS detector.

architecture, while preserving vector gains. It will be the basis of the third version of the GeantV core architecture, which is scheduled for an alpha release around the fourth quarter of 2017.

A number of high-performance components developed for GeantV have already been made available to the HEP community. The aim is to integrate GeantV with experiment frameworks by the end of 2018.

FairRoot/ALFA framework

FairRoot is a framework for simulation, reconstruction, and data analysis for particle and nuclear physics experiments. It is based on the data-processing framework ROOT, which was born at CERN. FairRoot is being developed at the GSI Helmholtzzentrum für Schwerionenforschung for the future experiments at the FAIR facility. In FairRoot, the reconstruction algorithms can be implemented as 'tasks' in ROOT, or as separate processes that communicate with each other via message passing. The latter method has been under development for a few years and enables easy integration of co-processors and/or different hardware architectures.

The commonalities between the FAIR experiment and the ALICE experiment at CERN — particularly in terms of their computing requirements — have led to the development of a common software framework called ALFA. ALFA is a concurrency framework for high-quality parallel data processing and reconstruction on heterogeneous computing systems. It was jointly developed by the ALICE online-offline and FairRoot teams. ALFA provides a data-transport layer and the capability to coordinate multiple data-processing components. It is a flexible, elastic system. It balances performance with reliability and ease of development by using multi-processing in addition to multi-threading. The Knights Corner generation of Intel Xeon Phi processors was integrated into ALFA's dynamic deployment system (DDS) in 2016. The DDS is a toolset that automates and simplifies the deployment of user-defined sets of processes and their dependencies on any available resources.

At the end of 2016, work to get the code running fully on the Knights Landing generation of Intel Xeon Phi processors began. Tests will be carried out in 2017 to compare the performance of the processing and communication.

BioDynaMo

The BioDynaMo project, which engages with the life sciences community, is a collaboration between CERN, Newcastle University, Innopolis University, Kazan Federal University, and Intel. Its goal is to design and build a scalable and flexible cloud-based computing platform for rapid simulation of biological tissue development. It foresees three main phases: the consolidation, optimisation, and further extension of biological simulation code to run efficiently on modern multi-core and many-core platforms; the deployment of a cloud-based platform using state-of-the-art cloud-based high-performance computing technologies; the creation of a shared ecosystem of tools, datasets, processes, and human networking in the field of biological simulation.

The project focuses initially on the area of brain tissue simulation, drawing inspiration from existing, but low-performance software frameworks. Late 2015 and early 2016 saw algorithms already written in Java code ported to C++. Once porting was completed, work was carried out to optimise the code for modern computer processors and co-processors, so as to make the best possible use of the many available cores. The optimisations will be tested over the first months of 2017 and support for additional cell types and behaviours will be added. The next step will then be to extend the system to run in a cloud-computing environment, thus making it possible to harness many thousands of computer processors to simulate very large biological structures.

Three project workshops were held in 2016: in Newcastle, Kazan, and at CERN.

Members of the BioDynaMo project came to CERN for a technical workshop on 1-2 December.

GeneROOT

Like BioDynaMo, the GeneROOT project is part of CERN openlab's ongoing investigations around transferring computing knowledge and technologies to other research communities. The aim of the project is to use ROOT — a data-processing framework created at CERN and used widely in the high-energy physics community — to analyse large genomics datasets and to share methods and results across a distributed community of scientists.

In collaboration with project partner King's College London, GeneROOT is initially making use of sequences from the TwinsUK registry. However, the project has the goal of extending to other similar datasets hosted by facilities across the globe.

The project will investigate whether the ROOT system can be successfully used to extend the capabilities of existing analysis and file-access tools. Increased efficiency is required to handle the ever increasing amounts of data produced by modern sequencing techniques. The field of genomics has very well established working practices, meaning that any new system will have to be highly performant in order to ensure widespread community adoption.

The project kicked off in May 2016. The 300TB genomics dataset has been transferred to the CERN EOS storage system to ensure data redundancy and to provide local access to CERN experts.

The GeneROOT project aims to use the ROOT data-processing framework to analyse large genomics datasets.

Initial genome sequence alignment jobs have been run to improve understanding of the computing requirements. The project's next step is to integrate the standard data format used in genomics into ROOT, so that analysis benchmarking can take place. Work will then start on adapting the ROOT analysis tools to the genomics community's needs and ways of working.

Cisco Data Plane Computing System (DPCS)

The aim of the Cisco Data Plane Computing System (DPCS) project, which came to a close in June 2016, was to enhance the performance of distributed applications through the elimination of kernel processing in the data path. This was achieved using the ultra-low latency Ethernet cards from Cisco (usNIC) that enable direct communication between the user-space application and the network interface controller (NIC), bypassing the kernel transmission control protocol (TCP) stack and effectively offloading this work from the CPU.

The main outcome of the project is the implementation of a new transport mechanism for an open-source NanoMsg library that uses a libfabric abstraction and a Cisco driver for the usNIC fabric. This makes it possible for the ALICE experiment to use its unmodified software and benefit from low-latency access to network hardware, bypassing the kernel software stack.

As part of the project, we worked to develop a new transport mechanism that would enable the current ALICE software to use the usNIC fabric. The NanoMsg library was chosen, and extended with features enabling the creation and management of zero-copy messages. The team also submitted patches to the maintainers of the library.
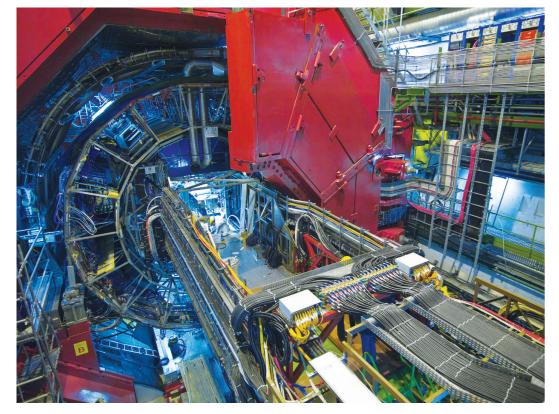
The chosen solution for the network transport layer was based on the libfabric library, developed by the Open Fabric Alliance and for which Cisco provides and maintains a low-level driver. During the project's development phase, a series of issues were identified and reported back to the development team, along with proposed modifications and extensions.

Once development was completed, the prototype system for improved data transport was tested using standard ALICE benchmarks, running on hardware provided by Cisco with UCSVIC1285 cards connected with two 40-gigabit links. In order to evaluate the network performance improvements, a benchmark automation tool (roBob) that simplifies both the collection of measurements for repetitive tasks and the automatic creation of reports was developed and made publicly available on GitHub.
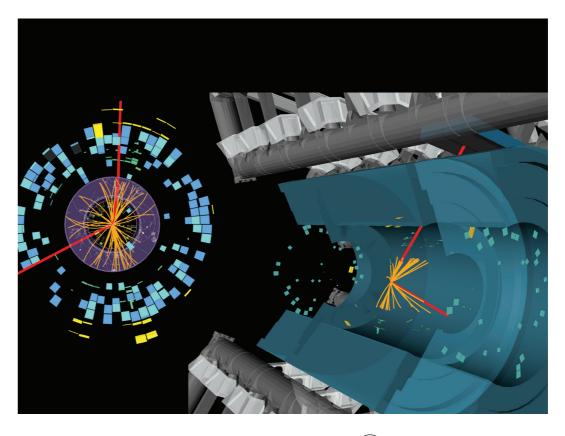

ALICE detects quark-gluon plasma, a state of matter thought to have formed just after the big bang.

ARM Porting and Benchmark Studies

The LHC experiments and the CERN computing and data infrastructure make use of a large number of software frameworks for simulation and reconstruction of collision events. It is important to continuously monitor advances and trends in technology and evaluate software on different computing platforms as they evolve. A new project was therefore launched midway through 2016 to port several widely used codes for running on ARMv8-A 64-bit architecture. As part of the project, studies are being carried out to test and measure performance, energy consumption, and other operational aspects, so as to understand the strengths and weaknesses of the architecture.

A cluster comprised of ARMv8-A 64-bit evaluation prototype servers was delivered in June. Each server contains 32 Cortex-A57 cores with 128 GB RAM connected in four fast memory channels. The project started with the successful porting of two key software packages used at CERN, Geant4 and ROOT. These programs are each of the order of around three million lines of C++ code and are both fundamental to the research carried out at CERN. Following this porting work, the project's efforts focused on use cases for the two large multi-purpose experiments on the LHC: ATLAS and CMS.

Members of the ATLAS collaboration first worked to port their software framework and validate the output from the ARMv8-A 64-bit servers against current technology used. Once the output was found to be in agreement, the team then worked to benchmark the ARMv8-A 64-bit servers. In terms of energy efficiency, the cluster was found to deliver a maximum of approximately 1150 event simulations per kilowatt hour.

Members of the CMS collaboration carried out similar benchmarking work using tools pertinent to their experiment. They found that the new ARMv8-A 64-bit cluster demonstrates up to roughly five times increased performance over comparable clusters consisting of previous generations of ARM platforms, which they have been experimenting with since 2013.

Updated ARM-based clusters are set to be made available to the experiments as part of the next phase of the project, so that further porting and benchmarking work can be carried out.

An event simulated on the ARM prototype server. This is in alignment with output from servers currently used in production.

Data Analytics

During the past decades, CERN and other international research laboratories have been gathering not only enormous amounts of scientific data, but also very large quantities of systems-monitoring data from their instruments. Curating, enriching, and managing this data enables its exploitation. The main challenges in data analytics for scientific and engineering applications involve technology, integration, and education.

Oracle Analytics-as-a-Service

During the first half of 2016, the team carried out a detailed evaluation of Oracle Big Data Discovery. This involved analysis of terabytes of technical engineering data produced by approximately 50,000 sensors and other monitoring devices in CERN's accelerator complex. In the second half of the year, we focused our efforts in two main areas: (i) the integration of Oracle Big Data Discovery into our production environment; (ii) establishing an architecture for reliability and availability analysis of the systems within the proposed Future Circular Collider (FCC). This infrastructure has

Oracle Big Data Discovery was used for fault-analysis studies relating to the proposed Future Circular Collider.

enabled the collaboration working on the FCC studies to perform a variety of analyses for accelerator conditions and modes, such as cooling down, warming up, and injecting beams of energy in a scalable manner. We also implemented a 'proof of concept' using Oracle R Advanced Analytics for Hadoop (ORAAH) for the reliability and availability studies carried out for the FCC. ORAAH is a framework that provides an R-language interface to transparently manipulate data in Hadoop and run high-performance machine-learning algorithms with Spark. We modelled some of the analyses carried out as part of the cryogenic valve degradation and fault-detection studies. Subsequently, ORAAH runs them in Hadoop for hundreds of valves in parallel. This work enabled us to easily develop and adapt models in R to run faster and process more data than traditional parallel computing solutions in R, thanks to the massive parallel capabilities of Hadoop and Spark.

During the final quarter of 2016, the team got involved in the specification, deployment, and validation phases for a scalable streaming analytics platform based on Apache Kafka. We started to study the integration of Kafka with Oracle Stream Explorer Platform, in order to simplify the near-real-time data analysis and identification of complex event patterns.

In terms of outreach, we shared information about our endeavours and experiences related to big data analytics with a range of organisations. We also participated in a number of important conferences, including the Oracle Big Data Analytics Innovation Summits in Switzerland, Norway, and Sweden.
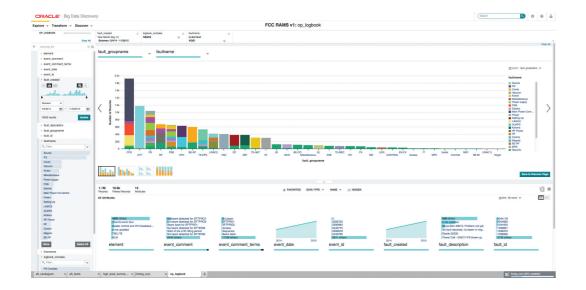

Siemens Industrial Control and Monitoring

Most of the CERN installations in the experiments, accelerators, and other technical infrastructures rely on a multitude of heterogeneous industrial control systems for proper functioning. These control systems produce enormous amounts of data related to both the systems they control and their own internal state. Together with our partners Siemens, we are seeking to apply big-data analytics techniques to this data, in order to improve the operational behaviour of the entire system. Our goal is to increase efficiency and develop new control models that improve reliability. In particular, specific algorithms have been designed and implemented for use-cases related to each of the following areas:
• Detection of faulty sensor measurements for cryogenics systems
• Performance measurement of process control systems using proportional–integral–derivative (PID) controllers
• Detection of alarm flooding and identification of responsible control devices
• Development of a recommendation system for users of WinCC OA (a SCADA tool widely used at CERN).

As well as important new use cases being identified in 2016, methods developed through this project were integrated into operational tools used in various control rooms at CERN.

In 2016, work began on a tool to provide operators and domain experts with information about the health of their processes — as determined from data-analytics algorithms. The goal is to display

the results of several types of analysis within WinCC OA operator panels. This entails location, extraction, and display of data-analysis results that have been created and stored within a Hadoop system. This display must be integrated with existing process 'faceplates'. A demonstrator system was developed with a focus on displaying the results of oscillation detection — this is particularly valuable for control valves, where oscillation can be an indicator of problems with the valve. The four use cases listed above will gradually be added to the tool.

Next Generation Archiver

Over the coming years, planned upgrades to CERN's accelerator complex and the experiments' detectors will lead to increased and more complex requirements in terms of storing monitoring data. Thus, in the context of our work related to WinCC OA, significant progress was made in 2016 towards the development of a next-generation archiver.

The work in this area is geared towards providing a modular, plugin-based archiver, wherein various storage backends can be implemented as necessary and used interchangeably. Of particular interest for CERN is a high-performance backend that enables connection to CERN's Hadoop-based data-analytics infrastructure. After a series of exploratory exercises using different technologies, a first proof-of-concept was implemented and its performance was checked to prove the feasibility of the modular approach. The project team aims to produce the first functioning prototype, with two different backends, in early-to-mid 2017.

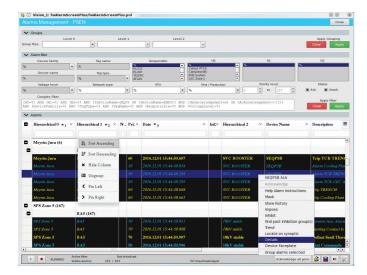Members of the team collaborating with Siemens pictured in the CERN Control Centre.

Advanced Alarm Screen

The complexity of CERN's industrial control systems demands advanced tools for the treatment of alarms. During 2016, work began to develop an improved alarm screen for WinCC OA. This will enable complex filtering, grouping of alarms by a range of criteria, and many possibilities for user customisation.

Two possibilities were evaluated for the design and implementation of this system: one based on the Qt C++ toolkit (QTitan DataGrid library) and the other based on HTML/JavaScript technology. For the HTML/JavaScript option, many different libraries were considered. It was finally decided to go for the HTML/JavaScript version based on Angular.js as it offered better flexibility and a more appropriate licensing model. A first prototype was developed and a performance evaluation study was carried out.

The new functions offered result in a performance penalty of around 15% on the client side. Work is now focused on minimising this client-side impact, as well as generalising the system to all applications. We will then deploy it in selected production applications.

Advanced alarm screen for WinCC OA.

Yandex Data Popularity at LHCb

Data collected by the LHCb experiment is stored across multiple datasets on both disks and tapes in the LHCb data storage grid. The storage systems used vary in terms of cost, energy consumption, and speed of use. The goal of this project is to design, develop, and deploy a 'data popularity estimator service' that would analyse the usage history of each dataset, predict future usage patterns, and provide an optimal scheme for data placement and movement.

During 2015, the project team developed an algorithm that is able to reduce the total amount of disk space needed by 40%. In 2016, work focused on determining the optimal number of replicas for each file. This investigation was divided into two main parts.

Firstly, a classifier based on decision trees was used to predict the likelihood of datasets being used in the coming six months. For this purpose, the access histories of the datasets — and their metadata for the last two and a half years — were fed into the system. Each dataset is described in terms of the following features: how recently it was last accessed, time elapsed between consecutive accesses, time of first access, creation time, access frequency, type, and size. Comparison with the 'Least Recently Used' (LRU) algorithm shows that our approach significantly decreases the number of files that are wrongly removed.
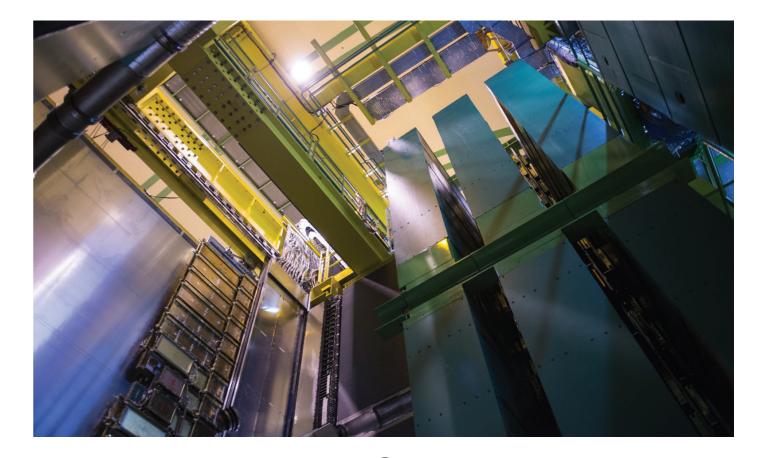
Secondly, Brown's model for exponential smoothing was used to predict the number of times each dataset would be accessed over the next month. This was then used to prioritise the datasets in terms of the number of replicas they each need. Thus, the service is able to make appropriate decisions about the cases in which replica datasets can and cannot be removed, in order to maximise available disk space — while keeping the risk of data loss to a minimum.

Yandex Anomaly Detection in LHCb Online Data Processing

Ensuring data quality is essential for the LHCb experiment. Checks are done in several steps, both offline and online. Monitoring is based on continuous comparison of histograms with references, which have to be regularly updated by experts.

In 2016, work continued to create a novel, autonomous monitoring service for data collection. The service is capable of identifying deviations from normal operational mode to help personnel responsible for monitoring data quality to find the reason for such deviations. It will, therefore, increase the efficiency of the system by reducing the amount of spoiled data that is erroneously stored for further analysis.

An automated data-certification assistant, called 'Roboshifter' was developed in 2016. A data pipeline was set up for feeding LHCb run history into the Roboshifter machine-learning algorithm, which predicts whether given data is good or bad. Roboshifter was also integrated into the web service for data-quality monitoring that is used by personnel to help them find the detector components responsible for any problems with the data.

The aim of the LHCb experiment is to record the decay of particles containing b and anti-b quarks, collectively known as 'B mesons'.

Education

A knowledge factory

CERN openlab is designed to create and disseminate knowledge.

CERN openlab conducts education and training at a number of different levels, with our experts contributing to a broad range of activities.

Workshops and seminars are regularly organised at CERN on advanced topics directly connected to the CERN openlab projects. These feature a mix of lecturers from both industry and research, and are an excellent example of two-way knowledge transfer in action. The workshops, which are listed on page 35, often combine theory with hands-on practice. Special courses are also organised for advanced CERN users in specific technical areas.

The students participating in this year's programme came from 21 different countries.

Summer success

The CERN openlab summer student programme is a major element of our educational work. The programme continues to go from strength to strength, with 39 participants — representing 21 different nationalities — selected from around 1500 applicants in 2016. This educational programme was launched in 2002 to enable undergraduate and masters students to gain hands-on experience with advanced ICT solutions and to learn more about the challenges being addressed through CERN openlab.

The 2016 students worked on ambitious projects using some of the latest hardware and software technologies, and were able to gain first-hand knowledge of the integral role cutting-edge ICT solutions play in supporting high-energy physics research. In addition, the students attended a series of lectures developed for the CERN openlab summer students, given by ICT experts on advanced CERN-related topics. Visits to various CERN facilities and experiments were also included in the programme, as well as to other research institutions and companies.

For the first time, CERN openlab played a major role in supporting and organising the CERN Student Webfest. The event is a weekend of online web-based creativity, modelled on the 'hackathons' that energise many open-source communities. It is a grassroots initiative that aims to spark new ideas that could innovate the future of web-based education about CERN, the LHC, particle physics, humanitarian matters, and health. Over 50 participants worked in 13 teams across the weekend (29 to 31 July).

Intel competition winners come to CERN

From 11 to 17 June, CERN hosted the 10 young students who won the CERN Special Award at the Intel International Science and Engineering Fair (ISEF) 2016. These winners were selected from the 1700 high-school students who participated in the competition. The competition, which is a programme of the Society for Science and the Public, is the world's largest pre-university science competition. It offers high school students from across the globe a chance to showcase their research into a range of fields. The CERN Special Award — funded jointly by Intel and the CERN IT Department — is now in its eighth year.

The 10 winners of the CERN Special Award at the Intel ISEF 2016.

Two of the first-place winners from the Intel Modern Code Developer Challenge, pictured on their visit to CERN in September.

During their time at CERN, the students presented five-minute 'lightning talks' on their projects. These focused on areas from IT to materials science and from robotics to astrophysics. The winning students — aged between 15 and 18 — spoke to researchers at the laboratory and had the opportunity to see exciting technology up close. Sites visited include the AMS experiment, the Data Centre, the LEIR accelerator, the Synchrocyclotron, the CERN Control Centre, and the Antiproton Decelerator.

The winners of the Intel Modern Code Developer Challenge also came to CERN in 2016. Two of the top winners participated in the CERN openlab summer student programme, while two others came to the laboratory in September for an extended tour.

As part of the competition, students were tasked with optimising code that is integral to the BioDynaMo project. The goal of the initiative was to spur advancement in parallel coding and the science it supports, as well as encouraging students to pursue careers in the field of high-performance computing. The challenge attracted over 17 000 students from around the world, representing over 130 universities across 19 countries. Over 1200 students downloaded the code and over 1000 participated in the free online training programme, choosing from over 20 hours of educational materials.

ICE-DIP concludes

ICE-DIP, the Intel-CERN European Doctorate Industrial Programme came to a conclusion at the end of 2016. This Marie Curie Actions project, within the European Union's 7th Framework Programme,

built on CERN's long-standing relationship with Intel through CERN openlab. The focus of the project, which was launched in 2013, was the development of next-generation techniques for acquiring and processing data that are relevant for the trigger and data-acquisition systems of the LHC experiments.

The project brought together CERN, Intel, and universities to train five early-career researchers. These students were funded by the European Commission and were granted CERN fellowships while enrolled in doctoral programmes at the partner universities, Dublin City University and the National University of Ireland Maynooth. They each completed 18-month secondments at Intel locations around the world, gaining in-depth experience of the very latest generations of Intel hardware.

Through this public-private partnership, the researchers have been preparing the techniques necessary for acquiring and processing hundreds of terabits per second using and expanding the most innovative concepts available in the ICT industry today. They have been exploring new areas, including silicon photonics for network links in harsh operational conditions, and tight integration of reconfigurable logic with commodity processors to bring new approaches to data acquisition — all with increased performance and decreased cost in mind.

Specifically, the researchers have: developed a power-efficient, low-cost, and high-bandwidth data link based on silicon-photonics and capable of operating in harsh environments; designed a high-performance data pre-processing system that integrates

commodity CPUs and FPGAs; assembled a cost-effective, high-bandwidth data-acquisition network capable of multi-terabit lossless throughput using commodity components; designed and prototyped a massively scalable data-filtering system that fully exploits the parallelism of many-core processor architectures, and which is programmable using industry-standard x86 models. The findings open up new opportunities in the domain of high-performance data acquisition, going towards and beyond technological novelties such as silicon photonics, the Intel Xeon Phi co-processor, and the ever-growing family of Intel Xeon processors.

A workshop was held in September to mark the project coming to a close. The event was an opportunity to share the valuable knowledge gained. The developments made by the ICE-DIP researchers are of great interest for CERN's future computing facility upgrades, other research laboratories and, potentially, other business sectors too.

Dissemination

CERN openlab results were disseminated at a wide range of international conferences. These posters, presentations, and repports can be consulted on the CERN openlab website. The full list of presentations, publications and posters for 2016 is available on pages 36 to 39 of this annual report.

CERN openlab Topical Workshops

- **CERN openlab/Intel Hands-on Workshop on Code Optimization,** 05-06 April 2016, CERN, L. Duhem/Intel K.D. Oertel/Intel, G. Zitzlsberger/Intel Ulm, J. Arnold/Intel
- **CERN openlab Mini-Workshop on Floating-Point Computation,** 07 April, CERN, J. Arnold/Intel
- **Containers and Orchestration in the CERN Cloud (Seminar),** 08 April 2016, CERN, R. B. Da Rocha/CERN
- **CERN openlab Machine Learning and Data Analytics workshop,** 29 April 2016, CERN
- **ICE-DIP Closing Workshop,** 13-14 September 2016, CERN, B. Jones/CERN, A. Nowak/Tlk Services, D. Malone/Maynooth University, G. Jereczek/CERN, M. Zeiler/CERN, S. Sridharan/CERN, A. Santogidis/CERN, P. Karpinski/CERN, L. Barry/Dublin City University, A. Winstanley/Maynooth University, J. Butler/Intel, M. Nordberg/CERN
- **2nd CERN openlab/Intel Hands-on Workshop on Code Optimisation,** 24-25 November 2016, CERN, K. D. Oertel/Intel, G. Zitzlsberger/Intel, P. Karpinski/CERN
- **BioDynamo Workshop,** 01-02 December 2016, CERN, F. Rademakers/CERN, M. Manca/CERN, R. Bauer/ Newcastle University, L. Breitwieser/CERN, V. Drobny/Innopolis University, I. Dmitrenok/Innopolis University, A. Trusov/Kazan Federal University, L. H. Nielsen/CERN, A. Kosterin/Innopolis University, A. Di Meglio/CERN
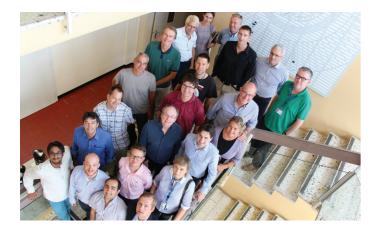
- **CERN openlab Technical Workshop,** 08-09 December 2016, CERN, M. Girone/CERN, S. Baymani/CERN, O. Awile/CERN, N.Neufeld/CERN, E. Meschi/CERN, G. Aielli/ University and INFN of Roma Tor Vergata, A. L. Krajewski/ CERN, E. Martelli/CERN, S. N. Stancu/CERN, S. Trigazis/ CERN, T. Bell/CERN, C. Green/ Seagate, X. E. Curull/CERN, A. J. Peters/CERN, A. Nappi/CERN, L. D. Lupsa/CERN, K. M. Dziedziniewicz-Wojcik/CERN, A. Purcell/CERN, F. Carminati/ CERN, L. Atzori/CERN, J. W. Smith/ Georg-August-Universitaet Goettingen, A. Salzburger/CERN, M. Schulz/CERN, A. H. De Campos/ Universidad Nacional de Educacion a Distancia UNED, F. M. Tilaro/CERN, J. Guzik/CERN, P. J. Seweryn/ CERN, F. Ratnikov/ Yandex School of Data Analysis, E. Tejedor Saavedra/CERN, Oliver Gutsche/Fermi National Accelerator Lab, L. Canali/CERN, D. Duellmann/CERN, M. Pierini/CERN, M. Manca/CERN, L. Breitwieser/CERN

CERN openlab Summer Student Programme Teaching Series, July-September 2016:

- Computing in High Energy Physics, H. Meinhard
- DAQ - Filtering Data from 1 PB/s to 600 MB/s, N. Neufeld
- Finding the Higgs Boson : the story from the statistical methods perspective, W. Verkerke
- Finding the Higgs Boson : the story from the software and computing perspective, K. Bloom
- Computing Security, S. Lopienski
- How to give presentations and pitches, M. Cirilli, M. Dehli
- Computing from Grids to Clouds-Part 1, I. Fisk
- Computing from Grids to Clouds-Part 2, I. Fisk
- Machine Learning - Part 1, S. Gleyzer
- Machine Learning - Part 2, S. Gleyzer
- Best practices: the theoretical and practical underpinnings of writing code that's less bad, A. Naumann
- Volunteer computing, C. A. Bourdarios
- Introduction to code optimizations + practical examples, S. Vallecorsa
- Intel challenge experience, M. Gravey
- Evolution in Computing Hardware - Part 1, S. Jarp
- Evolution in Computing Hardware - Part 2, A. Nowak

CERN openlab Summer Students 2016, with Nation State, Home Institute and Project Topic:

- A. Aghabayli, Azerbaijan, Baku State University, Azerbaijan, "Data Quality Monitoring at CMS with Machine Learning"
- M. Arora, India, Indraprastha Institute of information Technology, India, "A Configuration Management and Monitoring Toolkit for EOS and Kinetic Drive Cluster"
- M. Bandi, Bosnia and Herzegovina, International Burch University, Bosnia and Herzegovina, "Hue Application for Big Data Ingestion"
- J. Banjac, Serbia, Faculty of Technical Sciences University of Novi Sad, Serbia, "FPGA Based Data Smoother for Sensor Data"
- S.S. Baveja, India, Netaji Subhas Institute of Technology, India, "Test of Oracle JSON support in the view of CMS JSON data"
- R. Bisht, India, Vellore Institute of Technology, India, "Oracle R technologies for data analytics and machine learning in hybrid data systems"
- A.C. Chelba, Romania, National Institute of Applied Sciences of Lyon, France, "Consolidation and Performance measurements of ROOT Multiproc Core"
- D. Christidis, Greece, University of Patras, Greece, "Automating WLCG Job Accounting Validation"
- M. Dhar, India, International Institute of Information Technology – Hyderabad, India, "Gotham Remote Logins

The ICE-DIP project built on CERN's long-standing relationship with Intel through CERN openlab.

Monitoring System"

- D. Ernst, Germany, Friedrich-Alexander University Erlangen-Nürnberg, Germany, "Performance studies on different accelerators using OpenCL"
- S.N. Fatima, Pakistan, National University of Sciences and Technology, Pakistan, "High Speed DAQ with DPDK"
- M. Gravey, France, University of Lausanne, Switzerland, "Implement a MPI-based framework to steer NUMA-aware GeantV workload for many-core systems"
- R. Gupta, India, University of Petroleum and Energy Studies, India, "Modernizing the Monitoring of Mass Storage systems"
- S. Gupta, India, Maharaja Agrasen Institute of Technology, India, "Database Consistency Verification for CMS metadata"
- N. Hardi, Serbia, University of Novi Sad, Serbia, "A graphical Linux-perf based tool that reports application performance information"
- A. Hesam, Afghanistan, Delft University of Technology, the Netherlands, "Integrating ROOT I/O in BioDynaMo Brain Simulation Platform"
- S.M.R. Jaffery, Pakistan, National University of Science and Technology, Pakistan, "Improvement of a Memory Profiling Tool, MALT"
- V. Janevski, Former Yugoslav Republic of Macedonia, Saints Cyril and Methodius University of Skopje, Former Yugoslav Republic of Macedonia, "Shadowserver reports automated tool"
- E. Kleszcz, Poland, Warsaw University of Technology, Poland, "OAuth protocol for CERN Web Applications"
- E. Kok, the Netherlands, Radboud University Nijmegen, the Netherlands, "Root based genomic project"
- J. Mahapatra, India, Veer Surendra Sai University of Technology, India, "Deep Learning for Imaging Calorimetry"
- N. Medjkoune, Algeria, University of Technology of Compiègne, France, "Integrating Docker containers into the CERN batch system"
- C.G. Moraru, Romania, University Politehnica of Bucarest, Romania, "MemProf - Memory Allocation Profiling Tool for Real-World Applications"
- K. Napoli, Malta, University of Malta, Malta, "Microservices Scheduling for ALICE O2 Facility"
- I. Nikoli, Croatia, University of Zagreb, Croatia, "Software Defined Networking Topology Service"
- I. Nurgaliev, Russia, Innopolis University, Russia, "Kibana, Grafana and Zeppelin on Monitoring data"
- D. Padmadas, India, Amrita University, India, "RapidIO and

Multicast: Investigations using a real-time chat"
- D.I. Panova, Bulgaria, Barcelona Graduate School of Econimics, Spain, "Evaluation of Intuitive Platforms for Data analytics"
- D. Petruk, Ukraine, AGH University of Science and Technology, Poland, "Puppet librarian and Git"
- F. Pressuti, Italy, California Institute of Technology, United States of America, "Distributed LHC Event – Topology Classification"
- V.K. Rathi, India, Jawaharlal Nehru University, India, "Container Networking"
- P. Ribes Metidieri, Spain, University of Barcelona, Spain, "Integrating Htopml in DAQPIPE"
- K. Šatara, Bosnia and Herzegovina, University of Novi Sad, Serbia, "Building custom plugin for Kibana to visualise Oracle database audit logs"
- D. Saxena, India, University of Delhi, India, "Resource Visualization"
- F.C. Schiavi, Italy, Polytechnic University of Milan, Italy, "Online syntax highlighting and checker for Control Unit Type of Finite State Machine"
- T. Shaffer, United States of America, University of Notre Dame, United States of America, "HEP Application Delivery on HPC Resources"
- M. Sharma, India, National Institute of Technology Delhi, India, "Explorer of Grid Load"
- I.M.S. Sutandi, Indonesia, University of Indonesia, Indonesia, "Oracle Traffic Director for High Availability Architectures in the CERN Web Application Ecosystem"
- E. Tapta, Greece, Athens University of Economics and Business, Greece, "Documentation System for WinCC OA applications"

CERN openlab Presentations and Publications

Presentations

- F. Rademakers/CERN, New technologies for HEP- The CERN openlab, ACAT 2016, Valparaiso, Chile, 18 January 2016
- J. Castro Leon/CERN, Identity Service @ CERN Cloud, Workshop on Federated Identity for Cloud Services, Zurich, Switzerland, 21-22 January 2016
- M. Girone/CERN, A successful public-private partnership, Presentation of CERN openlab at Oracle visit to CERN, CERN, Geneva, Switzerland, 27 January 2016.
- A.R. Marin/CERN, Oracle Big Data Discovery for CERN's Control Data, BIWA 2016, Oracle Conference Center at Oracle HQ Campus, Redwood Shores, CA (USA), 27 January 2016
- A.R. Marin/CERN, M.M. Marquez/CERN, Fault Detection using Advanced Analytics at CERN's Large Hadron Collider, BIWA 2016, Oracle Conference Center at Oracle HQ Campus, Redwood Shores, CA (USA), 28 January 2016
- T. Bell/CERN, CERN and Clouds, CCT-SIL-20160204 Data Production with the cloud, 04 February 2016
- A. Di Meglio/CERN, Big data for big discoveries: How LHC finds needles by burning haystacks, Presentation of CERN openlab at the High Performance Computing & Big Data 2016 Conference, Victoria Park Plaza, London, UK, 04 February 2016.
- A. Santogidis/CERN, Extending ZeroMQ and or NanoMSG with support for SCIF, ALICE CWG13 Meeting, CERN, Geneva, Switzerland, 26 February 2016
- W. Magro/Intel, Intel in HPC: Innovation Beyond the Core, Intel in HPC: Innovation Beyond the Core, CERN, Geneva, Switzerland, 29 February 2016
- A. Santogidis/CERN, Shared memory and Message passing revisited in the many-core era, inverted CERN School of Computing 2016, 02 March 2016

- L.C. Noll/Cisco, Computational Challenges of finding the largest prime, Computational Challenges of finding the largest prime, CERN, Geneva, Switzerland, 9 March 2016
- A.R. Marin/CERN, Oracle Big Data Discovery for CERN's Control Data, Oracle Business Analytics Big Data und Data Warehouse Konferenz 2016, Mainz, Germany, 16 March 2016
- P.H. Lensing/Seagate, Kinetic Drive Integration in EOS, Kinetic Drive Integration in EOS, CERN, Geneva, Switzerland, 18 March 2016
- T. Bell/CERN, CERN Computing Infrastructure, Soleil Synchrotron Visit, CERN, Geneva, Switzerland, 22 March 2016
- M. Girone/CERN, Big Data Analytics and the LHC, Big Data Analytics Europe (bda) 2016, Amsterdam, Netherlands, 22 March 2016
- I. Charalampidis/CERN, New libfabric based transport for nanomsg, ALICE offline week, CERN, Geneva, Switzerland, 1 April 2016
- L. Duhem/Intel, K.D. Oertel/Intel Munich, G. Zitzlsberger/Intel, J. Arnold/Intel, CERN openlab/Intel hands on workshop on code optimization, CERN openlab/Intel hands on workshop on code optimization, CERN, Geneva, Switzerland, 5-6 April 2016
- J. Arnold/Intel, An Introduction to Floating-Point Arithmetic and Computation, CERN openlab Mini-Workshop on Floating-Point Computation, CERN, Geneva, Switzerland, 7 April 2016
- J. Arnold/Intel, Floating Point Arithmetic and Computation: Instructions for Exercises, CERN openlab Mini-Workshop on Floating-Point Computation, CERN, Geneva, Switzerland, 7 April 2016
- R. B. Da Rocha/CERN, M. Velten/CERN, B. Noel/Ministere des Affaires Etrangeres et Europeennes (FR), Containers and Orchestration in the CERN Cloud, Containers and Orchestration in the CERN Cloud, CERN, Geneva, Switzerland, 8 April 2016
- T. White/Cloudera, Random Decision Forests on Apache Spark, Random Decision Forests on Apache Spark, CERN, Geneva, Switzerland, 12 April 2016
- X. Espinal Curull/CERN, CERN/IT EOS evolution, ALICE, ATLAS, CMS & LHCb Second Joint Workshop on DAQ@LHC, Geneva, Switzerland, 13 April 2016
- C. Faerber/CERN, LHCb R&D with the CPU - FPGA combination, ALICE, ATLAS, CMS & LHCb joint workshop on DAQ@LHC 2016, Château de Bossey, Switzerland, 14 April 2016
- M. K. Patterson/Intel, Trends in HPC and Data Center Power, Packaging and Cooling, Trends in HPC and Data Center Power, Packaging and Cooling, CERN, Geneva, Switzerland, 18 April 2016
- X. Espinal Curull/CERN, L. Mascetti/CERN, A. Pace/CERN, H. Rousseau/CERN, D. Van Der Ster/CERN, M. Lamanna/CERN, J. Iven/CERN, Storage at CERN: towards the Chamäleon, HEPiX 2016, DESY Zeuthen, Germany, 20 April 2016
- G. Swart/Oracle, Breakthroughs in security, efficiency, and performance with on-die hetero-processing, Breakthroughs in security, efficiency, and performance with on-die hetero-processing, CERN, Geneva, Switzerland, 28 April 2016
- M. Floris/CERN, Alice: ML and DA Challenges, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- D. Rousseau/CERN, ATLAS: ML and DA Challenges, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016 F.M. Tilaro/CERN, A. Voitier/CERN, M. Gonzalez Berges/CERN, Challenges for Industrial Control Systems, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- J.R. Vlimant/CERN, CMS: ML and DA Challenges, CERN

openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- A. Ustyuzhanin/CERN, LHCb: ML and DA Challenges, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- E. Fenoglio/Cisco, Machine Learning and Data Analytics at Cisco, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- T. White/Cloudera, Machine Learning and Data Analytics at Cloudera, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- A. Osterloh/Google, Machine Learning and Data Analytics at Google, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- C. Bekas/IBM, Machine Learning and Data Analytics at IBM, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- M.C. Sawley/Intel, Machine Learning and Data Analytics at Intel, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- A. Gattiker/Microsoft, Machine Learning and Data Analytics at Microsoft, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- M. Peter/Nvidia, Machine Learning and Data Analytics at Nvidia, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- V. Tresp/SIEMENS, Machine Learning and Data Analytics at Siemens, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- A. Ustyuzhanin/Yandex, Machine Learning and Data Analytics at Yandex, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- M. Girone/CERN, Welcome and Goals of the workshop, CERN openlab Machine Learning and Data Analytics workshop, CERN, Geneva, Switzerland, 29 April 2016
- M. Martin Marquez/CERN, Mike Olson/Cloudera, Modern data strategy and CERN Opening Keynote, Strata+Hadoop, London, UK, 31 May 2016
- C. Faerber/CERN, Cherenkov angle reconstruction with FPGA compute accelerators, 7th LHCb Computing Workshop, CERN, Geneva, Switzerland, 02 June 2016
- O. Awile/CERN, P. Szostek/CERN, Experiments with multi-threaded velopixel track reconstruction, 7th LHCb Computing Workshop, CERN, Geneva, Switzerland, 2 June 2016
- T. Bell/CERN, Frontiers of Science and Clouds, OpenStack Israel 2016, Tel Aviv, Israel, 02 June 2016
- S. Baymani/CERN, Emerging Technologies: RapidIO, 20th Real Time Conference, Padova, Italy, 05-06 June 2016
- G. Jereczek/CERN, A Lossless Network for Data Acquisition, 20th Real Time Conference, Padova, Italy, 7 June 2016
- O. Awile/CERN, P. Szostek/CERN, A Thread-Parallel Implementation of High-Energy Physics Particle Tracking on Many-Core Hardware Platforms, PASC, Lausanne, Switzerland, 8 June 2016
- L. Lupsa/CERN, Database Evolution and Monitoring (Oracle), CERN openlab open day, CERN, Geneva, Switzerland, 8 June 2016
- M. Castelli/Oracle, Oracle's Keynote: Oracle: Innovations with CERN, CERN openlab open day, CERN, Geneva, Switzerland, 8 June 2016
- M. Martin Marquez/CERN, Oracle's Keynote: Oracle Innovation with CERN Data Analytics Presentation, CERN openlab open day, CERN, Geneva, Switzerland, 8 June 2016
- M. Martin Marquez/CERN, Oracle Data Analytics Project at

CERN openlab Data Analytics Presentation, CERN openlab open day, CERN, Geneva, Switzerland, 9 June 2016

- M.S. Rahman/ISEF winner, A Smart Burn and Spill Proof "SAFE" Microwave that Spares the Salad: Novel Application of Levenberg-Marquardt Algorithms in Bayesian Analysis for Real-Time Numerical Thermodynamic Modeling, , ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- C.V. Yoke/ISEF winner, Constructive Interference of Seismic Surface Waves Antipodal to Crater Impact Sites on Terrestrial Bodies, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- S.W. Kanaski/ISEF winner, Creation of Additional Signal Regions to Increase Signal Sensitivity in the Search for Vector-Like Quarks at the LHC, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- S. Atzpodien/ISEF winner, Deriving an Analytical Algorithm for the Localization of Signal Sources in Orb Webs and Other Net Geometries: A Novel Mathematical Approach to Positioning Systems, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- R. Joshi/ISEF winner, Determining Network Robustness Using Region Based Connectivity, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- M. Truell/ISEF winner, Fido: A Universal Robot Control System Using Reinforcement Learning with Limited Feedback, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- J. Iven/CERN, Introduction, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- A.Z. Yang/ISEF winner, Orbital Recognition System for Space Debris Tracking Using Artificial Neural Networks: A Journey from Inner-Brain GPS to Outer-Space GPS, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- S.A. Maazouz/ISEF winner, Spectroscopic Analysis of Titanium Oxide Presence in Stars, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- M.T. Earle/ISEF winner, Two-Dimensional Mapping of Energy Transfer in Graphene/MoS2 Photodetectors, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- C. Baker/ISEF winner, User Authentication Based on Gait Analysis, ISEF CERN Special Award Winners Lightning talks, CERN, Geneva, Switzerland, 16 June 2016
- Romero Marin/CERN, M. Martin Marquez/CERN, Exploring Data-Driven Decision with Oracle Big Data Discovery Data Analytics Presentation, Cloudera & Oracle Big Data Analytics Innovation Summit, Zurich, Switzerland, 23 June 2016 M. Martin Marquez/CERN, CERN Big Data Strategy & Oracle Big Data Discovery Data Analytics Presentation, DW & Big Data Global Leader Program, Oslo, Norway, 29 June 2016
- O. Awile/CERN, Parallelism in Tracking, DS@HEP, Simons Foundation, New York, USA, 6 July 2016
- T. Bell/CERN, CERN Cloud Experience, PANDAAS2, Grenoble, France, 7 July 2016
- Castro Leon/CERN, GridKA School 2016, Karlsruhe, Germany, 30 August 2016
- L. Lobato Pardavila/CERN, CERN openlab collaboration: Highlights of the journey to Oracle GoldenGate 12c (ENG), Data Capital, Big Data and Data Integration Solution, CERN, Geneva, Switzerland, 13 September 2016
- M. Martin Marquez/CERN, Discovering Hidden Patterns with Oracle Big Data Discovery, Data Capital, Big Data and Data Integration Solution, CERN, Geneva, Switzerland, 13 September 2016

- A. Santogidis/CERN, Data transfer on manycore processors for high throughput applications, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- S. Sridharan/CERN, FPGAs for next DAQ and Computing systems at CERN, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- A. Nowak/CERN, ICE-DIP key technical points, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- M. Zeiler/CERN, Radiation-hard silicon photonics for high energy physics and beyond, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- D. Malone/CERN, G. Jereczek/CERN, Software Switching for Data Acquisition, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- P. Karpinski/CERN, The consequences of poor vectorization, ICE-DIP closing workshop, CERN, Geneva, Switzerland, 13-14 September 2016
- A. Romero Merin/CERN, M. Martin Marquez/CERN, EMEA Big Data breakfast presentation Data Analytics Presentation, Oracle Open World 2016, San Francisco, USA, 19 September 2016
- E. Grancher/CERN, M. Martin Marquez/CERN, What's new for Machine Learning with Oracle Database and Hadoop Data Analytics Presentation, Oracle Open World 2016, San Francisco, USA, 19 September 2016
- E. Pilecki/CERN, Oracle Database InMemory Customer Panel Database Panel, Oracle Open World 2016, San Francisco, USA, 21 September 2016
- J. van Eldik/CERN, The CERN OpenStack Cloud: Compute Resource Provisioning for the Large Hadron Collider, OpenStack Days Nordic Stockholm, Stockholm, Sweden, 22 September 2016
- E. Pilecki/CERN, Using Oracle Database In-Memory to Speed Up CERN Applications Database Presentation Oracle Open World 2016, San Francisco, USA, 22 September 2016
- C. Faerber/CERN, CERN Intel Xeon/FPGA status, Intel CERN HTCC meeting, Portland, USA, 7 October 2016
- X. Espinal Curull/CERN, CERN Data Services for LHC Computing, CHEP 2016, San Francisco, USA, 10 October 2016
- F. Rademakers/CERN, CERN openlab Knowledge Transfer and Innovation Projects, CHEP 2016, San Francisco, USA, 10 October 2016
- M. Martin/CERN, Leveraging Oracle Big Data Discovery to Master CERN's data, Data Analytics Presentation Cloudera & Oracle Big Data Summit, Oslo, Norway, 11 October 2016
- S. Baymani/CERN, RapidIO as a multi-purpose interconnect, CHEP 2016, San Francisco, USA, 11 October 2016
- R. Brito Da Rocha/CERN, Integrating Containers in the CERN Private Cloud, CHEP 2016, San Francisco, USA, 12 October 2016
- F. Rademakers/CERN, CERN openlab Researched Technologies That Might Become Game Changers in Software Development, CHEP 2016, San Francisco, USA, 12 October 2016
- M. Martin/CERN, Leveraging Oracle Big Data Discovery to Master CERN's data, Data Analytics Presentation Cloudera & Oracle Big Data Summit, Stockholm, Sweden, 12 October 2016
- C. Faerber/CERN, Acceleration of Cherenkov angle Reconstruction with the new Intel Xeon/FPGA compute platform for the particle identification in the LHCb Upgrade, CHEP 2016, San Francisco, USA, 13 October 2016
- D.H. Campora Perez/CERN, cross kalman, CHEP 2016, San Francisco, USA, 13 October 2016
- M. Girone/CERN, Engaging Industry for Innovation in the LHC run 3&4 R&D Programme: CERN openlab, CHEP 2016, San

Francisco, USA, 13 October 2016
- X. Espinal Curull/CERN, From Physics to Industry: EOS outside HEP, CHEP 2016, San Francisco, USA, 13 October 2016
- G. Jereczek/CERN, Software switching for the LHC experiments at CERN, Intel Software Professionals Conference, Gdansk, Poland, 18 October 2016
- S. Sridharan/CERN, FPGAs for next gen DAQ and Computing systems at CERN, Amrita University, Coimbatore, India, 21 October 2016
- T. Bell/CERN, OpenStack at CERN, OpenStack Summit 2016, Barcelona, Spain, 25 October 2016
- J. Hughes/Apple, Cassandra and Persistent RAM, Cassandra and Persistent RAM, CERN, Geneva, Switzerland, 27 October 2016
- S. Sridharan/CERN, FPGAs for next gen DAQ and Computing systems at CERN, Amrita University, Bangalore, India, 27 October 2016
- M. Martin/CERN, The Data Transformation, Data Analytics round table Oracle Digital Day, Madrid, Spain, 27 October 2016
- D.H. Campora Perez/CERN, cross kalman, 8th LHCb Computing Workshop, Paris, France, 17 November 2016
- C. Faerber/CERN, RICH reconstruction on FPGAs, 8th LHCb Computing Workshop, Paris, France, 17 November 2016
- O. Awile/CERN, Roofline: performance analysis and code optimization, 8th LHCb Computing Workshop, Paris, France, 17 November 2016
- A. Pace/CERN, A. J. Peters/CERN, J. Leduc/CERN, X. Espinal Curull/CERN, IT-ST: The Storage Group, IT-ST Group, CERN, Geneva, Switzerland, 18 November 2016
- G. Zitzlsberger/Intel, The Intel Xeon Phi Processor Architecture, 2nd CERN openlab/Intel hands-on workshop on code optimisation, CERN, Geneva, Switzerland, 24-25 November 2016
- K. D. Oertel/Intel, Which tool should I use?, 2nd CERN openlab/Intel hands-on workshop on code optimisation, CERN, Geneva, Switzerland, 24-25 November 2016
- H. G. Zimmermann/Siemens, Identification of Complex Dynamical Systems with Neural Networks, CERN, Geneva Switzerland, 5-6 December 2016

Publications:

- M. Zeiler/CERN, Data Acquisition at CERN: A future challenge, IEEE Xplore Digital Library, 9 May 2016
- M. Zeiler/CERN, Silicon Photonic Components for Optical Data Links in High Energy Physics Experiments, www.europractice-ic.com Activity Report 2015, June 2016
- S. Baymani/CERN, K. Alexopoulos/CERN, S. Valat/CERN, Exploring RapidIO technology, within a DAQ system event building network, 20th Real Time Conference, Padova, Italy, 05-06 June 2016
- C. Auffray/European Institute for Systems Biology and Medicine, R. Balling/Luxembourg Centre for Systems Biomedicine, A. Di Meglio/CERN, et al., Making sense of big data in health research: Towards an EU action plan, Genome Medicine 2016, vol. 8, no. 71.
- R. Bauer/Newcastle University, L. Breitwieser/CERN, A. Di Meglio/CERN, L. Johard/Innopolis University, M. Kaiser/ Newscastle University, M. Manca/CERN, M. Mazzara/ Innopolis University, M. Talanov/Kazan Federal University, The BioDynaMo Project, arXiv:1607.02717, 10 July 2016 2)
- A. Di Meglio/CERN, M. Manca/CERN, From Big Data to Big Insights: The Role of Platforms in Healthcare IT, in New Perspectives in Medical Records, Rinaldi, Giovanni (ed.), Springer Publishing, 2016

- M. Zeiler/CERN, S. Detraz/CERN, L. Olantera/CERN, C. Sigaud/ CERN, C. Soos/CERN, J. Troska/CERN, F. Vasey/CERN, A system-level model for high-speed, radiation-hard optical links in HEP experiments based on silicon Mach-Zehnder modulators, Journal of Instrumentation, vol. 11, no. 12, pp. C12059–C12059, 2016.
- M. Zeiler/CERN, S. Detraz/CERN, L. Olantera/CERN, C. Sigaud/CERN, C. Soos/CERN, J. Troska/CERN, F. Vasey/CERN, Comparison of the Radiation Hardness of Silicon Mach-Zehnder Modulators for Different DC Bias Voltages, in IEEE Nuclear Science Symposium/Medical Imaging Conference (NSS/MIC), 2016

Posters:

- F. Tilaro/CERN, A.Voitier/CERN, M. Gonzalez Berges/CERN, M. Roshchin/Siemens, Smart Data analysis of CERN Control Systems
- S. Baymani/CERN, K. Alexopoulos/CERN, S. Valat/CERN, Exploring RapidIO technology, within a DAQ system event building network, 20th Real Time Conference, Padova, Italy, 05-06 June 2016
- C. Faerber/CERN, N. Neufeld/CERN, R. Schwemmer/CERN, J. Machen/Intel, Particle identification on an FPGA accelerated compute platform for the LHCb Upgrade, 20th Real Time Conference, Padova, Italy, 07 June 2016
- L. Lupsa/CERN, CERN openlab in partnership with Oracle bringing early access to Oracle database 12c Release 2 features, CERN openlab open day, CERN, Geneva, Switzerland, 08-09 June 2016
- S. Trigazis/CERN, Containers and the CERN Cloud, CERN openlab open day, CERN, Geneva, Switzerland, 08-09 June 2016
- B. Voneki/CERN, S. Valat/CERN, J. Machen/Intel, N. Neufeld/ CERN, Evaluation of 100Gb/s LAN networks for the LHCB DAQ upgrade, CERN openlab open day, CERN, Geneva, Switzerland, 08-09 June 2016
- I. Charalampidis/CERN, L. Lazaridis/CERN, Optimizing the current transport layer of the ALFA framework using Cisco's usNIC fabric, CERN openlab open day, CERN, Geneva, Switzerland, 08-09 June 2016
- C. Farber/CERN, N. Neufeld/CERN, R. Schwemmer/CERN, J. Machen/Intel, Particle identification on an FPGA accelerated compute platform for the LHCb Upgrade, CERN openlab open day, CERN, Geneva, Switzerland, 08-09 June 2016
- O. Awile/CERN, N. Hardi/CERN, A. Santogidis/CERN, A graphical performance analysis and exploration tool for Linux perf, CHEP 2016, San Francisco, USA, 10-14October 2016
- M. Girone/CERN, CERN openlab: Engaging Industry for Innovation in the LHC Run3-4 R&D Programme, CHEP 2016, San Francisco, USA, 10-14 October 2016
- J. Wyatt Smith/University of Goettingen, G. A. Stewart/ University of Glasgow, R. Seuster/University of Victoria, A. Quadt/University of Goettingen, ATLAS software stack on ARM64, CHEP 2016, San Francisco, USA, 10-14 October 2016

The CERN openlab open day marked exactly 15 years since our unique public-private partnership was established.

Events and Outreach

Creating and disseminating knowledge

Together, we can share a vision for the future of scientific computing.

CERN openlab undertakes outreach and communications activities in support of its technical work. The collaboration's results were disseminated at a wide range of international conferences in 2016. CERN openlab provides CERN and its members with a means to share a vision of the future of scientific computing. This is made possible through the organisation of joint workshops and events. The vision is communicated to a wide audience, including partner clients, members of the press, and the public.

Top delegations from governments and industry frequently tour CERN: 142 protocol visits were organised in 2016. The CERN openlab concept and projects are systematically presented to the guests visiting the CERN IT department.

CERN openlab partners regularly organise customer and press visits. These groups are briefed about CERN openlab in a dedicated VIP meeting room known as the CERN openlab 'openspace'. More than 200 press articles were published about our work over the course of the year.

An important element of CERN openlab's outreach strategy is the annual 'open day' event. The 2016 edition was the perfect opportunity to celebrate our 15th anniversary. As well as featuring both talks and posters on the achievements of CERN openlab's diverse range of projects, the two-day event offered attendees the chance to participate in hands-on technology demonstrations from the collaborating companies.

Over 50 workshops, lectures, visits and other events related to CERN openlab's work were held throughout the year. Visits and guest seminars are listed on pages 42 and 43. Further information about these can be found on our website, including many recordings.

On the CERN openlab website, you can also find a complete collection of press coverage, 'sponsor spotlights', press releases, and other case studies from our collaborators. In addition, technical documents, presentations, and reports can be consulted on the site: **cern.ch/openlab**.

Representatives of CERN openlab's collaborators gave presentations at the CERN openlab open day.

CERN's Charlotte Warakaulle presents an award to Peter Gleissner of Intel in recognition of 15 years of fruitful collaboration with the company through CERN openlab.

Visit to Oracle HQ
25-27 January 2016
Visit to Oracle's headquarters in California, USA. CERN participants: Eric Grancher, Katarzyna Maria Dziedziniewicz-Wojcik, Artur Wiecek, and Antonio Romero Marin. Oracle participants: Cris Pedregal, Pauline Gillet-Mahrer, and Jean-Philippe Breysse

Oracle VIP visit to CERN
27 January 2016
Visit of Joel Cherkis and Rachael Hartley from Oracle.

Cisco visit to CERN
8 March 2016
Visit of Artur Barczyk from Cisco, accompanied by Torsten Alt from FIAS — Frankfurt Institute for Advanced Studies and Mikolaj Krzewicki from Johann-Wolfgang-Goethe University, also in Frankfurt, Germany.

ISEF 2016 students visit to CERN
13-17 June 2016
The 10 winners of the 'CERN special award' came to CERN for an extended tour and educational programme.

Google interns visit to CERN
12 August 2016
Interns from Google visit CERN for a tour full-day tour of the laboratory's facilities.

Intel Modern Code Developer Challenge Winners' tour of CERN
18-21 September 2016
CERN tour for first-place winners, accompanied by Russel Beutler of Intel.

Seminars by CERN openlab guests

Intel in HPC: innovation beyond the core
29 February 2016
William Magro of Intel.

Computational challenges of finding the largest prime
09 March 2016
Landon Curt Noll of Cisco.

Kinetic drive integration in EOS
18 March 2016
Paul Hermann Lensing of Seagate.

Random Decision Forests on Apache Spark
12 April 2016
Tom White of Cloudera.

Trends in HPC and data centre power, packaging, and cooling
18 April 2016
Michael K. Patterson of Intel.

Breakthroughs in security, efficiency, and performance with on-die hetero-processing
28 April 2016
Garret Swart of Oracle.

'Lightning talks' by ISEF CERN Special Award winners
16 June 2016
Sophie Atzpodien, Chloe Baker, Michael Thomas Earle, Rucha Joshi, Sloan Wayne Kanaski, Sarah Amina Maazouz, Muhammad Shahir Rahman, Michael Truell, Amber Zoe Yang, Camille Virginia Yoke.

Cassandra and persistent RAM
27 October 2016
James Hughes of Apple.

Identification of complex dynamical systems with neural networks
5-6 December 2016
Hans Georg Zimmermann of Siemens.

In 2017, the LHC will continue to produce a wealth of data at unprecedented energies. As well as enabling more precise definition of the characteristics of the Higgs boson, this data will be a significant boon to the search for new physics.

Many exciting scientific and technological accomplishments lie ahead, and CERN looks forward to another year of brilliant performance across the varied work of the laboratory. With the LHC experiments exploring new territories of energies, alongside experiments exploring antimatter, astroparticle physics, and more, CERN's researchers look set to reap fascinating new results in 2017.

With a new three-year phase of CERN openlab set to launch at the start of 2018, work is now underway to identify key areas for future collaboration. In early 2017, workshops are being held to discuss the ICT challenges faced by 'big science' over the coming years. In September, CERN openlab will publish a whitepaper based on the outcomes of these workshops. With the LHC and the experiments set to undergo major upgrade work in 2019 and 2020, CERN openlab's sixth phase offers a clear opportunity to develop solutions that will make a tangible difference for researchers at CERN when the upgraded LHC and experiments come online in 2021.

2016 marked 15 years since the founding of CERN openlab, thus offering an opportunity to look back at the work carried out through this unique public-private partnership in support of research at CERN and beyond. This anniversary also offers an excellent opportunity to look forward.

We're currently seeing the emergence of a whole new set of technology paradigms, from pervasive ultra-fast networks of smart sensors in 'the internet of things', to machine learning and 'the optimisation of everything'. These technologies have the power to revolutionise the way big science is done, particularly in terms of data analysis and the control of engineering processes. They have enormous potential, not just for research, but also for wider society. The knowledge and expertise at CERN can play a key role in ensuring this potential is realised, with these technologies being put to use for the benefit of both science and society. CERN openlab — with its research members across a range of scientific fields and its unique collaboration with several of the world's leading ICT companies — is ideally positioned to help make this a reality.

Work is underway to identify the ICT challenges that will be addressed during CERN openlab's sixth phase.

The Future

Preparing for our sixth phase

With CERN openlab's next three-year phase set to begin at the start of 2018, work is now underway to define the ICT challenges we will tackle together with our collaborators.