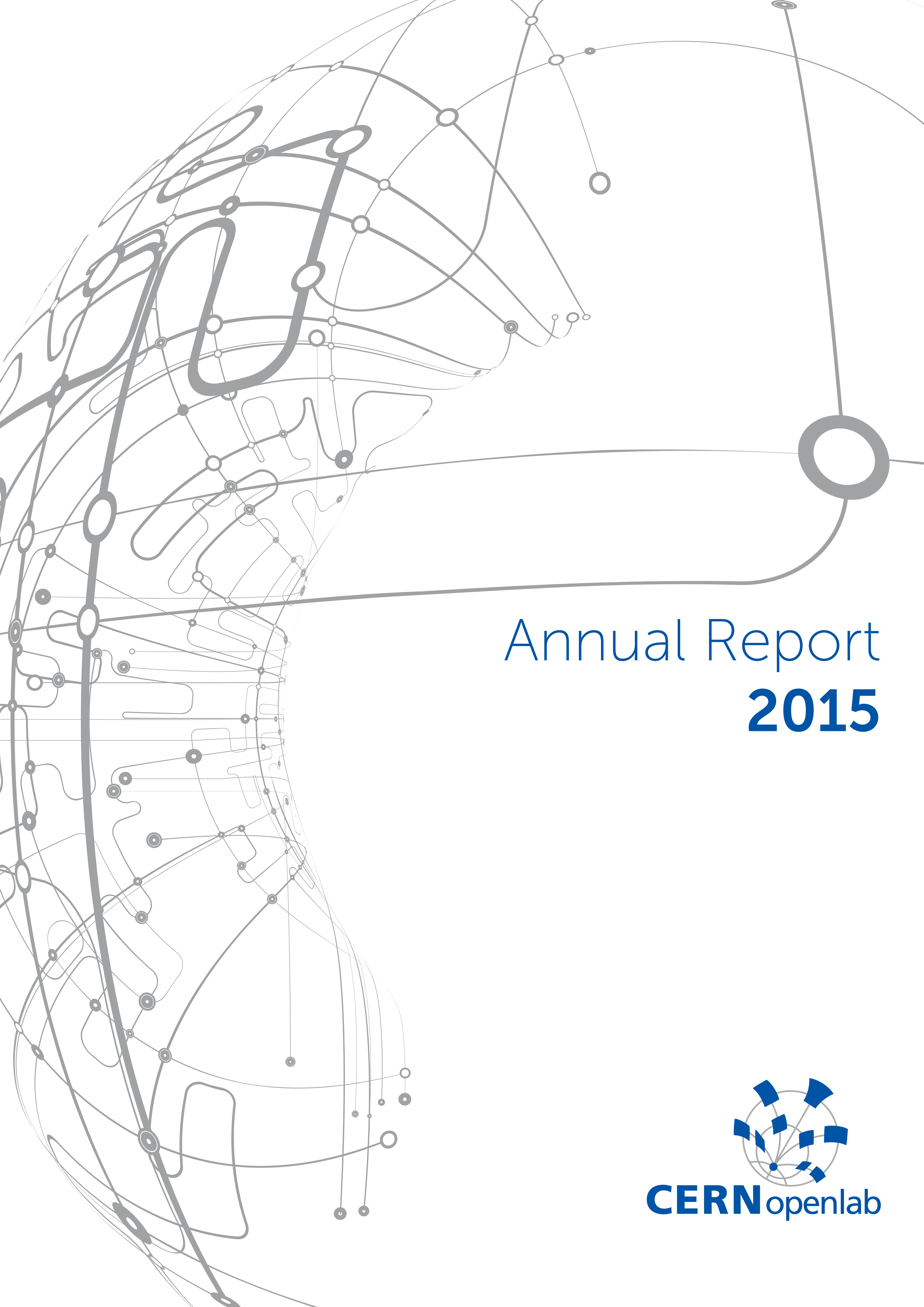
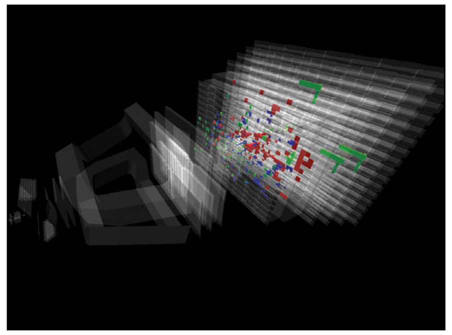
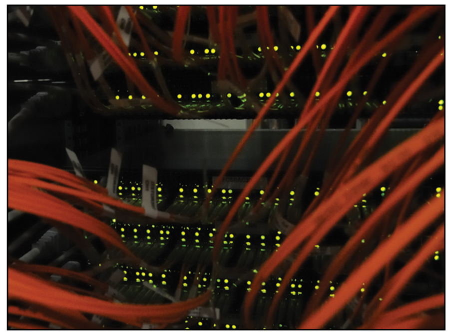
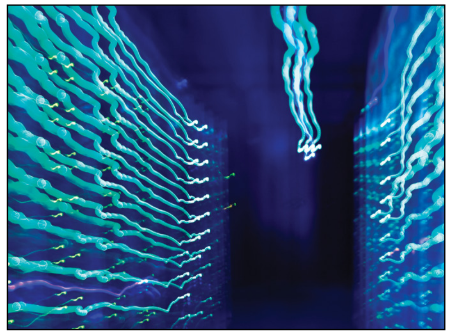
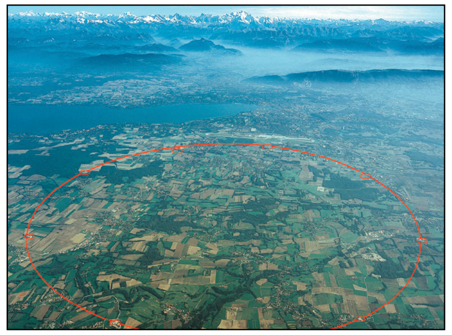
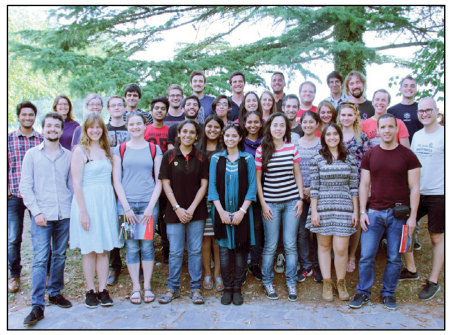
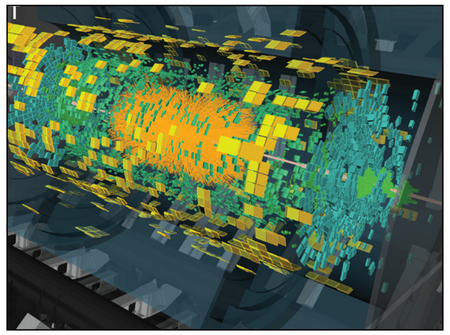


Annual Report
2015

CERN openlab


Contents

A Word from the DG 5

The Context 6

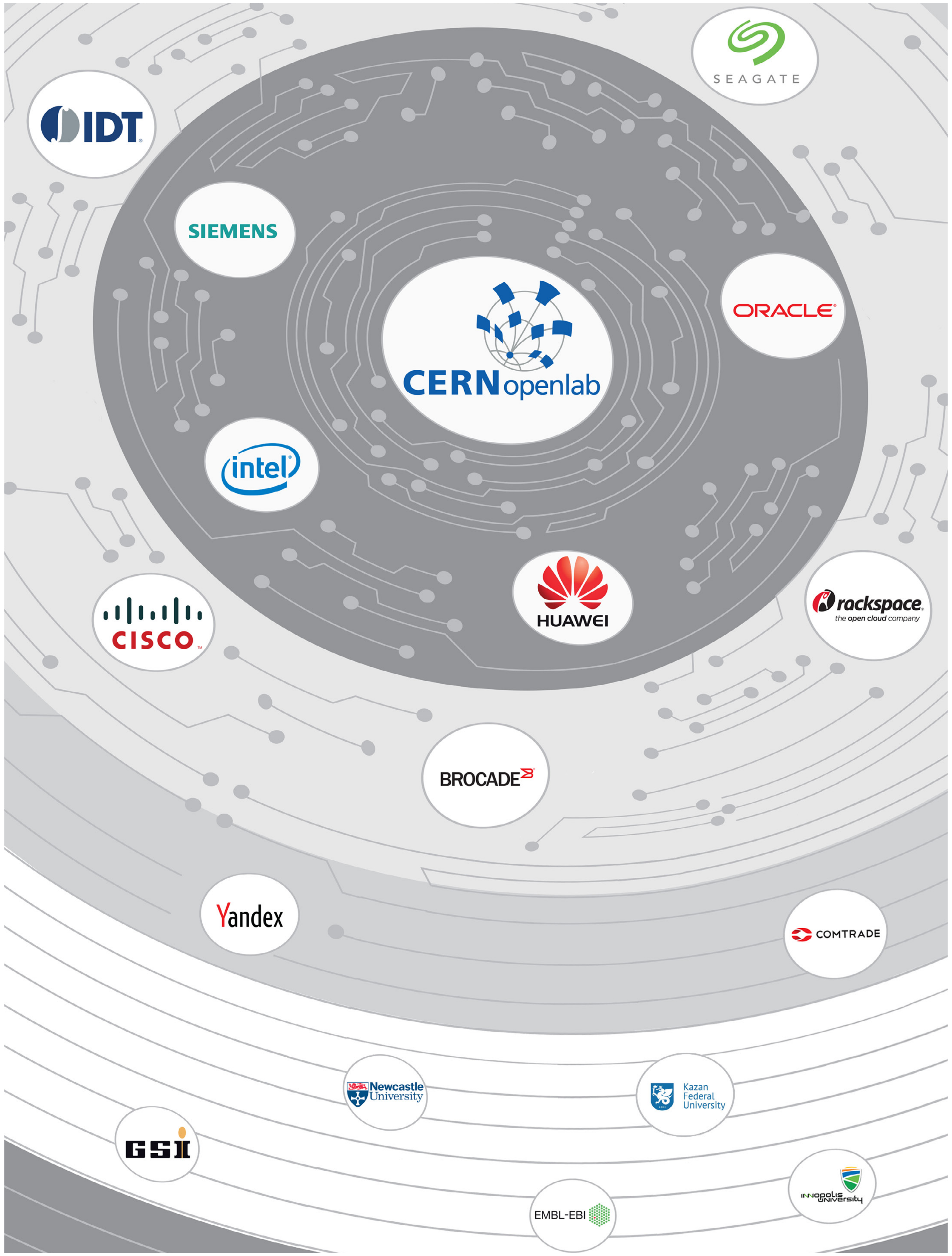
The Concept 10

The Results 16

Education 34

Events and Outreach 40

The Future 44



A Word from the DG

CERN openlab started a new phase in 2015, with several new companies joining and the collaboration broadening to include other research institutions for the first time. Founded in 2001 to develop the innovative systems needed to cope with the unprecedented ICT challenges of the LHC, CERN openlab unites science and industry at the cutting edge of research and innovation.

CERN openlab has a proven track record of ensuring that members of CERN's scientific community have access to the very latest ICT solutions to help them carry out their ground-breaking physics research. By collaborating with scientific institutions on ICT challenges common across large research infrastructures, CERN openlab is able to ensure maximum relevance of its work in this new phase.

Within the CERN openlab framework, CERN provides access to its complex ICT infrastructure and its engineering experience. 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.

The flexibility of CERN openlab makes it possible for partnerships to flourish when the time is right for both CERN and industry. The results presented in this report provide tangible evidence of the virtuous circle linking basic and applied science.

The work carried out in 2015 has been of paramount importance in ensuring CERN's ICT capacity continues to meet the evolving challenges posed by the ambitious LHC programme, including the planned upgrades.

Education and training are also central to CERN openlab's mission. The young researchers hired by CERN and funded by the collaborating companies play a key role in CERN openlab's work. These young researchers and students demonstrate remarkable curiosity and bring fresh new viewpoints to the collaboration; they are the catalyst for many creative ideas.

I would like to thank all of our current and past members for their support, and I wish all of the CERN openlab collaborators the very best of luck as we continue this exciting new phase together.

The Context

Scenes in the CERN Control Centre as stable beams mark the start of Run 2 physics at the LHC.

New horizons

Following extensive upgrade work, the LHC gets off to a flying start for its new three-year running period.

A new beginning

On 12 January 2015, following two years of work on the entire CERN accelerator complex, the team in charge of the colossal Long Shutdown 1 (LS1) project handed the symbolic key to the LHC over to the operations team. During the two years of LS1, an impressive amount of work was accomplished in preparation for running the LHC at 13 TeV. Eighteen of the machine's 1 232 dipole magnets, which guide the beams around their 27-kilometre orbit, were replaced due to wear and tear. More than 10 000 electrical interconnections between magnets were fitted with shunts to provide an alternative path for the 11 000-amp current, protecting the interconnection if there is a fault. Many of the machine's electronic components were replaced, the vacuum system that keeps the beam pipe clear of stray molecules was upgraded, and the cryogenics systems were refurbished.

With the LHC restarting not only at higher energy, but also with higher luminosity (a measure of the rate of particle collisions delivered to the experiments), the LHC experiments were also busy during LS1. To prepare for the challenge of more collisions, the experiments carried out full consolidation and maintenance programmes, including upgrades to their subdetectors and data-acquisition systems.

Bigger data

To prepare for the increased amount of data in Run 2, the LHC experiment teams and the Worldwide LHC Computing Grid (WLCG) collaboration have upgraded the computing infrastructure and services. During LS1, the IT Department doubled the capacity available to the LHC experiments, with the addition of some 100 petabytes (PB) of disk storage and almost 60 000 new cores. The compute capacity of the CERN private cloud has nearly doubled during 2015, now providing over 150 000 computing cores. LS1 offered an ideal opportunity to migrate the archived data from legacy cartridges and formats to higher density ones. This involved migrating around 85 PB of data, and was carried out in two phases during 2014 and 2015. As an overall result of this two-year migration, no less than 30 000 tape cartridge slots were released to store more data.

But it was not just the LHC, experiments, and computing facilities

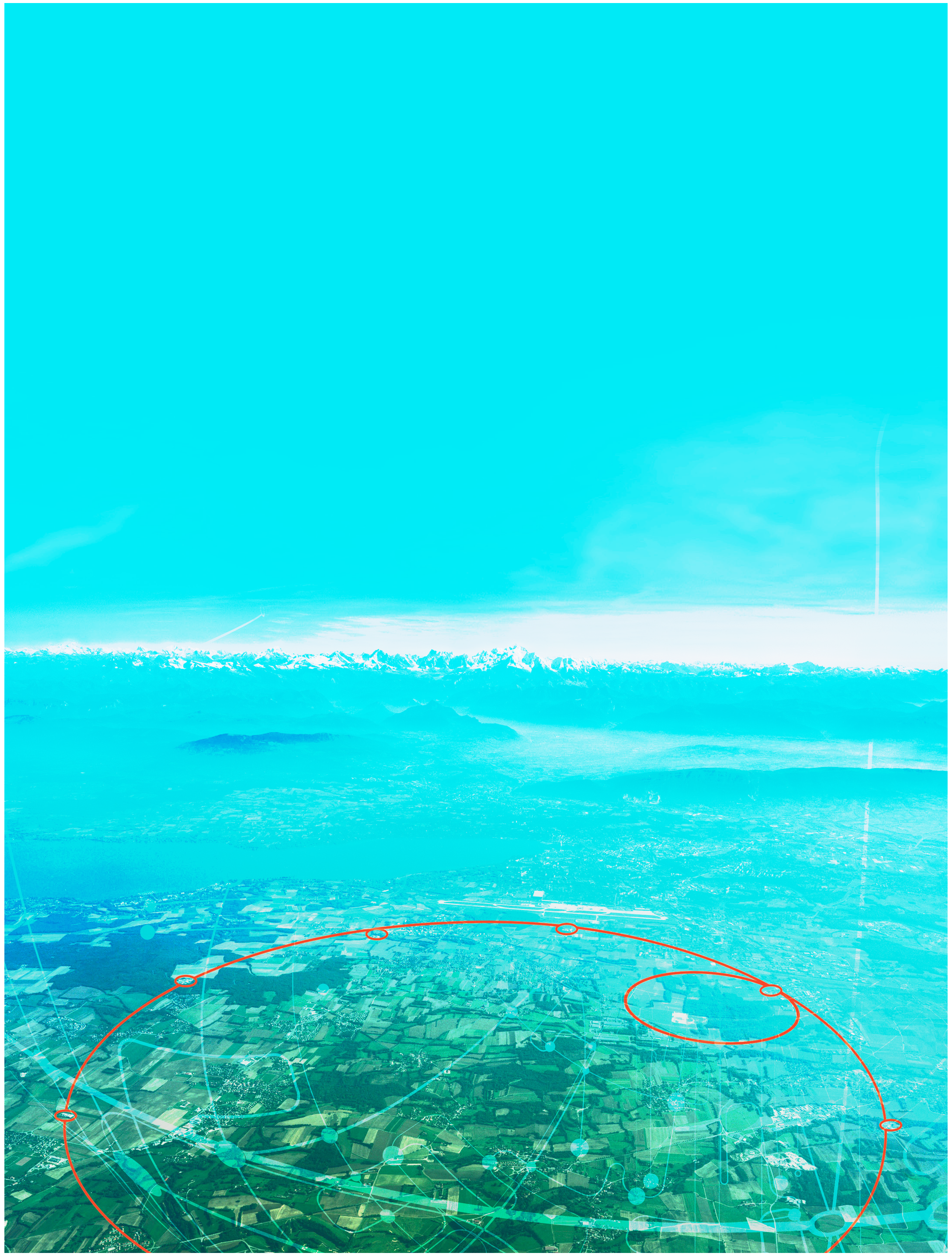
In preparation for the increased data rates of Run 2, significant upgrades were made to the various ICT systems that support the LHC.

that underwent rejuvenation during LS1. CERN's accelerators upstream of the LHC support a vibrant research programme, as well as serving as the injector chain for the LHC itself. The oldest accelerator still in operation, the Proton Synchrotron, first started up in 1959, and LS1 provided an ideal opportunity to carry out essential maintenance to ensure optimum performance and reliability for the future. When the key was handed over on 12 January, it was to an entirely renovated accelerator complex.

New run, new records

Three months later, all the hard work of LS1 paid off as proton beams circulated in the LHC on 5 April, an important milestone on the way to the start of physics data-taking at 13 TeV on 3 June. The Brout-Englert-Higgs mechanism, dark matter, antimatter, and quark-gluon plasma are all on the menu for LHC Run 2. After the discovery of the Higgs boson in 2012, physicists are now putting the Standard Model of particle physics to its most stringent test yet as they continue their search for new physics.

With the start of Run 2, new data-taking records were achieved. 40 PB of data were written on tape at CERN in 2015. Out of the 30 PB from the LHC experiments, a record-breaking 7.3 PB were collected in October, and up to 0.5 PB of data were written to tape each day during the heavy-ion run. By way of comparison, CERN's tape-based archive system collected in the region of 70 PB of data in total during the first run of the LHC. The WLCG also set a new record in 2015 by running a total of 51.1 million jobs in October.



The Concept

Delegates at the 2015 meeting of the CERN openlab Collaboration Board.

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 complex 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. 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 first 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. These needs were identified following an in-depth cross-community consultation exercise at the end of CERN openlab's fourth phase, involving a series of technical meetings and in-depth discussions with representatives of diverse research organisations. This exercise

Members of the CERN openlab management team.

resulted in the publication of the 'CERN openlab Whitepaper on Future IT Challenges in Scientific Research'.

CERN openlab is growing 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 held a public 'open day' event for the first time in 2015, enabling more people than ever before to learn about our exciting work.

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 13, while the positioning of CERN openlab activities within CERN is detailed on pages 14 and 15.

The distributed team structure permits 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 CERN openlab Summer Student Programme, either directly to CERN openlab activities or more widely to WLCG and other CERN activities in the IT Department.

At the start of 2015, Alberto Di Meglio succeeded Bob Jones as head of CERN openlab.

CERN openlab Management

Rolf Heuer	CERN Director General, Chair of CERN openlab Collaboration Board
Frédéric Hemmer	Head of CERN IT Department
Alberto Di Meglio	Head of CERN openlab
Fons Rademakers	Chief Technology Officer
Maria Girone	Liaison with LHC experiments
Kristina Gunne	Administrative and Finance Officer
Konstantinos Papangelopoulos	Junior Administrative Officer
Mélissa Gaillard	Communications Officer
Andrew Purcell	Communications Officer
Petya Georgieva	Junior Communications Officer

CERN openlab personnel (collaborator indicated)

Sepo Akila	Huawei (Fellow)
Maria Arsuaga Rios	Huawei Fellow)
Pawel Lukasz Szostek	Intel (Fellow)
Omar Awile	Intel (Staff)
Christian Faerber	Intel (Fellow)
Andrei Gheata	Intel (Staff)
Karel Ha	Intel (Technical Student)
Sebastien 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)
Lorena Lobato Pardavila	Oracle (Fellow)
Flavia Castro Alves	Siemens (Technical Student)
Carlos Garcia Calatrava	Siemens (Technical Student)
Pavel Fiala	Fellow (Siemens)
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)
Paul Lensing	Seagate (Cooperation Associate)
Lukas Breitwieser	Newcastle University, Kazan Federal University, Innopolis University (Technical Student)

CERN contributing personnel (collaborator indicated)

David Collados	Oracle
Andrei Dumitru	Oracle
Katarzyna Maria Dziedziniewicz-Wojcik	Oracle
Artur Wiecek	Oracle
Nicolas Marescaux	Oracle
Zbigniew Baranowski	Oracle
Lorena Lobato	Oracle
Piotr Golonka	Siemens
Axel Voitier	Siemens
Edoardo Martelli	Brocade
Stefan Stancu	Brocade
Marco Manca	Newcastle University, Kazan Federal University, Innopolis University

ICE-DIP Fellows

Grzegorz Jereczek
Przemyslaw Karpinski
Aram Santogidis
Srikanth Sridharan
Marcel Zeiler

CERN openlab project coordinators (collaborator indicated)

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

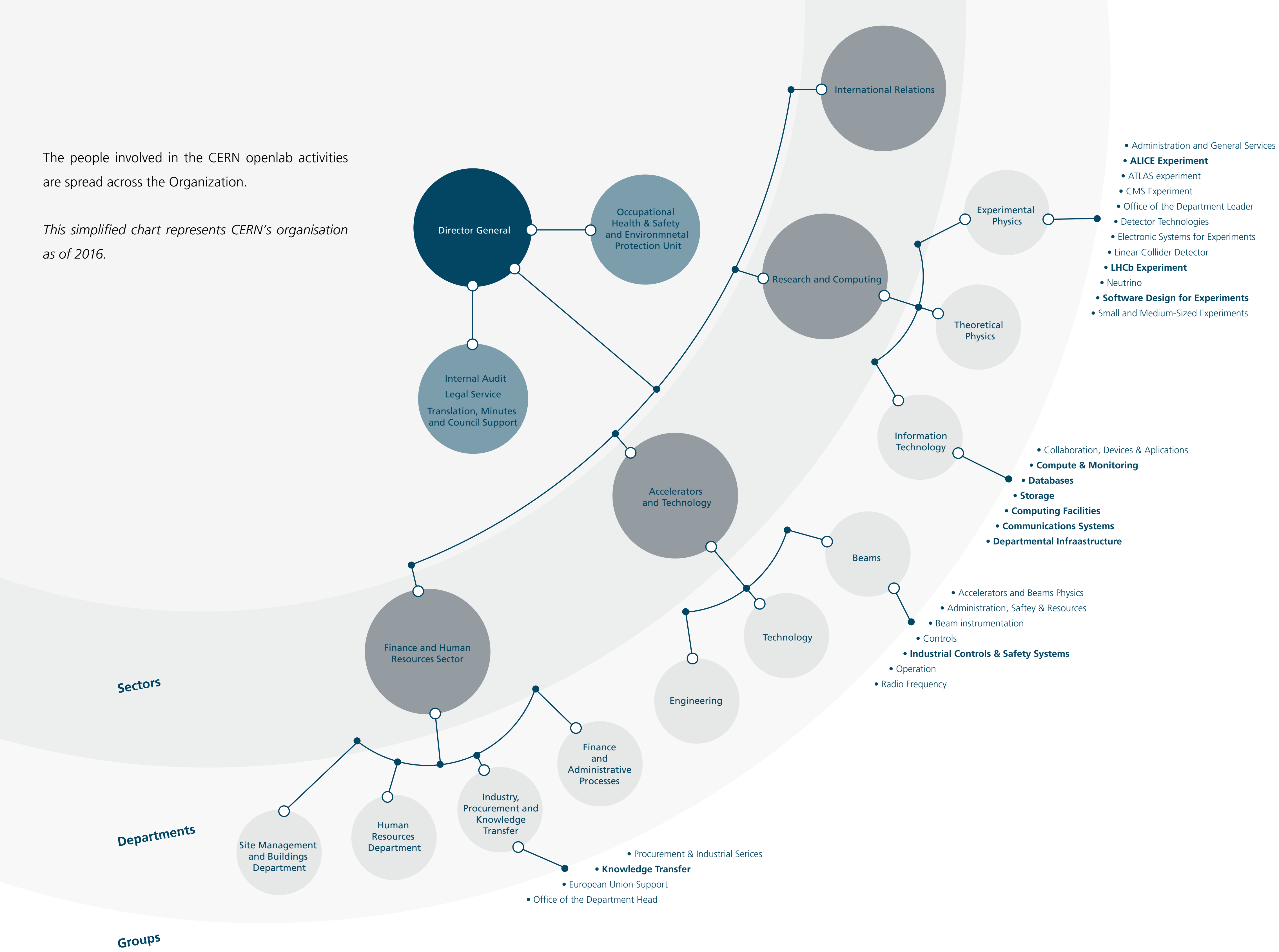
Collaborator liaisons with CERN

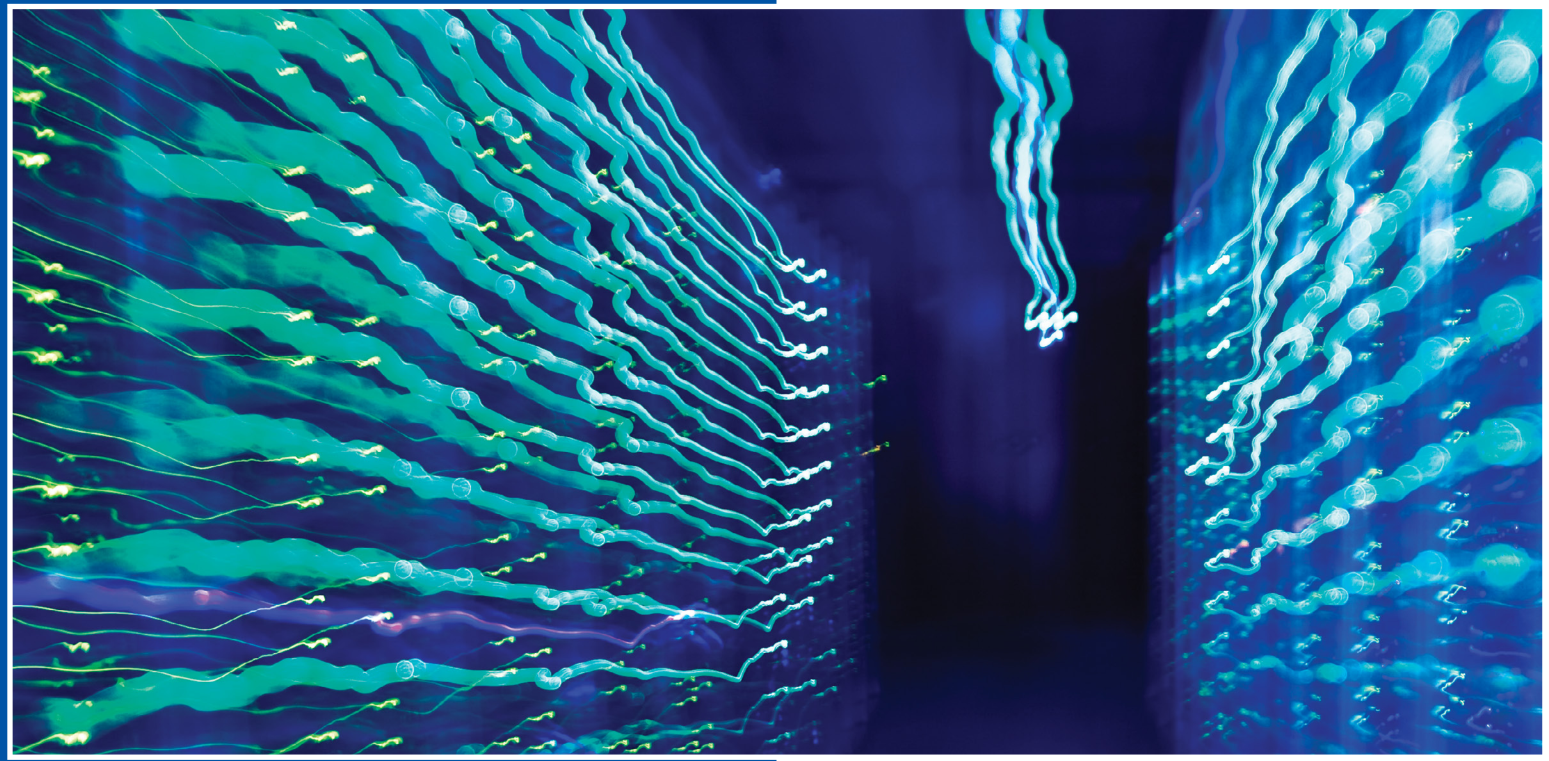
Davis Wu	Huawei
Yuan Yuan	Huawei
Karl Solchenbach	Intel
Claudio Bellini	Intel
Laurent Duhem	Intel
Michal Dzoga	Intel
Stephan Gillich	Intel
Jonathan Machen	Intel
Marie-Christine Sawley	Intel
Monica Marinucci	Oracle
Jean-Philippe Breysse	Oracle
Dmitrij Dolgušin	Oracle
Pauline Gillet-Mahrer	Oracle
Rachael Hartley	Oracle
Kevin Jernigan	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
Devashish Paul	IDT
Giri Fox	Rackspace
Paul Voccio	Rackspace
Adrian Otto	Rackspace
Paul Kusbel	Seagate
Goran Garevski	Comtrade
Gregor Molan	Comtrade
Andrey Ustyuzhanin	Yandex
Steven Newhouse	EMBL-EBI
Mohammad Al-Turany	GSI
Manuel Mazzara	Innopolis University
Max Talanov	Kazan Federal University
Roman Bauer	Newcastle University

Positioning CERN openlab activities at CERN

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

This simplified chart represents CERN's organisation as of 2016.





The Results

The results of the 17 CERN openlab projects active in 2015 are organised by technical challenge.

Data Acquisition (online) ***Page 18***

Intel High-Throughput Computing Collaboration (HTCC)

IDT RapidIO for Data Aquisition

Networks and Connectivity ***Page 20***

Brocade Flow Optimiser Software

Data Storage Architectures ***Page 21***

Oracle Database Technology and Monitoring

Huawei Cloud Storage Appliance

Seagate Alternative Storage Architecture 'Kinetic'

Comtrade EOS Productisation

Compute Management and Provisioning ***Page 24***

Oracle Java EE

Oracle Database Cloud

Rackspace Cloud Federation

Computing Platforms (offline) ***Page 26***

Intel Code Modernisation

Cisco Data Plane Computing System (DPCS)

Data Analytics ***Page 29***

Oracle Analytics-as-a-Service

Siemens Industrial Control and Monitoring

Yandex Data Popularity at LHCb

Yandex Anomaly Detection in LHCb Online Data Processing

Innovation and Entrepreneurship ***Page 32***

Intel Innovation Management and Entrepreneurship

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

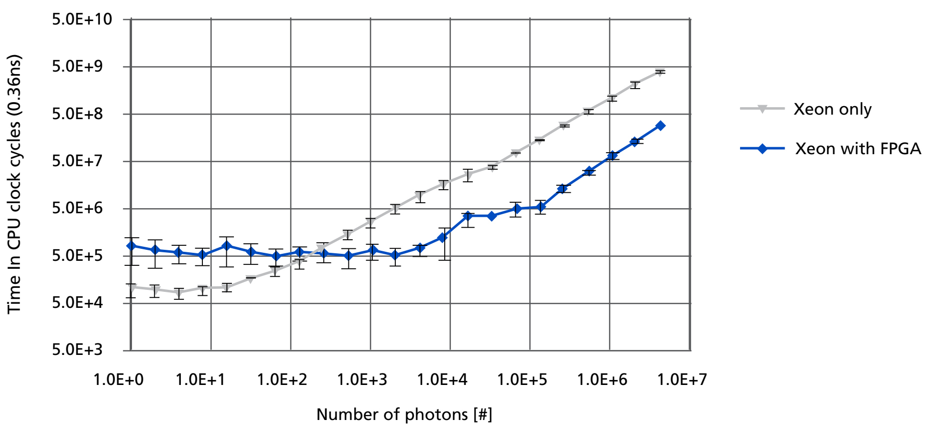
In the last 10 years or so, particle physicists have started to rethink and revise code and algorithms to get more performance out of the computing hardware. Throughout 2015, the High-Throughput Computing Collaboration (HTCC) studied three key technologies: Intel Omni-Path, Intel Xeon/FPGA, and Intel Xeon/Phi.

Intel Omni-Path was used for data-exchange at 100 gigabits per seconds (Gb/s) between servers. We have developed a fabric-agnostic multi-protocol test-suite that allows us to run event-building traffic in various scenarios on networks of practically any scale. Our performance measurements for a full event-building emulation on 16 nodes shows that using Omni-Path and libfabric a bandwidth of up to 73 Gb/s per port (full duplex) can be achieved. 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 understand and test this new type of accelerator, a number of performance studies were conducted. The first such test case, an event-sorting algorithm, achieved an acceleration factor of 50 over a dual-Xeon v2 system. Our second test case was an algorithm for particle identification: using FGPA, a 15-25 factor speedup over a vectorised CPU-only version was achieved.

Intel Xeon/Phi was used to accelerate those algorithms that can be (re)written in an efficient parallel form. One of the algorithms studied relates to straight-line particle tracking in the VeloPixel subdetector. Our prototype code allows us to explore different implementation choices, such as the programming model (OpenMP or TBB) and a number of different parameters. The Intel Xeon/Phi platform, with its large memory bandwidth, is an interesting candidate for event sorting: benchmarks show a throughput of more than 50 gigabytes per second (GB/s) can be achieved on an Intel Knight's Landing platform.

In modern particle physics experiments, huge amounts of data need to be transported and filtered. The processing of the data comprises low-level data-manipulation, pattern-extraction, and sophisticated statistical treatment. Many of these are ideal candidates for acceleration by FPGA or Xeon/Phi. In conjunction with a fast data-transport using Omni-Path and Xeon-processors, HTCC is developing a blueprint for the next generation of high-throughput event-selection.

Comparison of runtimes for Cherenkov angle reconstruction with Xeon only and Xeon with FGPA.



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 primarily revolves around three use cases, in which different aspects and capabilities of RapidIO will be explored. Each use case focuses on a different area: (i) data analytics, (ii) data acquisition, and (iii) real-time triggering systems.

For the first use case, which was the primary focus of 2015, we are looking at RapidIO from a big-data point of view to see if it is possible to improve the efficiency of data analytics at the CERN Data Centre, as well as in other data-analytics frameworks.

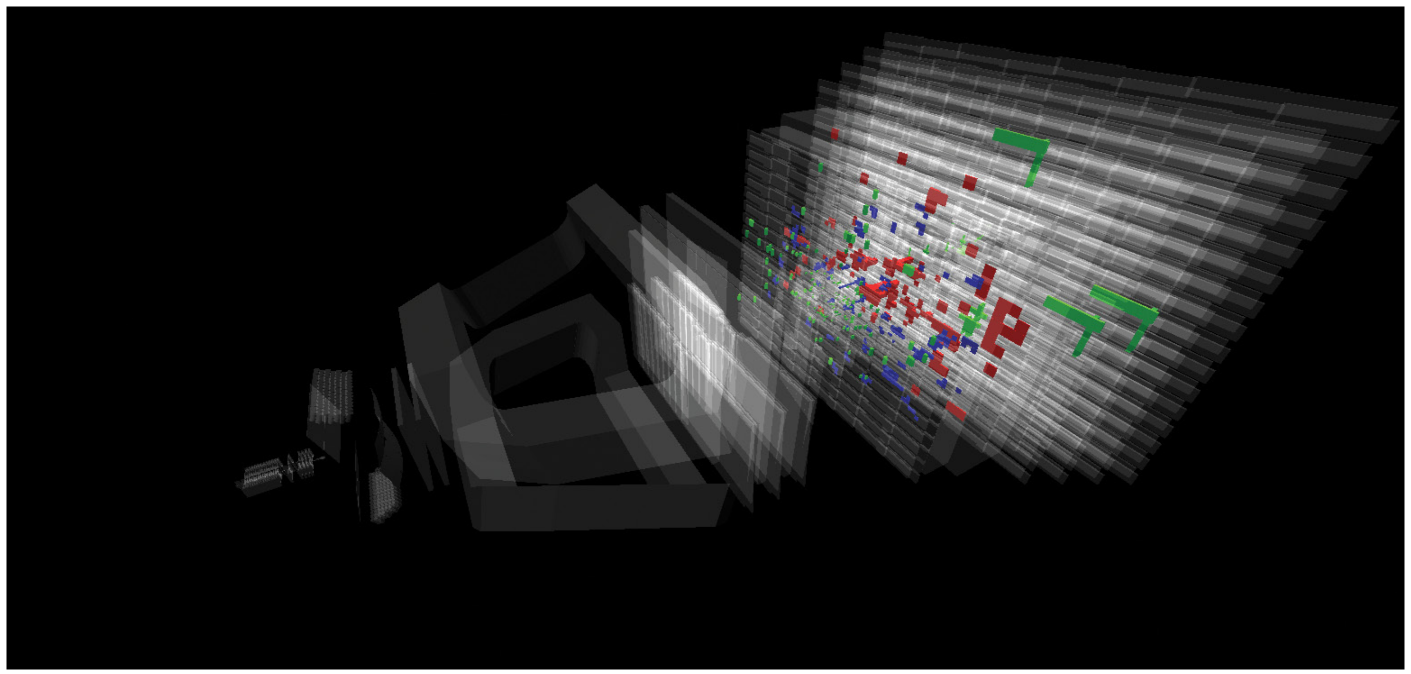
In use case two, we are using RapidIO in a data-acquisition network, primarily that of the LHCb experiment. As traffic in data-acquisition networks tends to be bursty, it requires a high sustained bandwidth from the underlying interconnect. We are investigating whether RapidIO can support this kind of traffic.

For the third use case, the goal is to implement a real-time triggering system with RapidIO used as the back plane. Triggering systems require low and deterministic latency. In the specific case of the LHCb triggering system, this is required in a heterogeneous computing environment in which all units are connected through a uniform back plane.

The project kicked-off in mid-2015 with training by IDT personnel at CERN and set up of a RapidIO test cluster on site: standard x86 servers were connected to a RapidIO-switched network, after which development for the data analytics use case could begin. We started with implementation of a file transfer application, to explore the low-level APIs and to benchmark data transfer rates in a native environment. We also installed Hadoop on the test cluster to evaluate TCP/IP emulation over RapidIO. From there, work started to port the physics data analysis framework ROOT to RapidIO. This work continued until the end of 2015, at which point IDT launched a data analytics 'kit' based on some of the activity of the CERN researchers.

In November 2015, a visit was organised to IDT's offices in Ottawa, Canada. During this week-long visit we held a project evaluation meeting, as well as additional training on RapidIO.

An event recorded by the LHCb detector during first 13 TeV collisions on 21 May 2015.



Networks and Connectivity

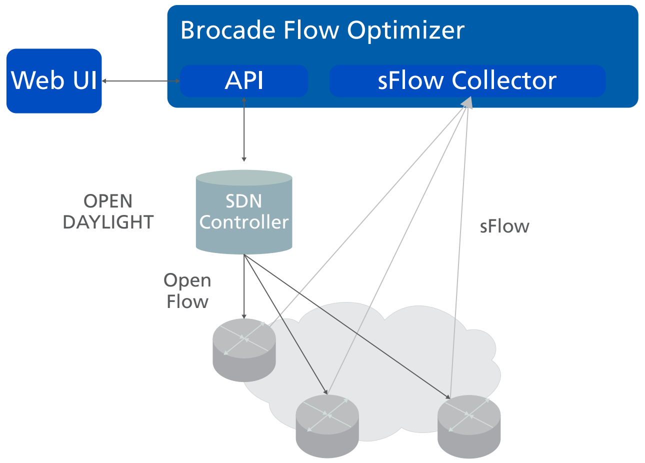
Today, the ever increasing external data traffic (more than 100 Gb/s) is putting pressure on the CERN firewalls. CERN and other large organisations require robust, scalable network solutions that provide high bandwidth for data transfers while maintaining appropriate levels of security. Networking also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

Brocade Flow Optimiser Software

The Brocade Flow Optimizer (BFO) project aims to create an intelligent system that can optimise and secure the routing of data traffic entering and leaving the Organization. The project aims to enhance and generalise the BFO application, to deliver a system capable of identifying whether traffic is legitimate, thus routing it optimally, or part of a network attack, thus routing it for special analysis. This decision is based on information coming from the network itself, from a database of trusted applications, and from other data sources (e.g. black-lists of known threats).

BFO is a software-defined networking (SDN) application that uses the OpenFlow protocol for 'programming' the fast specialised application-specific integrated circuits (ASICs), which serve as the hardware-forwarding engines in network devices.

A diagram showing the proposed BFO system.



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

The system proposed would consist of the following:

- A set of OpenFlow-capable network devices placed at the border of an organisation connected to the Internet upstream on one side, and facing the internal network on the other.
- A flow collector to gather real-time information of flows crossing the network.
- A set of data sources, which can provide information about routing, link utilisation, trusted sites, trusted applications, and known threats.
- An intelligent system that can analyse all the information coming from the different data sources and program the SDN controller managing the network.

CERN openlab has embedded an engineer, Adam Krajewski, in the Brocade development team in charge of designing and implementing new BFO features. In 2015, the enhancements to BFO necessary to meet CERN's requirements were identified, and prioritised. Subsequently, two new versions of BFO were released that partially address CERN's use cases. Both versions include contributions developed by Adam.

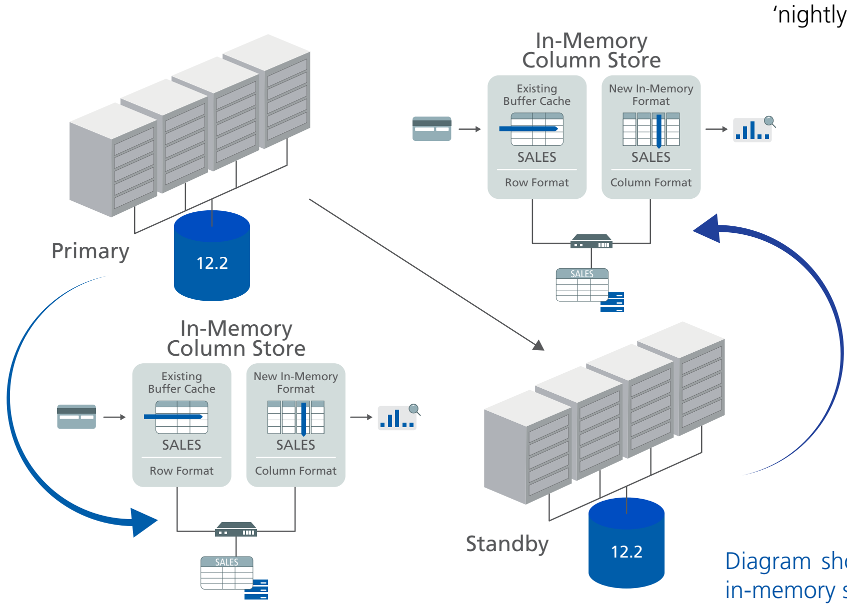
Data Storage Architectures

Every year, the four large-scale LHC experiments create tens of petabytes of data that 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 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

The new Oracle Enterprise Manager 13c was announced in 2015. This offers strengthened ability to manage the Oracle estate in two directions: vertically across the stack, and horizontally — both within and across clouds. A number of new capabilities and features offered by the product have been evaluated in the context of CERN openlab, as part of a beta-testing programme.

Special effort has been dedicated to testing the integration of the new Oracle Enterprise Manager 13c with the CERN Single Sign-On (SSO) system. Work was carried out closely with the Oracle Enterprise Manager security team in order to define the configuration needed to achieve integration with our SSO infrastructure.



At the Oracle OpenWorld event held in October, the beta availability of Oracle Database 12c Release 2 was announced. We are currently evaluating and testing new features from which we could benefit for the CERN deployment. We have put special focus on testing Oracle Grid Infrastructure in order to adopt newly introduced ASM-related requirements to our use cases and existing installation models.

One of the main challenges for database work is to speed up analytical queries while keeping hardware costs under control. The tests of Oracle In-Memory features, therefore, continued — with special focus on in-memory compression. Test results were very promising, both in terms of query performance and in terms of the efficiency of compression. Evaluation of new features added to Oracle In-Memory has continued as a part of the beta testing of Oracle Database 12c Release 2. This work will continue in 2016, with a focus on support for Active Data Guard.

Huawei Cloud Storage Appliance

The evaluation of the Huawei UDS cloud storage appliance reached its final planned stage in 2015. There were two main aspects of this work:

- To complete the study of physics use-cases, the area of end-user analysis was addressed using the S3 protocol from the ROOT analysis framework.
- Deployment of a first prototype S3 service as a backend storage for the CERN VM file system (CernVM-FS) and analysis of its stability and performance as part of an experimental 'nightly build' system.

Diagram showing the architecture of Oracle Database 12.2 with in-memory support.

The evaluation of the Huawei UDS cloud storage appliance reached its final planned stage in 2015.

To address the physics analysis use case, a scalable, distributed I/O generator was constructed. This was deployed within the CERN agile infrastructure and enables creation of repeatable multi-client analysis load patterns, similar to those obtained from real production jobs. The use of ROOT-based client programs and experiment data files in ROOT format has also led to accurate simulation of sparse vector reads, which are typical for end-user analysis. With this distributed load, we have been able to fully utilise the available storage server bandwidth of 2 Gb/s with some 400 clients. The graceful handling of overload conditions with up to 1024 clients was also tested.

In a second study we have applied the CernVM-FS/S3 plugin that was developed earlier in the project and applied it to the larger scale CernVM-FS service scenario of the LHCb 'nightly build' system. Over a period of 110 days, a total repository of 18 million unique software files was created, which represents a daily software release with on average some 170,000 files per day. During this period we evaluated the UDS system for volume- or time-dependent changes in upload- or delete-performance, obtaining satisfactory general stability. More detailed quantitative results are available in a paper presented at the 21st International Conference on Computing in High Energy and Nuclear Physics (Using S3 cloud storage with ROOT and CvmFS, M. Arsuaga Rios, S. Heikkila, et al., CHEP 2015, Okinawa).

Seagate Alternative Storage Architecture 'Kinetic'

The integration of support for Seagate 'Kinetic' network-attached hard-drive technology into the EOS storage system started in March 2015. A one-petabyte Kinetic cluster was installed in the CERN Data Centre in Meyrin in May 2015. It consists of 21 Supermicro 12-drive Kinetic chassis, containing a total of 252 first-generation 'Lombard' Kinetic drives.

In order to enable EOS to support different I/O targets, the EOS I/O path has been modularised. In addition to locally attached hard drives and network-attached Kinetic drives, this enables EOS to, for example, store data using the S3 protocol or use any XrootD-compatible storage system as an I/O endpoint.

The implementation of the Kinetic I/O module is encapsulated in a stand-alone open-source library available on GitHub. The Kinetic I/O module provides clustering functionality for Kinetic drives: data can be replicated or erasure-coded within a cluster, while clusters continue operating gracefully in case of both transient and non-transient drive failures. Data corruption is detected using checksumming and automatically repaired. Very high single-stream performance can be achieved by striping the data of a file to multiple drives of a cluster.

Management of Kinetic clusters has been integrated into the EOS management console. A number of 'EOS Kinetic' commands are supplied to modify cluster configuration, as well as to scan for and repair data corruption or drive errors. A first EOS deployment with Kinetic integration has been set up to provide storage for up

A one-petabyte Kinetic cluster was installed in the CERN Data Centre in Meyrin in May 2015.

to 400 terabytes of genome data in collaboration with the CERN openlab project with the European Bioinformatics Institute (EMBL-EBI). In summary, the integration of network-attached Kinetic hard drives into EOS is functionally complete and provides additional redundancy, robustness, and performance features. It has been deployed to verify stability and performance of the integration.

Comtrade EOS Productisation

The scope of this project — undertaken in collaboration with CERN openlab associate member Comtrade — 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 installation kit includes the necessary instructions, the tools for operations, and a first version of the administration and user guides.

The installation set up the required repositories and enabled instantiation of a fully compliant EOS instance with master/slave head nodes and an arbitrary number of storage nodes. After the setup, an installation summary is displayed summarising the installed resources and a simple functional test is run to check the system is ready to use.

The first beta version was released in March 2015 and EOS and was subsequently tested and installed at several external institutions, including the University of Vienna, INFN Trieste, and the European Commission's Joint Research Centre in Ispra, Italy.

Future efforts will focus on finishing a white paper on the project and closely following the rollout of the system at the external institutions. In the pipeline, there are also plans to run erasure coding across several locations and on distributed environments, embed simple monitoring in the installation kit, and develop an interface for administration, among many other things.

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.

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 was a very busy year for the Java EE project. At the Oracle OpenWorld event in October, Oracle also announced the release of the latest version of Oracle Weblogic (12.2.1). CERN was part of the beta-testing programme for this product, thoroughly investigating and evaluating its new features. A summary of this work was presented during the conference, in collaboration with Oracle.

A complete client for the new Oracle WebLogic Restful Management Services has also been developed, with the aim of simplifying the work of the WebLogic administrators. This client is available on GitHub and an article on this has been published in the summer 2016 issue of the Oracle Scene magazine.

CERN's usage of Oracle APEX is continuously evolving. Our goal is to use Oracle APEX to move from a centrally managed infrastructure to a true 'platform-as–a-service' model, in which the user becomes the owner of the service. This work is described in the 'Oracle APEX in the CERN Java Cloud' article published in the spring issue of the Oracle Scene magazine.

In 2014, our private Java platform-as-a-service — 'middleware on demand' (MWOD) — was put into production. During 2015, work was carried out to consolidate the platform. One of the main improvements was the deployment of the Oracle Java Diagnostics Agent in all the user containers. This agent is helping us to chase those undesirable performance issues in our applications.

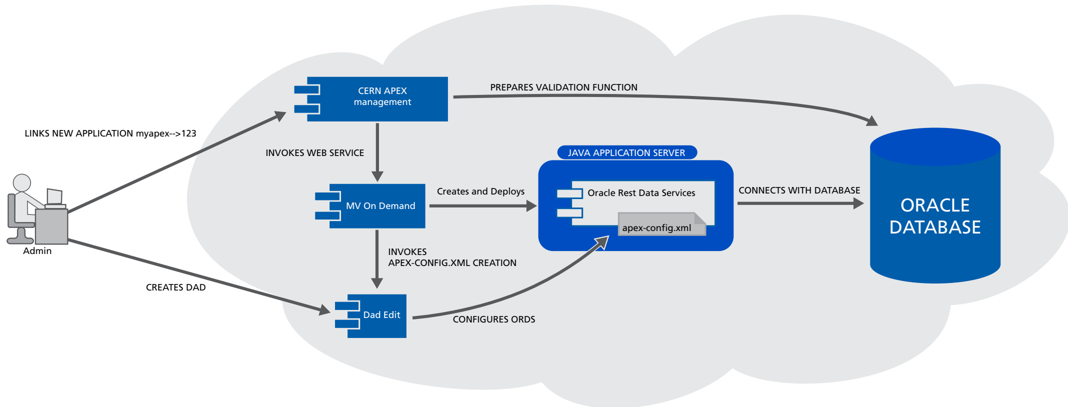
Oracle Database Cloud

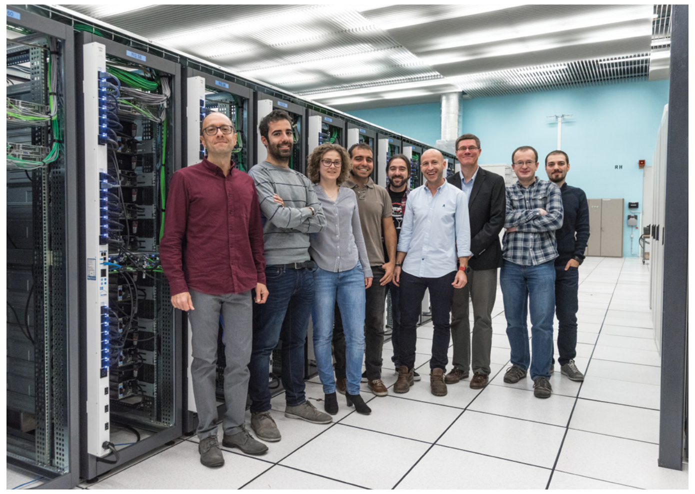
Cloud and virtualisation technologies are important areas of work. What is an optimal infrastructure to run Oracle analytics? How to make it scalable, resilient, and flexible enough to meet all requirements — while also staying within budget? The aim of this project is to investigate how to conceive such platform and to share the experience gained.

The main work carried out in 2015 was the integration of compute nodes with Oracle VM into CERN's OpenStack environment, in order to homogenise infrastructure across all the CERN IT systems.

The next steps are to test the new Oracle OpenStack 2.0 release (based on OpenStack Kilo and available since November) and Oracle VM 3.4, for which we are part of the beta-testing program. Work will also be carried out to integrate compute nodes (Oracle VM) with the new Oracle OpenStack release in the CERN's OpenStack deployment.

Oracle APEX in the CERN Java Cloud.





Members of the CERN IT Department's Database Services Group participated in a number of collaborative projects with Oracle.

The collaboration with Rackspace has focused on cloud federation.

Rackspace Cloud Federation

In 2013, Rackspace and CERN first identified a common need for cloud federation: the ability for a user to log in once and use resources from multiple public and private clouds seamlessly. With contributions from Marek Dennis, the CERN fellow on the project, the Rackspace technical team, and other members of the OpenStack community, this functionality was developed throughout 2015 and released as a standard part of the open-source product. These experiences have been shared at the six-monthly OpenStack summits, which attract over 5,000 attendees.

The federation functions are already in use in the production cloud at CERN, enabling single sign-on to the web portal for authorised users with the EduGain identity federation from hundreds of academic and research organisations. It has also been a key technology supporting collaborations such as the INDIGO DataCloud Horizon 2020 project with 26 diverse partners. Even identities such as Google and Facebook can be used, subject to the cloud security policies. With over 30 companies announcing support for OpenStack federation, there is a bright future for further enhancements by the community following the project's completion.

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's 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. These codes are written in a mix of FORTRAN, C/C++, and Python.

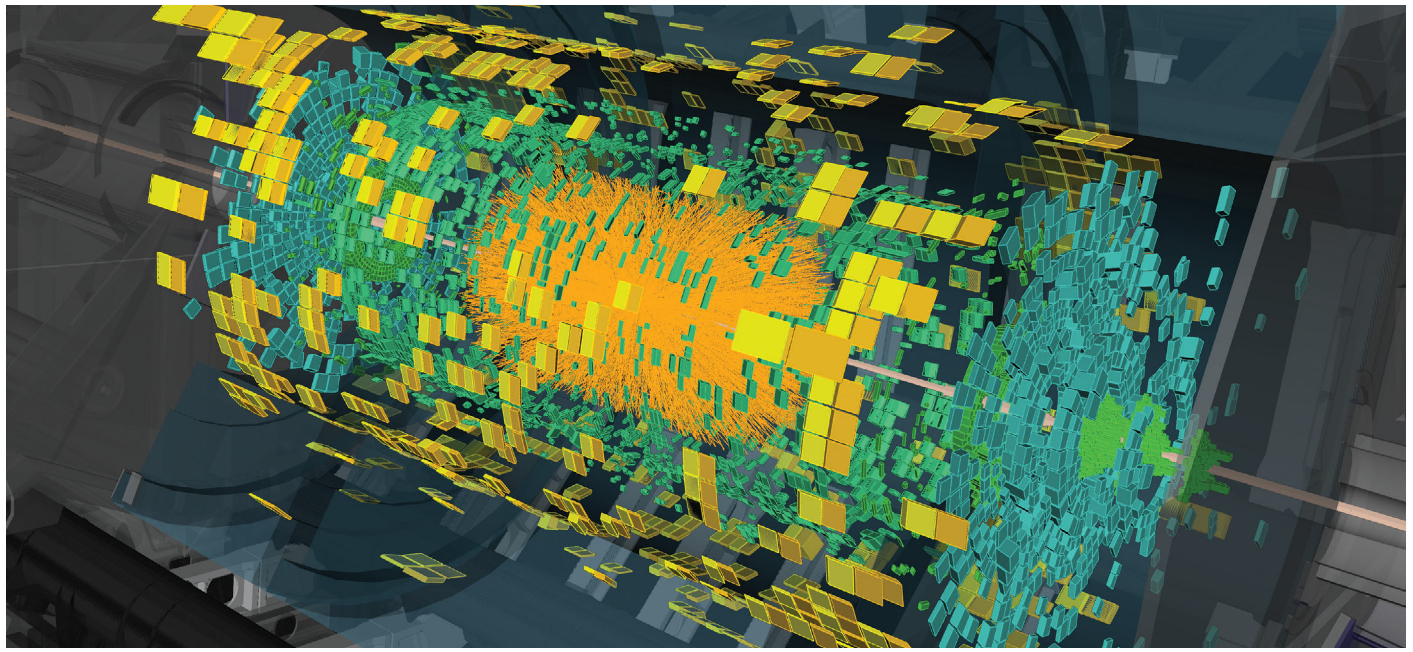
Geant software simulation toolkit

The GeantV work aims to develop the next-generation simulation software used for describing the passage of particles through matter. The goal is to develop and disseminate new approaches in the high-energy physics community for maximising application efficiency through adaptation to modern hardware.

The GeantV concept for vectorisation extends from CPU's with vector units to many-core processors, such as the Intel Xeon Phi architecture, by developing a type-based templated backend interface using the concept of traits. This enables several algorithms — with either scalar or vector inputs — to be expressed with the same code, and has been implemented as an independent library called 'VecCore'. GeantV processes several events at the same time, using multithreading to regroup particles for maximising the cache coherence and favouring vectorisation.

This enables GeantV on several architectures and makes possible use of both generic and device-specific libraries in a transparent way, insulating the core application and algorithms from the technology layer. VecCore is a general solution that can be adopted by any application, requiring maintenance only of a limited set of backends.

One of the first heavy-ion collisions with stable beams recorded by ATLAS in November 2015



Another important achievement of GeantV is the development of vectorised geometry algorithms deployed as the 'VecGeom' library. This not only performed better than existing solutions, but also made its way into the essentially non-vectorised Geant4 and ROOT as a replacement for their native geometries.

A full-scale GeantV-based CMS application using tabulated physics has shown a preliminary speedup of more than three times in single-thread mode compared to its equivalent in Geant4, coming mainly from improved cache coherence and vectorisation. Several new avenues are now being explored to improve this speedup yet further.

FairRoot 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.

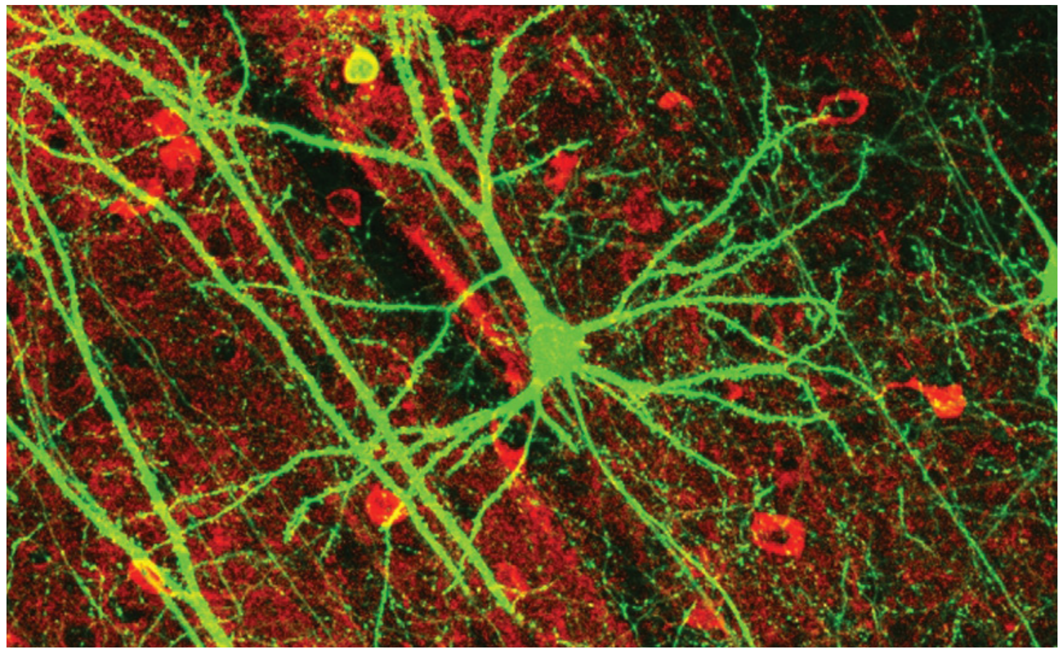
FairRoot is being developed at the GSI Helmholtzzentrum für Schwerionenforschung for the future experiments at the FAIR facility. In FairRoot, the reconstruction algorithms can be implemented as 'tasks' in ROOT, or as separate processes that communicate with each other via message passing. The latter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

In 2015, we started porting some compute-intensive code to Xeon Phi processors and used them via traditional offloading. In this implementation, the message passing is exclusively done on the host code, which passes the data to the Xeon Phi via SCIF protocol. A more general way is under investigation, where we try to extend the ZeroMQ and nanomsg transport engines to support direct transport of data to the processes running on the Xeon Phi.

At the framework level, work to make the dynamic deployment system (DDS) Xeon Phi aware also started. DDS is a toolset that automates and simplifies the deployment of user-defined sets of processes and their dependencies on any available resources. Work to extend the topology properties in DDS to support Xeon Phi code is ongoing.

BioDynaMo: Biology Dynamic Modeller

CERN openlab's mission rests on three pillars: technological investigation, education, and dissemination. The collaboration with research communities and laboratories outside the high-energy physics community brings together all these aspects. As part of the ongoing general effort to develop methods to modernise and optimise software code, the BioDynaMo project was established in September 2015 to transfer ideas, methods, and tools from high-energy physics to the life sciences.



The BioDynaMo project focuses initially on simulating the growth of brain tissue.

BioDynamo is a collaboration between CERN, Newcastle University, Innopolis University, Kazan Federal University, and Intel to design and build a scalable and flexible computing platform for rapid simulation of biological tissue development. It foresees three main phases: (i) the consolidation, optimisation, and further extension of biological simulation code to run efficiently on modern multi-core and many-core platforms; (ii) the deployment of a cloud-based platform using state-of-the-art HPC-on-cloud technologies; (iii) 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 the Cx3D software framework. During the last quarter of 2015, an initial core set of software libraries written in Java have been rewritten in C++ and made ready to be optimised for multi-core CPUs.

Cisco Data-Plane Computing System

The aim of the Cisco Data-Plane Computing System (DPCS) project is to enhance the performance of distributed applications through the elimination of kernel processing in the data path. This is done using the ultra-low latency Ethernet cards from Cisco (usNIC) and the libfabric library developed by the Open Fabric Alliance. The latter enables direct communication between the user-space application and the network interface controller (NIC), bypassing the kernel transmission control protocol (TCP) stack, effectively offloading this work from the CPU.

It is necessary to make this new transport mechanism available to the experimental software without any modification to the latter. The FairMQ abstraction for network communication used by the ALICE experiment's software framework allows us to do this by extending the underlying libraries it uses: ZeroMQ or NanoMsg. After an initial trial with ZeroMQ, the NanoMsg library was chosen for the work due to its clearer API design, which allows modification to the underlying transport mechanism in a straightforward way.
Within a short time we developed a prototype implementation and successfully compiled FairMQ with the new transport. The final, optimised version of the transport was obtained after profiling and improvements in both our implementation and the Cisco provider code for libfabric.

A stable libfabric-based transport in the NanoMsg framework was developed, and published on Github in December. The final version of the transport is stable and is not affected by the performance penalties of the other transports in NanoMsg. That is because we don't rely on the file descriptor-based polling mechanisms of NanoMsg, but on the high-level APIs of libfabric.

As expected, memory copy — especially with big message sizes — is quite expensive, so our main focus has been to achieve zero-copy throughout the message path. NanoMsg had some primitive support for such operations, however we had to extend it in order to be usable by the transport. Our customisations were pushed to the upstream repository and are now awaiting approval.

Not all features of the usNIC driver in libfabric are yet in a stable state, therefore our test case was valuable debugging material for Cisco developers. The issues were reported in the group meetings with Cisco engineers, as well as on the libfabric forum.

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

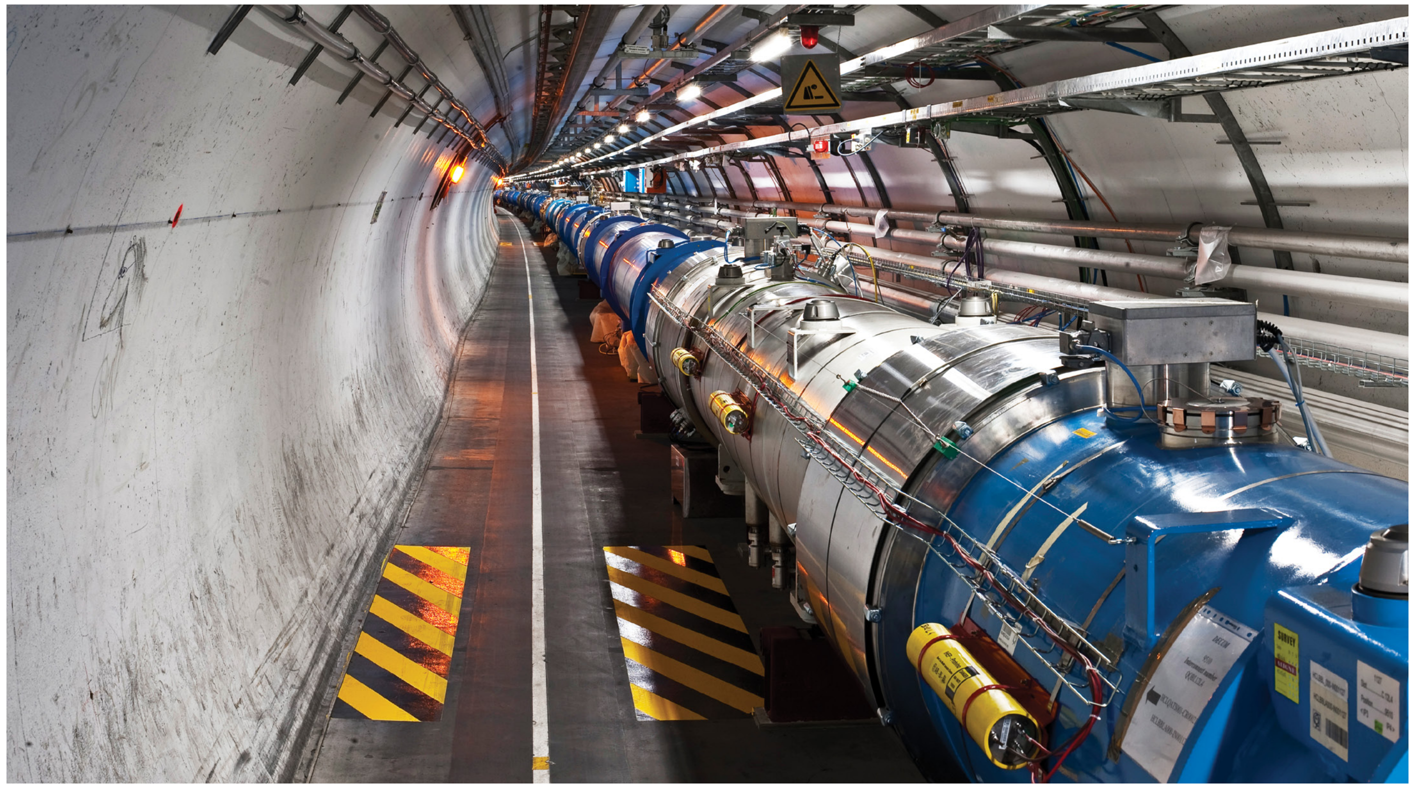
In terms of data analytics, 2015 has been an exciting year. Driven by the increasing needs of new paradigms for fast processing and analysis of large amounts of data, most of our efforts were focused on exploring and implementing solutions to real-life big data problems. Within this context, the so-called 'Hadoop ecosystem' has played an important role.

An example of this is the reliability, availability, and maintainability (RAM) study carried out for the Future Circular Collider (FCC). This study explores industrial approaches to modelling and simulating the reliability and availability of the entire particle accelerator complex. This project has revealed the need for flexible data-management processes and adaptable analytics environments that permit integration of the heterogeneous data sources involved to enable the analysis and exploration of large amounts of data.

One of the challenges with current big-data technologies is to provide an easy-to-use platform through which users with different background domains can interact with the data. A new Oracle technology, Oracle Big Data Discovery, was released last year to address such problems. We implemented a 'proof of concept' for this, and — based on the results obtained from the different parts involved — we are currently preparing its deployment in a production environment.

Another area of focus was hybrid platforms for advanced and streaming analytics. Some of our use cases present response-time restrictions, which cannot compromise the accuracy of the results. The recent integration of Oracle R Advanced Analytic technologies and Oracle Stream Explorer provides a promising platform through which users can write models using R, train them transparently in

Enormous amounts of data are produced by the industrial control systems used in the operation of the LHC.



hundreds of nodes using Hadoop and Spark, and finally deploy the models automatically in Oracle Stream Explorer for real-time scoring of streaming data.

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. In 2015, we worked together with Siemens on ways to handle these large datasets and extract insights that can lead to improved operational efficiency. The work was arranged into two main areas:

- The data flow from WinCC Open Architecture (a SCADA tool widely used at CERN) to a high-performance storage system
- The analysis of the stored data with Siemens analytics tools

In the area of WinCC OA, work was started to make a generic archiver through which one could plug in different systems. We are currently considering the use of Hadoop and Spark as the main technologies. This will enable improved handling of the data produced by our control systems.

For the analysis of the stored data, work was carried out to enhance detection of faulty sensor measurements, to enable better measurement of the performance of control processes (PID), and to develop a new alarm system for flooding detection (as well as identifying the responsible control devices).

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

Our results have shown that most control systems can obtain significant benefits from the use of big-data analytics.

Yandex Data Popularity at LHCb

During 2015, we continued our collaboration with Yandex, and its big-data analytics division, known as ‘Yandex Data Factory’.

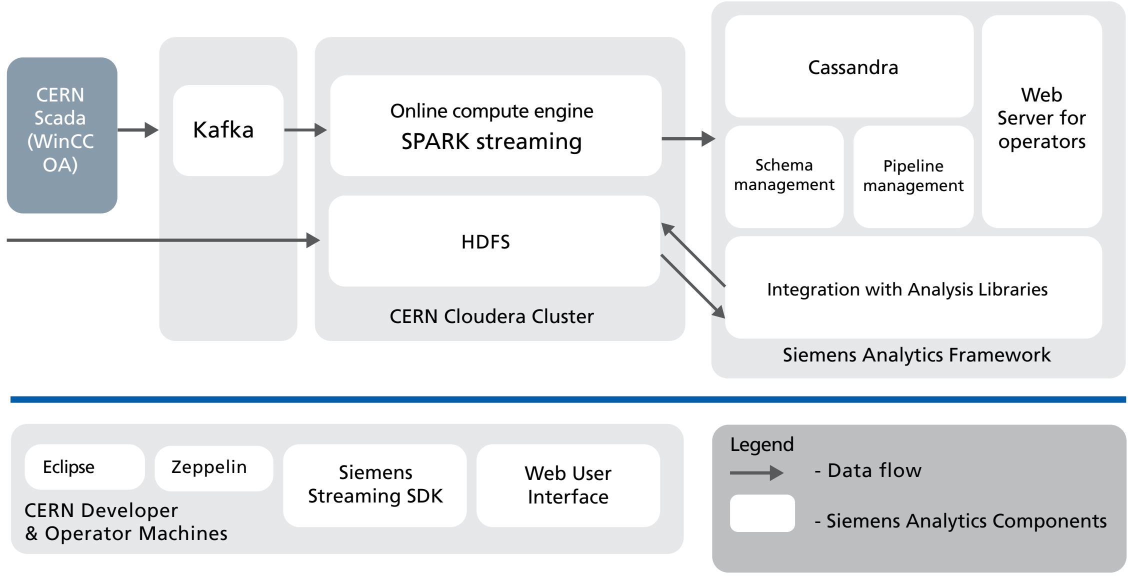
Designing high-performance and cost-effective data-storage systems often means making use of hybrid configurations, in which fast technologies are combined with ones that may be slower but cheaper. An example of such a system is the LHCb grid data storage system, which makes use of both hard-disk drives and magnetic tapes. These hybrid systems require effective data placement and data-movement algorithms to attain maximum performance and cost effectiveness.

A lot of algorithms currently exist for data management in hybrid systems. However, many don’t take workload patterns into account, or apply static assumptions that don’t allow maximum performance to be reached.

In 2015, we therefore designed an algorithm to achieve improved performance over conventional approaches. We made use of Markov Chains, Artificial Neural Networks, Autoregressive–moving-average models, and other machine-learning and time-series-analysis techniques.

This enabled us to produce an algorithm that reduced the total amount of disk space needed by 40%. There are, however, two main drawbacks with this algorithm: it’s difficult to interpret and users have to provide several parameters upfront in order to make use of it. In 2016, we’re therefore planning to simplify the algorithm design and make it much easier to interpret without sacrificing performance.

This project is now set to evolve into a wider scale project that is going to collect statistics from various grid devices and processes. Analysis of collected data will provide insights into bottlenecks and will help with decision making for grid optimisation.



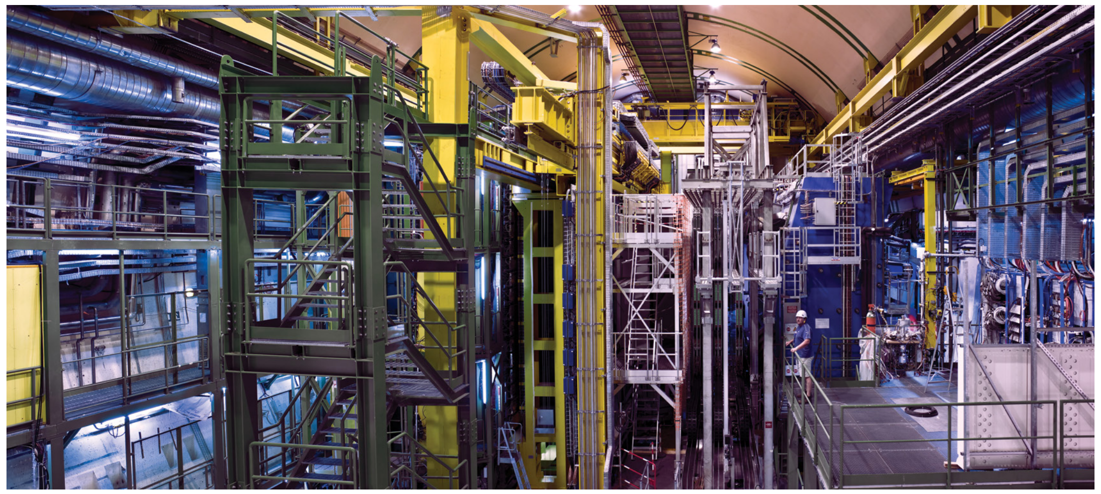
Conceptual design for the use of Siemens analytical tools at CERN.

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 2015, work was carried out to create a novel, autonomous monitoring scheme that doesn’t require references or manual comparison. Taking the high-level physical variables reconstructed by the trigger system, machine-learning techniques are used to check whether their multivariate distribution changes with time. The detector is designed to measure the constant physical processes, so a distribution difference means a difference in detector state: an anomaly.

A pilot version of the anomaly-detection algorithm has been designed. During 2016, we plan to improve performance so it can be used to deal with a wider set of counters, and hence a wider spectrum of anomalies. Our ultimate goal is to be able to find all the anomalies spotted by humans, and even more.



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

Innovation and Entrepreneurship

Discoveries in fundamental physics require constant technological innovation in such diverse fields as computing, electronics, materials science, and industrial sensing and control systems. CERN provides a stimulating environment where talented minds from multiple disciplines can meet to make such innovation happen. As part of CERN openlab's mission to promote innovation, a collaboration with Intel, the CERN Knowledge Transfer Group, and IdeaSquare was launched in July 2015.

Intel Innovation Management and Entrepreneurship Project

The objective of the new innovation and entrepreneurship project is to understand where and how innovation happens across the many different activities run in the laboratory and to help young talents develop new ideas from concepts to possible entrepreneurial ventures. During the second half of the year, experts from CERN openlab and the KT Group have collaborated in a number of innovation and entrepreneurship initiatives, such as the KT Group's monthly entrepreneurship meet-ups and the quarterly 'lightning talk' events in the IT Department. The entrepreneurship meet-ups have been an occasion for young innovators to discuss their ideas with peers and technology experts, and to promote the creation of a wider culture of innovation.

The innovation and entrepreneurship project will continue to run throughout this phase of CERN openlab, with the overarching objective of establishing a permanent process of innovation management, as well as instruments to help motivated engineers and scientists from CERN transition from research to industry.

Innovation meets entrepreneurship at CERN

The project collaborators jointly organised a new innovation and entrepreneurship event at CERN on 26 November. It was attended by over 80 engineers, scientists, students, and start-up enthusiasts. The event featured presentations from invited experts on commercialisation, public-private partnership, intellectual property, and other topics related to innovation and entrepreneurship.

The event proved popular, with over 80 people attending.

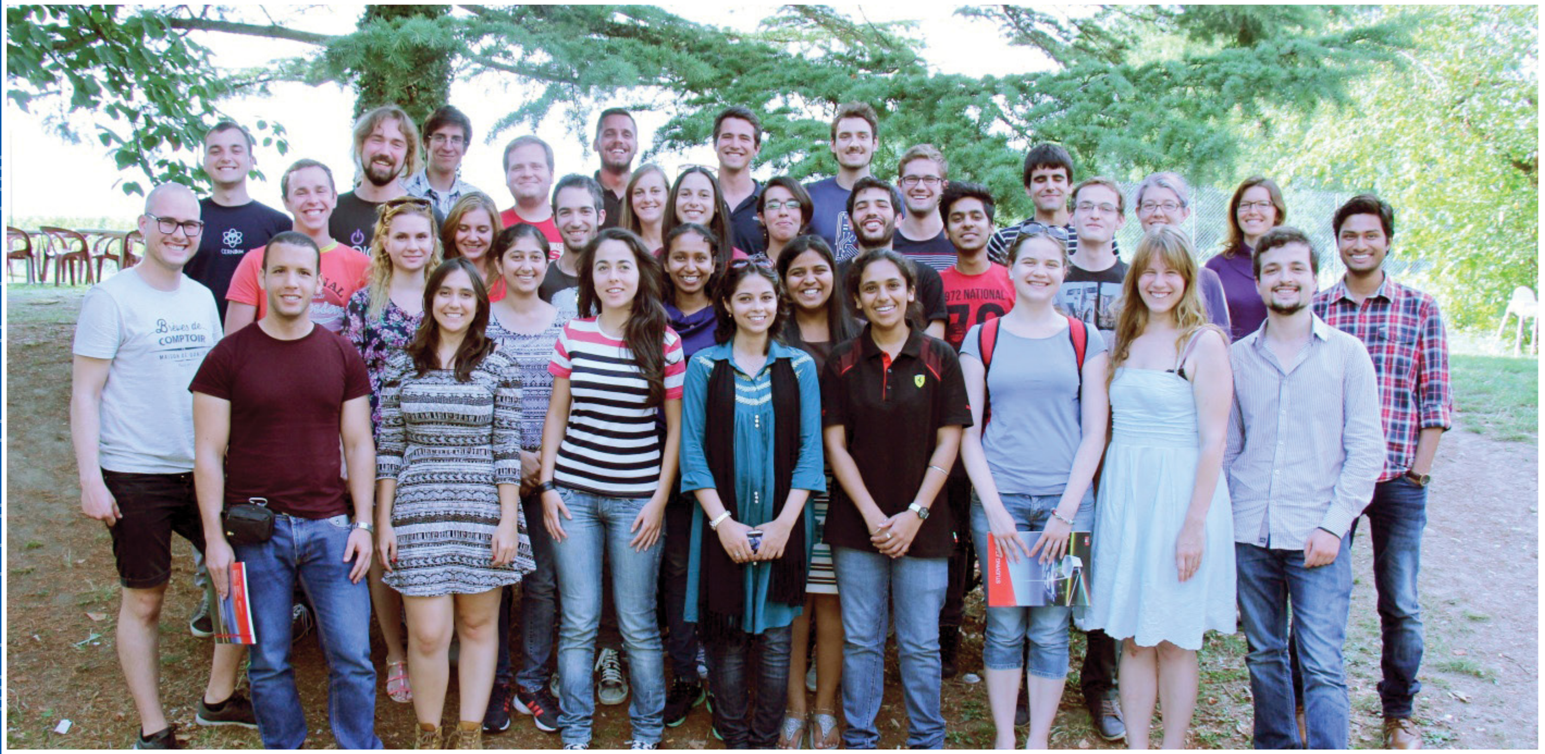
Andrzej Kusmierz, an assistant professor in the strategy department at the Kozminski University in Poland, spoke at the event about the importance of building balanced teams for start-up companies, explaining that a researcher who has an idea must find people to work with who have different, business-related skills. This was echoed by Pawel Bochniarz, who is in charge of innovation and public aid practice at PwC Poland. He stressed the importance of researchers with business ideas building relationships, adding that they must be ready to cede control to others — those with the appropriate skills — when it comes to commercialisation. During his talk, which gave attendees a brief overview of what industry is looking for in partnerships with science, he also spoke of the need for companies of all sizes to innovate.

Ophir Marko, a patent attorney specialising in start-up companies, presented attendees with an overview of intellectual property laws. And Bert Quint, a business development director with a strong background in life sciences, gave attendees an overview of some of the emerging healthcare-related areas in which new start-up businesses could potentially provide innovative solutions. Another perspective was provided by Enrico Giuliani, who is the founder and CEO of a start-up company called Neuron Guard. He spoke of his personal experience as an entrepreneur and he too emphasised the importance of building a good team when starting out.

As part of the event, participants had the opportunity to discuss their own business ideas one-to-one with the experts. In 20-minute one-to-one coaching sessions, the experts provided tailored advice and helped the participants to assess the technical and business feasibility of their proposals.

Expert coaches were on hand to give tailored advice to budding entrepreneurs.

Education



The 40 CERN openlab summer students spent nine weeks at CERN working on ambitious projects using some of the latest hardware and software technologies.

A knowledge factory

CERN openlab is designed to create and disseminate knowledge.

Building human capital

CERN openlab's educational and training work covers several lines of action. 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 CERN, thus exemplifying the CERN openlab principle of two-way knowledge transfer through active collaboration. Many of the workshops combine theory with hands-on practice. Special courses have also been organised for advanced CERN users in specific technical areas. In addition, CERN openlab experts contribute to a range of broader education and training activities.

Student success

The CERN openlab Summer Student Programme continues to go from strength to strength, with 40 participants — representing 27 different nationalities — being selected from around 1500 applicants this year. This educational programme was launched in 2002 to enable undergraduate, masters, and doctoral students to gain hands-on experience with advanced ICT solutions and to learn more about the challenges being addressed through CERN openlab.

The students worked on ambitious projects using some of the latest hardware and software technologies, and were able to see how advanced ICT solutions are used in high-energy physics. 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 laboratories (ETH, EPFL, ESRF) and companies (Google, OpenSystems).

ICE-DIP, the Intel-CERN European Doctorate Industrial Programme, also continued throughout 2015. This Marie Curie Actions project within the European Union's 7th Framework Programme builds on CERN's long-standing relationship with Intel through CERN openlab. It brings together CERN, Intel, and universities to train five early-career researchers. These students are funded by the European Commission and are granted CERN fellowships while enrolled in doctoral programmes at the partner universities, Dublin City University and the National University of Ireland Maynooth. As part of their three-year training programme, the doctoral students are taking part in extended secondments to Intel Labs at locations across Europe. The project is working to develop unparalleled capabilities in the domains of high-throughput, low-latency and online data-acquisition. These topics form a solid foundation for future cutting-edge data-processing technologies of high interest to a broad range of industries. The research themes addressed by ICE-DIP are critical to the sustained forward momentum of European high-energy physics and numerous other disciplines of science and technology.

Competition winners come to CERN

From 14 to 19 June, CERN hosted the 11 young students who won the CERN Special Award at the Intel International Science and Engineering Fair (ISEF) 2015. These winners were selected from the 1700 high-school students who participated in the competition, making it the world's largest pre-university science competition.

Another competition to take place in 2015 was the Intel Modern Code Developer Challenge. Students were tasked with optimising code that is integral to the BioDynaMo project. The goal of the competition 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. Four of the winning students have been invited to come to CERN in 2016.

Dissemination

CERN openlab results have been disseminated at a wide range of international conferences. These publications, presentations, and reports can be consulted on the CERN openlab website. The full list of presentations, publications, posters, and reports for 2015 is available on pages 38 and 39 of this annual report.

CERN openlab Topical Workshops

CERN openlab/Siemens, Identification of Complex Dynamical Systems with Neural Networks, 27-29 April 2015, CERN, Dr. H.-G. Zimmermann/Siemens
CERN openlab/Intel Software Tools Workshop 2015, 9-10 July 2015, Z. Matveev/Intel, G. Zitzlsberger/Intel, H. Pabst/Intel
Future Computing Technology, 28-30 October 2015, Andrzej Nowak (TIK Services, Switzerland)
CERN openlab Technical Workshop, 5-6 November 2015, CERN, A. Di Meglio/CERN, N. Neufeld/CERN, A. Gheata/CERN, A. R. Marin/CERN, L. R. Fernandez/Universidad de Oviedo (ES), L. D. Lupsa/CERN, S. Baymani/CERN, E. Martelli/CERN, M. K. Denis/CERN, P. H. Lensing/Seagate (CH), M. Cattaneo/CERN, P. Buncic/CERN, X. E. Curull/CERN, F. M. Tilaro/CERN, A. R. Purcell/CERN
CERN openlab Innovation and Entrepreneurship Event, 26 November 2015, P. Bochniarz/PwC, Poland, A. Di Meglio/CERN, K. Gawrysiak, E. Guiliani/Neuron Guard, P. Gleissner/Intel, A. Kusmierz/Kozminski University, M. Ophir

The high-school students who won the CERN Special Award at the Intel International Science and Engineering Fair (ISEF) spent a week at CERN in June.

CERN openlab Summer Student Programme Teaching Series, July-August 2015

- Computing in HEP - From Detector to Publication, H. Meinhard
- Simulating Physics and Detectors, J. Apostolakis
- DAQ - Filtering Data from 1 PB/s to 600 MB/s, N. Neufeld
- Computer Security, S. Lopienski
- Advanced Networking, E. Martelli
- Computing on the Grid and in the Cloud, L. Field
- Reconstructing Events, from Electronic Signals to Tracks, M. Elsing
- Data Storage and Data Distribution, A. Pace
- How to write bad code, A. Naumann
- Discovering the Higgs - Finding the Needle in the Haystack, W. Verkerke
- Oral presentation training
- Big Data in Biomedical Sciences, S. Newhouse
- Data Mining, Data Analytics and Big Data Z. Baranowski, A. R. Marin, K. Surdy
- Why SCADA Security is Not like Computer Centre Security, S. Lueders
- Transferring Knowledge from CERN to Industry, N. Ziogas

CERN openlab Summer Students 2015, with Nation State, Home Institute and Project Topic

- J. Aguilar Mena, Cuba, Higher Institute of Technologies and Applied Sciences (InSTEC), Cuba, "Kalman Filter Improves Using GPGPU and Autovectorization for Online LHCb Triggers"
- M. Akuruyejo, Nigeria, University of Lagos, Nigeria, "Continuous Integrated Testing of Oracle Databases on CERN Agile Infrastructure"
- S. Amrouche, Algeria, Ecole Nationale Supérieure d'Informatique, Algeria, "Evaluation of Apache Mesos for Data Analytics"
- G. Arnau Antoniucci, Spain, Universitat Politècnica de Catalunya, Spain, "FLUKA Pre-Optimizer for a Monte Carlo Treatment Planning System"
- G. Azzopardi, Malta, University of Malta, Malta, "Statistical Reports and Data Analytics with Distributed Computing"
- A. Bose, India, International Institute of Information Technology, India, "CMS Data-Services Ingestion into CERN's Hadoop Big Data Analytics Infrastructure"
- L. Breitwieser, Austria, Graz University of Technology, Austria, "Porting a Java-based Brain Simulation Software to C++"
- A. L. Brisighello Filho, Brazil, UNICAMP – Universidade Estadual de Campinas, Brazil, "Using Docker to Execute Offline Instances of CMSSW"
- Ł. Chrzaszcz, Poland, AGH University of Science and Technology, Poland, "Photo Albums in Invenio 2.x"
- J. Costa, Portugal, Universidade do Porto, Portugal, "RapidIO Usage in a Big Data Environment"
- J. Delgado Fernandez, Spain, Universidad de Oviedo, Spain, "Processing of the WLCG job monitoring data using ElasticSearch"
- J. Domšić, Croatia, University of Zagreb, Croatia, "Tracing and Accounting of Physical Resources in the Computer Centre"
- P. Doncheva, Bulgaria, FDIBA - Technical University, Bulgaria, "Flow Visualizer"
- S. Ganju, India, National Institute of Technology Hamirpur, India, "Evaluation of Apache Spark as Analytics as framework for CERN's Big Data Analytics"
- I. Grigorescu, Romania, University Politehnica of Bucharest, Romania, "Batch Monitoring Dashboards"
- E. Guiraud, Italy, Università degli Studi di Milano, Italy, "Enhancements to Multiprocessing in ROOT"
- S. Gupta, India, University of Petroleum and Energy Studies, India, "Non-Intrusive User Interaction Monitoring for WinCC OA based Applications"
- T. Kappé, the Netherlands, Universiteit Leiden, the Netherlands, "Towards EOS Namespace Persistence in NVRAM: Hashtable Benchmarks"
- E. Katsomallos, Greece, University of Thessaly, Greece, "Combining Grid, Cloud and Volunteering Computing"
- L. Kojo, Finland, Tampere University of Technology, Finland, "Monitoring Commercial Cloud Service Providers"
- J. Kvita, the Czech Republic, Vysoke Uceni Technicke v Brne, the Czech Republic, "Archiving OpenStack Cloud Volumes"
- S. Miras, Greece, Technological Educational Institute of Peloponesse, Greece, "Source Code Review Using Static Analysis Tools"
- A. Nadeem, Pakistan, NUST School of Electrical Engineering and Computer Science, Pakistan, "Evaluation and Implementation of SQRL and U2F as 2nd Factor Authenticators for CERN Single Sign On"
- R. Nevatia, India, Sardar Patel Institute of Technology, Mumbai University, India, "Intelligent Workload Management across Database Replicas"
- C. Newey, United Kingdom, Aberystywth University, United Kingdom, "Streamlining Infrastructure Monitoring and Metrics in IT- DB-IMS"
- P. Pamula, Poland, Wroclaw University of Technology, Poland, "Keystone Identity Service in Openstack"
- M. Patrascoiu, Romania, University Politehnica of Timisoara, Romania, "Manila – OpenStack File Sharing Service"
- I. Pejeva, former Yugoslav Republic of Macedonia, University of Goce Delcev – Shtip, former Yugoslav Republic of Macedonia, "Evaluating the Performance of Seagate Kinetic Drive Technology and its Integration into the CERN EOS Storage System"
- M. Pinkute, Lithuania, The University of Edinburgh, United Kingdom, "APEX Backup Tool Development"
- A. M. Popescu, Romania, Politechnica University of Timisoara, Romania, "Vectorisation and GPUs Extensions of ROOT::Math Routines"
- M. Potter, United States of America, Simmons College, United States of America, "Zenodo Information Architecture and Usability"
- M. Priisalu, Estonia, Lund University, Sweden, "Aggregating Labels in Crowdsourcing Data"
- C. Quast, Germany, Technical University of Berlin, Germany, "Implementing the Network Debugging Infrastructure for the New Detector Readout Board"
- S. Rai, India, University of Petroleum and Energy Studies, India, "Dynamic Web Infrastructure Visualisation Tool + Analysis of "Performance Demo Issues" webapp with tomcat and JMeter"
- S. Seetharaman, India, Amrita School of Engineering, Amritapuri, India, "Upgrading the Huawei Cloud Storage Benchmark Framework for ROOT6 Compatibility"
- E. Serrano, Spain, Universidad Carlos III de Madrid, Spain, "Optimization Studies of Millepede - a Detector Alignment Application"
- H. Urhan, Turkey, Fatih Sultan Mehmet Vakif University, Turkey, "Evaluation of Oracle Big Data Integration Tools"
- T. Van Steenkiste, Belgium, Ghent University, Belgium, "Integration of Key-Value Database in the Parameter

Daniel Vea (second from left) and Pablo Gonzalez De Aledo Marugan (second from right), two of the first-place winners of the Intel Modern Code Challenge, pictured with their partners and Russel Beutler of Intel (far left) during their visit to CERN.

Management System of the ALFA Framework"
- J. Vicente Cantero, Spain, Jaume I University, Spain, "Building a Real-time Notification System"
- H. Zafar, Pakistan, NUST School of Electrical Engineering and Computer Sciences, Pakistan, "Improving File Transfer Service FTS3"

CERN openlab Presentations and Publications

Presentations

- G. Jereczek/CERN, CERN European Laboratory for Particle Physics, Presentation of CERN to Polish High School Students, I LO Kartuzy, Kartuzy, Poland, 9 January 2015
- A. R. Marín/CERN, Real-Time Analytics: Making Better and Faster Business Decisions, London, UK, 29 January 2015
- A. R. Marín/CERN, CERN Data Analytics Project, India Analytics & Big Data Summit 2015, Mumbai, India, 3 February 2015
- A. R. Marín/CERN, CERN Data Analytics Project, India Analytics & Big Data Summit 2015, Bangalore, India, 5 February 2015
- M. Zeiler/CERN, Principles of Mach-Zehnder modulators, Presentation at CERN opto team meeting, CERN, 4 March 2015
- M. Zeiler/CERN, How to use PhoeniX software for photonic chip design, Presentation at CERN opto team meeting, CERN, 4 March 2015
- M. M. Marquez/CERN, Information Discovery as a Service: Better Decisions for more Efficient Operations at CERN, German Oracle Business Intelligence and Data Warehousing Conference, Hannover, Germany, 10 March 2015
- A. R. Marín/CERN, Big Data Analytics for the Exploitation of the CERN Accelerator Complex, Milan, Italy, 11 March 2015
- A. Di Meglio/CERN, Big Data Challenges - Power and Responsibilities, HEPTech Conference on Big Data, Budapest, Hungary, 30 March 2015
- T. Bell/CERN, Cloud Computing Infrastructure at CERN, HEPTech Conference on Big Data, Budapest, Hungary, 30 March 2015
- S. Sridharan/CERN, Evaluation of 'OpenCL for FPGA' for DAQ and Acceleration in HEP applications, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015), Okinawa, Japan, 13 April 2015
- M. Arsuaga-Rios/CERN, S. Heikkila/CERN, D. Duellmann/CERN, J. Blomer/CERN, R. Meusel/CERN, B. Couturier/CERN, Using S3 cloud storage with ROOT and CernVMFS, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015), Okinawa, Japan, 14 April 2015
- M. Denis/CERN, Cloud Federation Prelude to Hybrid Clouds, 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015), Okinawa, Japan, 16 April 2015
- M. M. Marquez, Data Science for Improving CERN's Accelerator Complex Control Systems, International Conference on Computing in High Energy and Nuclear Physics (CHEP) 2015, Okinawa, Japan, 16 April 2015
- A. Santogidis/CERN, Optimizing the Transport Layer of ALFA on the Intel Xeon Phi Coprocessor, Okinawa Institute of Science and Technology (OIST), Okinawa, Japan, 17 April 2015
- A. Santogidis/CERN, The ICE-DIP project at CERN and my role, System and Media Laboratory (SyMLab), HKUST (Hong Kong University of Science and Technology), Hon Kong, China, 22 April 2015
- L. L. Pardavila/CERN, Oracle GoldenGate at CERN, 12c Event, Baden, Switzerland, 28 April 2015
- M. M. Marquez/CERN, Information Discovery as a Service: Better Decisions for more Efficient Operations at CERN, Rittman Mead BI Forum, Brighton, UK, 8 May 2015
- M. Denis/CERN, New Advances in Federated Identity and Federated Service Provider Support for OpenStack Clouds, OpenStack Summit, Vancouver, BC, Canada, 19 May 2015
- A. Santogidis/CERN, High Performance I/O with the ZeroMQ (ØMQ) Messaging Library, CERN School of Computing, Geneva, Switzerland, 22 May 2015
- M. M. Marquez/CERN, Big Data & Analytics: Improving the Control of CERN's accelerator Complex by Data-Driven Decisions, Pharma Exabyte, Berlin, Germany, 27 May 2015
- A. Di Meglio/CERN, CERN Science and Technology, CIOCity Conference, Brussels, Belgium, 4 June 2015
- T. Bell/CERN, Clouds and Research Collide at CERN, Rackspace: Solve Global Conference Series, London, UK, 4 June 2015
- A. R. Marín/CERN, Information Discovery, Exploring Data-Driven Decision-Making for Improving the Control of CERN's Accelerator Complex, 16th Swiss Big Data User Group Meeting, Zurich, Switzerland, 8 June 2015
- P. Hermann Lensing/Seagate, J. Hughes/Seagate, Seagate: Expanding CERN's EOS Storage with Seagate Kinetic Disk Technology, CERN openlab Open Day, CERN, 10 June 2015
- E. Bakany/Siemens, F. Maria Tilaro/CERN, Siemens: Smart Technologies for Large Control Systems, CERN openlab Open Day, CERN, 10 June 2015
- S. Newhouse/EMBL-EBI, European Bioinformatics Institute: Research Infrastructure Needed for Life Science, CERN openlab Open Day, 10 June 2015
- M. Hanif/Brocade, Brocade: Optimal Flow Placement in SDN Networks, CERN openlab Open Day, CERN, 10 June 2015
- K. Leong Yong/DSI, S. Ruocco/DSI, Data Storage Institute: Speeding-up Large-Scale Storage with Non-Volatile Memory, CERN openlab Open Day, CERN, 10 June 2015
- G. Fox/Rackspace, Rackspace: Significance of Cloud Computing to CERN, CERN openlab Open Day, CERN, 10 June 2015
- A. Gheata/CERN, Intel: GeantV - Taking Up the Technology Challenge, CERN openlab Open Day, CERN, 10 June 2015
- D. Paul/IDT, IDT: RapidIO Interconnect based Low Latency Multi Processor Heterogeneous Solutions, CERN openlab Open Day, CERN, 10 June 2015
- M. Al Turany/GSI, ALFA: Next generation Concurrent Framework for ALICE and FAIR Experiments, CERN openlab Open Day, CERN, 10 June 2015
- K. Jernigan/Oracle, L. Lobato Pardavila/CERN, M. Martin Marquez/CERN, Oracle: Database and Data Management Innovations with CERN, CERN openlab Open Day, CERN, 10 June 2015
- Z. Susnjar/Cisco, Cisco: Next Generation Operating Systems, CERN openlab Open Day, CERN, 10 June 2015
- N.Neufeld/CERN, Intel: High Throughput Computing Collaboration: _A CERN openlab / Intel Collaboration, CERN openlab Open Day, CERN, 10 June 2015
- A. Di Meglio/CERN, CERN openlab a Successful Public-Private Partnership, CERN openlab Open Day, CERN, 10 June 2015
- A. Di Meglio/CERN, Core Skills for the digital economy, The Helix Nebula Initiative & PICSE: Towards a European Open Science Cloud, CERN, 26 June 2015
- S. Stancu/CERN, Brocade Flow Optimizer, Layer123 SDN & OpenFlow World Congress, Düsseldorf, Germany, 15 October 2015
- A. Di Meglio/CERN, Big and Smart Data Analytics, Medical Record Update Workshop, AdriHealthMob project, Pesaro, Italy, 20 October 2015
- A. Wiecek/CERN, Oracle WebLogic Server 12.2.1 Mul5tenancy, Oracle Open World 25-29 October 2015, San Francisko, USA, 27 October 2015
- G. Jereczek/CERN, A Lossless Switch for Data Acquisition Networks, IEEE Conference on Local Computer Networks 2015, Clearwater Beach, Florida, United States Of America, 28 October 2015
- A. Gheata/CERN, Porting LHC detector simulations on Intel Xeon and Xeon Phi architectures, Intel HPC Developers Conference - Austin, TX, 15 November 2015

Publications

- The LHCb collaboration, R. Aaij, B. Adeva, M. Adinolfi, A. Affolder, Z. Ajaltouni, S. Akar, J. Albrecht, F. Alessio and 693 more, Search for the lepton flavour violating decay ć– ćμ – μ + μ –, link.springer.com - Journal of High Energy Physics, 18 February 2015
- A. Rogozhnikov/Yandex, A. Bukva/Faculty of Physics, Belgrade, Serbia, V. Gligorov/CERN, A. Ustyuzhanin/Yandex, M. Williams/ Massachusetts Institute of Technology, Cambridge, MA, U.S.A., New Approaches for Boosting to Uniformity, IOPscience.iop.org - Journal of Instrumentation, Volume 10, 30 March 2015
- L. R. Fernandez/CERN, Oracle APEX in the CERN Java Cloud, OracleScene magazine, issue 56, 13 April 2015
- J. Apostolakis/CERN, M. Bandieramonte/University of Catania and INAF (IT), G. Bitzes/University of Athens (GR), R. Brun/CERN, P. Canal/Fermi National Accelerator Laboratory (US), F. Carminati/CERN, J. C. De Fine Licht/ University of Copenhagen (DK), L. Duhem/Intel Corporation, V. D. Elvira/Fermi National Accelerator Laboratory (US), A. Gheata/CERN, S. Y. Jun/Fermi National Accelerator Laboratory (US), G. Lima/ Fermi National Accelerator Laboratory (US), M. Novak/CERN, R. Sehgal/Bhabha Atomic Research Center (IN), O. Shadura/ National Technical University of Ukraine, "Kyiv" Politechnic Institute, S. Wenzel/CERN, Adaptive track scheduling to optimize concurrency and vectorization in GeantV, IOPscience 16th International workshop on Advanced Computing and Analysis Techniques in physics research (ACAT, Prague, Czesh Rebublic), 22 May 2015
- G. Jereczek/CERN, G. Lehmann Miotto/CERN, David Malone/Maynooth University, Maynooth, Irleand, Analogues between tuning TCP for Data Acquisition and Datacenter Networks, IEEE International Conference on Communications 2015, London, United Kingdom, 8 - 12 June 2015
- T. Likhomanenko/Yandex, A. Rogozhnikov/Yandex, A. Baranov/Yandex, E. Khairullin/Yandex, A. Ustyuzhanin/Yandex, Improving Reproducibility of Data Science Experiments, AutoML 2015 workshop @ ICML 2015, Lille, France, 11 July 2015
- G. Jereczek/CERN, G. Lehmann Miotto/CERN, David Malone/Maynooth University, Maynooth, Irleand, M. Walukiewicz/Intel Corp. Gdansk, Poland, A Lossless Switch for Data Acquisition Networks, IEEE Conference on Local Computer Networks 2015, Clearwater Beach, Florida, United States Of America, 29 July 2015
- M. Zeiler/CERN, CERN's LHC Explores Silicon Photonics for Data links, SPIE professional Magazine, October 2015
- A. Voitier/CERN, M. Gonzalez-Berges/CERN, F.M. Tilaro/CERN, M. Roshchin/Siemens, Formalizing Expert Knowledge in Order to Analyse CERN's Control Systems, The International Conference on Accelerator and Large Experimental Physics Control Systems, 23 October 2015
- S. Sridharan, P. Durante, N. Neufeld, D. Campora, Evaluation of 'OpenCL for FPGA' for Acceleration and DAQ in HEP LHCB, LHCB Technical Note, CERN, 02 November 2015
- A. Santogidis/CERN, A. Hirstius/Maynooth University, Intel GmbH Germany, S. Lalis/University of Thessaly, Evaluating the transport layer of the ALFA framework for the Intel[Symbol] Xeon PhiTM Coprocessor, iopscience.iop.org - 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015), 13 November 2015
- T. Likhomanenko/Yandex, A. Rogozhnikov/Yandex, A. Baranov/Yandex, E. Khairullin/Yandex, A. Ustyuzhanin/Yandex, Reproducible Experiment Platform, iopscience.iop.org - 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015), 13 November 2015

Posters

- A. Santogidis/CERN, Optimizing the Transport Layer of the Alfa Framework for the Intel Xeon Phi coprocessor, International Conference on Computing in High Energy and Nuclear Physics (CHEP), Okinawa, Japan, 13 April 2015
- M. M. Marquez/CERN, Big Data Analytics as a Service Infrastructure, International Conference on Computing in High Energy and Nuclear Physics (CHEP) 2015, Okinawa, Japan, 15 April 2015.
- M. Arsuaga-Rios/CERN, S. Heikkila/CERN, D. Duellmann/CERN, Evaluation of the Huawei UDS Cloud Storage Systems for HEP Applications, CERN openlab Open Day, CERN, 10 June 2015
- F. Tilaro/CERN, A. Voitier/CERN, M. Gonzalez Berges/CERN, Smart Data Analysis of CERN Control Systems, CERN openlab Open Day, CERN, 10 June 2015
- L. L. Pardavila/CERN, Evolution of Database Replication Technologies for WLCG, CERN openlab Open Day, CERN, 10 June 2015
- C. Färber/CERN, K. Ha/CERN, N. Neufeld/CERN, R. Schwemmer/CERN, P. Durante/CERN, Feasibility Study to Use Intel Xeon Phi and Xeon-FPGA Computing Accelerators with the Omni-Path Network in the Event Filter Farm for the LHCb Upgrade, CERN openlab Open Day, CERN, 10 June 2015
- P. Szostek/CERN, Evaluation of Power Efficiency and Performance of Multi-core Platforms Using HEP Workloads, CERN openlab Open Day, CERN, 10 June 2015
- P. Fiala/CERN, P. Golonka/CERN, M. Gonzalez Berges/CERN, K. Szkudlarek/CERN, F. Varela/CERN, Success Story: CERN –Siemens/ETM Collaboration to Develop WinCCOpen Architecture, CERN openlab Open Day, CERN, 10 June 2015
- M. Denis/CERN, Scaling beyond your cloud Hybrid OpenStack Clouds , CERN openlab Open day, CERN, 10 June 2015
- G. Jereczek/CERN, Analogues between tuning TCP for Data Acquisition and Datacenter Networks, IEEE International Conference on Communications 2015, London, United Kingdom, 8 - 12 June 2015
- A. Gheata/CERN, G. Amadio/UNESP, São Paulo, C. de Paula Bianchini/UNESP, São Paulo, Mackenzie Presbyterian University, F. Carminati/CERN, S. Vallecorsa/CERN, S. Wenzel/CERN, GeantV - Next Generation Simulation Prototype, SC15, Austin TX, 17 November 2015

Sverre Jarp, former CTO of CERN openlab, presenting in CERN's main auditorium.

Events and Outreach

Creating and disseminating knowledge

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

As well as the excellent technical results that CERN openlab provides, the collaboration gives CERN a means to share a vision of the future of scientific computing with its collaborators, through joint workshops and events, as well as to disseminate this to a wider audience, including partner clients, members of the press, and the general public.

Top delegations from governments and industry regularly tour CERN. In 2015, 126 protocol visits were organised at CERN. 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'. In addition, the general public is also widely introduced to the CERN openlab activities, including through our first ever CERN openlab 'open day'. More than 500 press articles were published about our work over the course of the year, with 2015 also seeing the launch of our dedicated CERN openlab social media channels.

On the CERN openlab website, you can find news about our activities in 2015, including workshops and seminars — many of which have been recorded and are available for viewing. On the site, you can also find a complete collection of press coverage, 'sponsor spotlights', press releases, and other case studies from collaborators. **cern.ch/openlab**

Rolf Heuer, Director General of CERN from 2009 to 2015, addresses CERN openlab collaborators.

Siemens VIP visit to CERN

13 January 2015

Visit of Bernard Reichl, CEO of ETM, and Wolfgang Hesoun, CEO of Siemens Austria, to the CMS experiment.

Siemens press and VIP visit to CERN

5 February 2015

Visit of Wolfgang Heuring, head of Siemens Research and Technology Centre; Thomas Hahn, head of software development at Siemens Corporate Technology; Benno Estermann from the Siemens Press office; and members of the German and Swiss press.

Intel customer visit to CERN

28 May 2015

Visit of Intel customers to CERN for a tour of the laboratory's facilities.

ISEF 2015 students visit CERN

14-19 June 2015

The 11 winners of the 'CERN special award' came to CERN for an extended tour and educational programme.

Google interns visit CERN

3 August 2015

48 interns from Google's Zurich office visited CERN for a tour of the laboratory's facilities.

Australian Bureau of Meteorology visit

18 September 2015

Lesley Seebeck, CIO of the Australian Bureau of Meteorology, visited CERN for discussions on data storage, distributed computing, and industrial collaboration.

CERN openlab visit to ILL and ESRF

23 September 2015

Members of the CERN openlab team and the CERN IT Department visited the Institut Laue-Langevin (ILL) and the European Synchrotron Radiation Facility (ESRF) in Grenoble, France.

Sergio Bertolucci, CERN's Director for Research and Computing from 2009 to 2015, speaks at the CERN openlab Open Day.

Seminars by CERN openlab guests

The Science of Prediction with Neutral Networks

29 April 2015

H.G. Zimmermann of Siemens.

CERN openlab/Intel in Big Data and IoT

30 April 2015

Parviz Peiravi of Intel.

Unlock performance secrets of next-gen Intel hardware

12 May 2015

Zakhar A. Matveev of Intel.

Intel OmniPath

4 June 2015

Representatives of the LHC experiments.

Corporate big data and analytics, an innovation journey

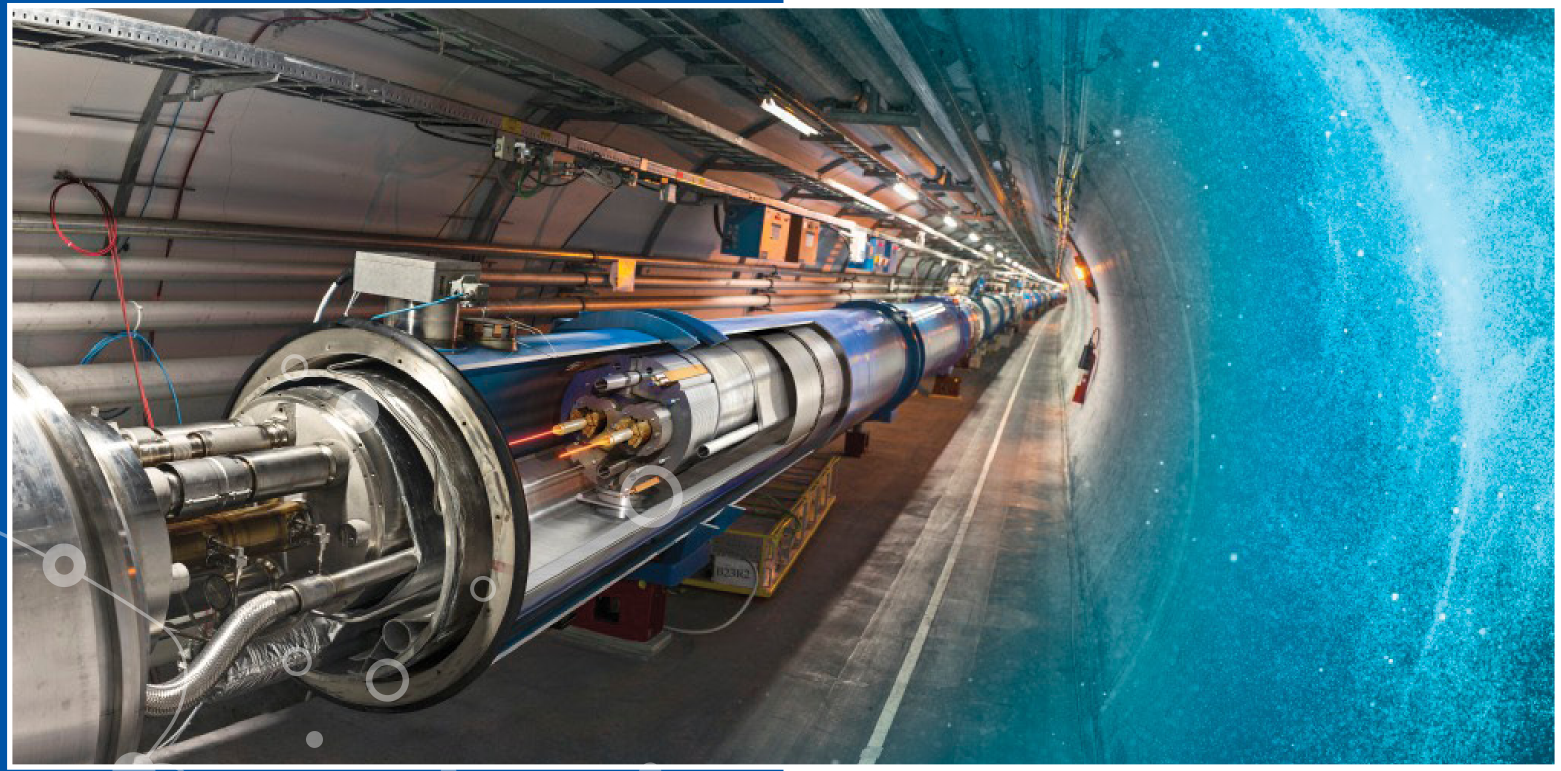
29 June 2015

Massimo Mercuri of DSM.

Balanced Technical Computing: Solving Big Data problems a novel way

2 December 2015

Landon Noll and Rene Raeber of Cisco Systems Inc.



The LHC's ambitious upgrade programme has already seen the collision energy almost double since Run 1.

The Future

The next decade and beyond

CERN openlab looks to support the LHC research community as the accelerators and the experiments are upgraded over the next decade.

In 2016, the LHC will generate up to 1 billion collisions per second in the experiments, producing a torrent of data. The operators will increase the number of particles circulating in the machine and the squeezing of the beams in the collision regions. The LHC will also continue to run at a collision energy of 13TeV, almost double the collision energy of Run 1.

In 2016, the ATLAS and CMS collaborations — who announced the discovery of the Higgs boson in 2012 — will study this particle in depth. But there are still several questions that remain unanswered by the Standard Model, such as why nature prefers matter to antimatter, and what dark matter consists of, despite it potentially making up one quarter of our universe.

The huge amounts of data from the 2016 LHC run will enable physicists to challenge these and many other questions, to probe the Standard Model further, and to possibly find clues about the physics that lies beyond it.

CERN openlab is working to support the LHC research community in 2016 and beyond, with a particular focus on work that will be carried out during the planned upgrade phases that are scheduled to take place over the next decade. With the data rates from the experiments set to increase significantly, we're focusing on supporting the work to overhaul and modernise their data-acquisition systems, while also ensuring that the maximum benefits are gained from the hardware available to CERN's teams by making sure the software running on it has been fully optimised.

In 2016, some of our specific areas of focus include the following, application of high-throughput computing technologies to data acquisition and reduction, investigation of tools to enable the deployment of software-defined networks to improve security and robustness, modernisation of the popular Geant software used to simulate the passage of particles through matter, and further expansion of our work to distribute and replicate data across the WLCG network. We're also beginning work to investigate emerging data-analytics and machine-learning tools — both those coming from industry and from open-source communities — and to identify specific experimental use cases for testing these tools, so that concrete activities can be carried out in this area in the second half of CERN openlab's fifth phase.

CERN openlab 2015 Annual Report
Editor: Andrew Purcell
Design and Production: CERN Graphic Design Service

All photos provided by the CERN Audio-Visual Production Service, except:
Christian Beutler (page 5), Veronika McQuade (page 17), LHCb collaboration (pages 19 and 31), Sylvain Fasel (page 20), ATLAS Collaboration (page 26), Wei-Chung *et al.* (2005)/PLOS Biology (CC BY 2.5) (page 27), CERN openlab communications office (pages 32, 33, 34, 36, 37, 40, 42, 43).

Graphic used on front cover (and throughout the report) courtesy of Shutterstock.

An electronic version of this report is available on the CERN openlab website:
www.cern.ch/openlab/resources/annual-reports

ISBN 978-92-9083-435-9
Published by CERN

CERN openlab

www.cern.ch/openlab