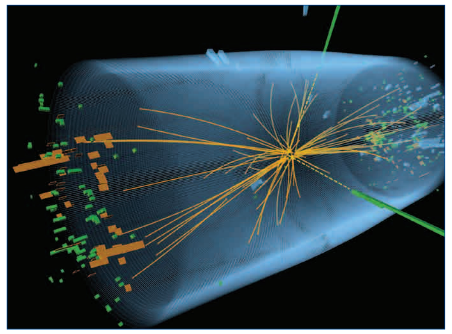
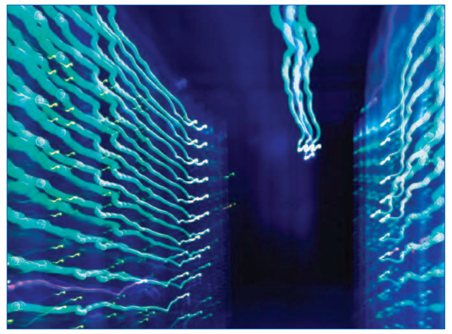
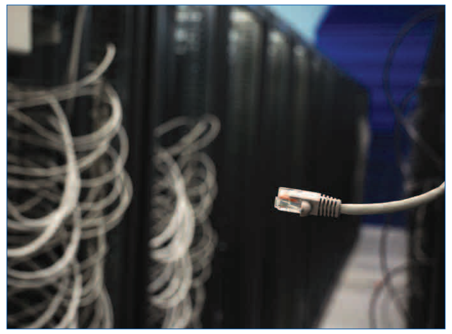
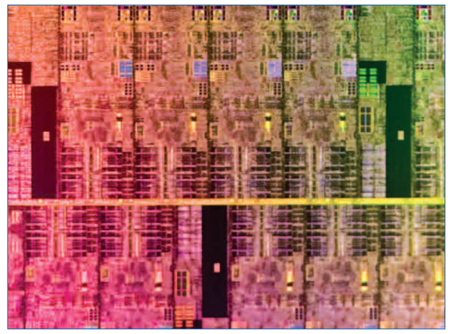
Annual Report 2012

CERN openlab



Contents

A Word from the DG

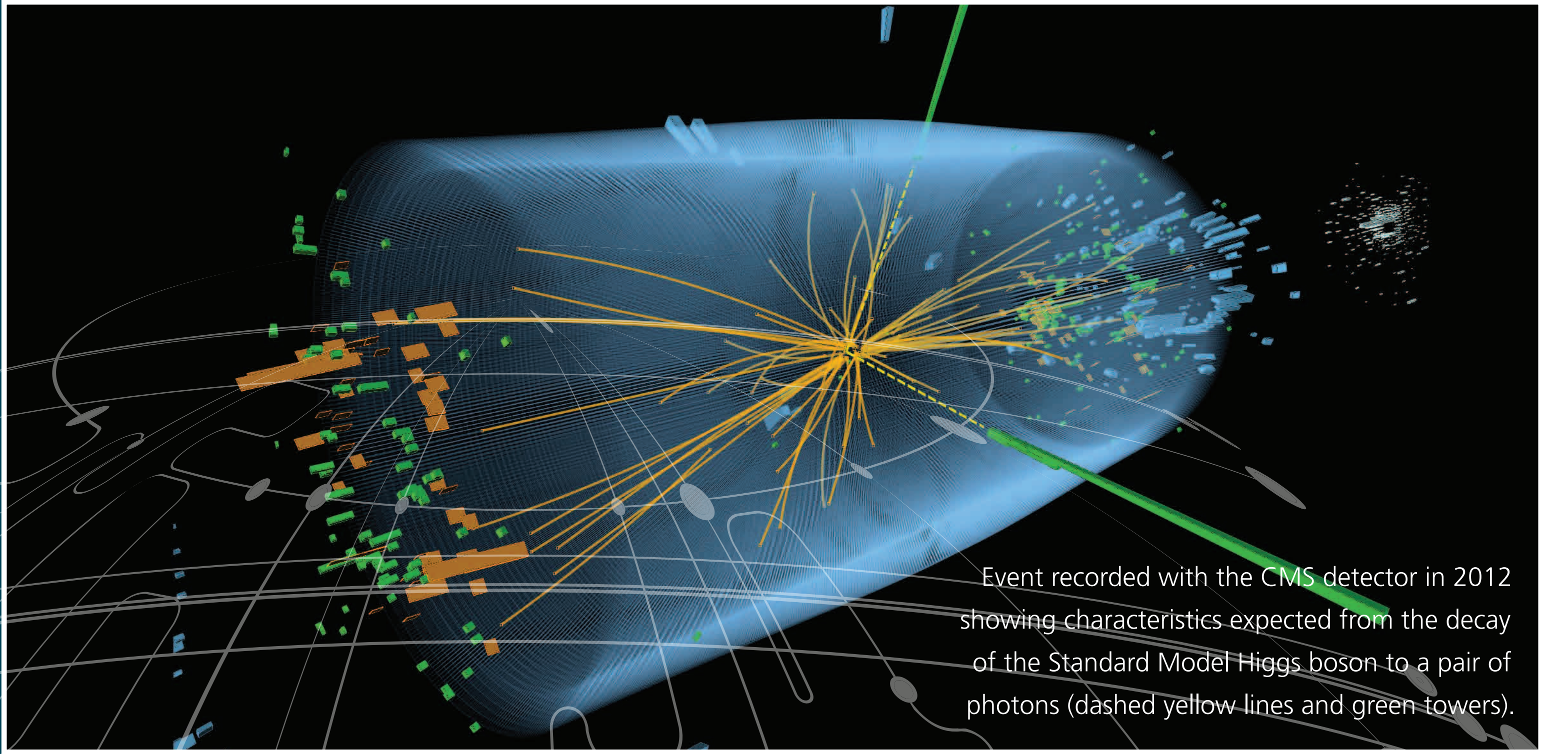
CERN openlab entered its fourth phase of three-years in 2012, and continues to go from strength to strength. Founded more than a decade ago to develop the innovative IT systems needed to cope with the unprecedented computing challenges of the LHC, openlab unites science and industry at the cutting edge of research and innovation.

CERN openlab is a flexible collaboration, allowing partnerships to flourish when the time is right for both CERN and industry. For that reason, we have welcomed new partners along the way, said the occasional, we hope temporary, farewell, and enjoyed long-standing relationships with IT companies whose aims are closely aligned with ours. Our longest-standing relationship pre-dates openlab by many years. In 2012, we celebrated the 30th anniversary of the Oracle-CERN collaboration, which helped pave the way for new and sophisticated IT solutions to support the experiments at LEP, one of which I joined in 1984. In 2012 we also celebrated ten years of HP's CERN openlab membership, while Intel passed that milestone in 2011.

It will, I hope, have come to your attention that CERN experiments announced a major discovery on 4 July 2012, a day that in some circles is now being referred to as Higgs-dependence day! The particle we have discovered was first discussed by scientists including Peter Higgs, Robert Brout, François Englert, Gerry Guralnik, Carl Hagan and Tom Kibble in the first half of the 1960s. Today, it bears the name Higgs boson, and is linked to the mechanism that gives fundamental particles their mass. Without it, we simply would not be here, so to put it mildly, it is important. CERN openlab played a vital part in that discovery. Without the ability to sift, record and process the huge volumes of data generated by the Large Hadron Collider (LHC), it would have been impossible for the experiments to sort the wheat – a few hundred Higgs candidates – from the chaff among the trillions of particle collisions that took place.

It is for reasons like this that openlab is so important, and why I am so pleased to see it in such good health as it enters its fourth phase. HP, Intel, Oracle and Siemens are confirmed partners for the full three years, with a new friend, Huawei, joining for one year initially. The synergy that CERN openlab creates with leading IT companies is vital and I thank all our partners and contributors for their support. Over the next three years, CERN will prepare the LHC for even greater performances, and this fourth phase of openlab will be instrumental in ensuring that our IT capacity meets the challenge.

The Context



Event recorded with the CMS detector in 2012 showing characteristics expected from the decay of the Standard Model Higgs boson to a pair of photons (dashed yellow lines and green towers).

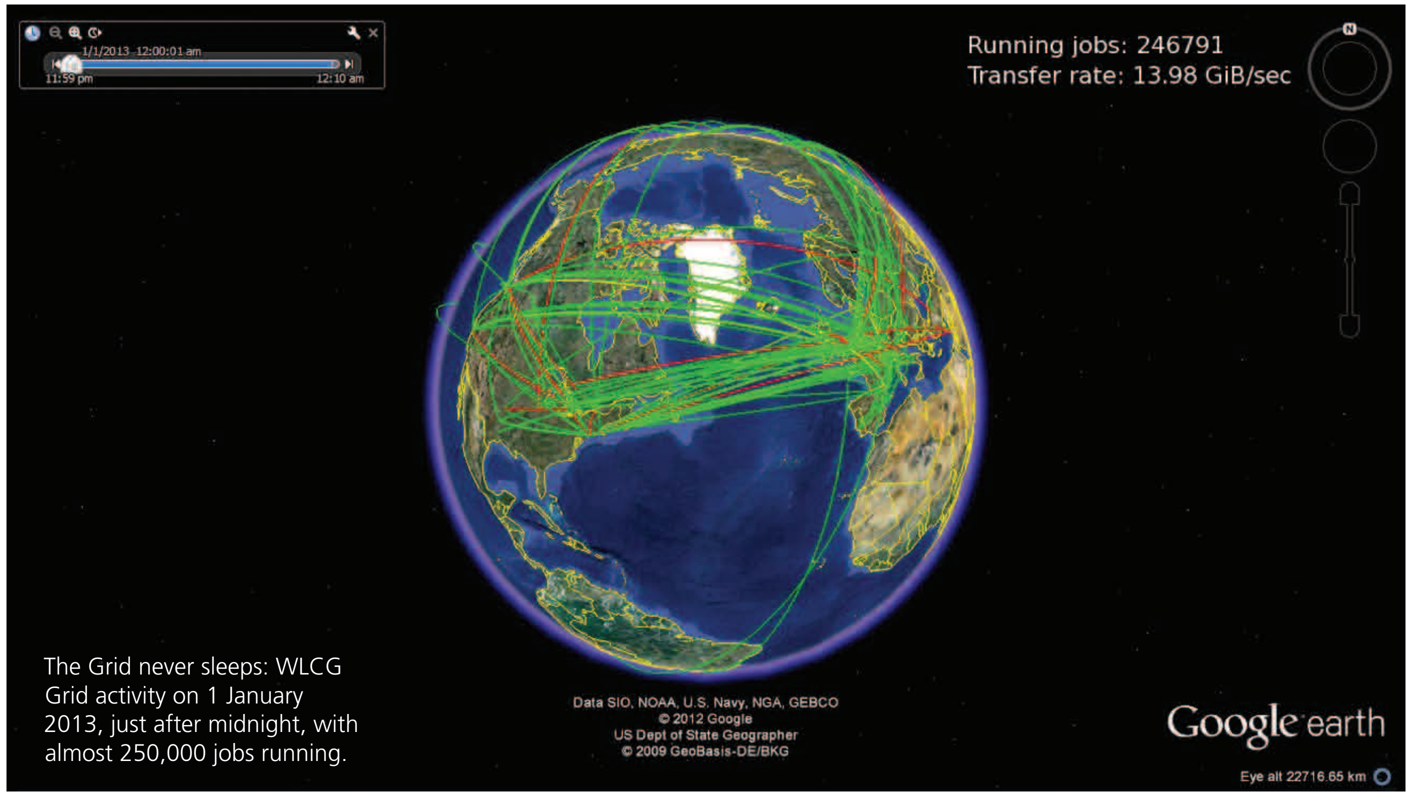
A historical year

The Higgs-like particle discovery: a defining moment in the history of science and the first of the LHC's major discoveries.

The Large Hadron Collider (LHC) project started more than 20 years ago, with the aim of preparing the next major phase in the on-going quest for a deeper understanding of the fundamental laws of nature. Now that it is up and running, it is the world's most powerful particle accelerator and also the largest and most complex scientific instrument ever built. Located in a 27 km long circular tunnel buried 50-150 m below ground, it accelerates particles to more than 99.9% the speed of light to energies never reached before. Some 9600 superconducting magnets operating at just 1.9 degrees above absolute zero (-271.3˚C), colder than outer space, provide the very strong magnetic fields needed to keep the particles on the right orbit. Two beams of particles travel inside two ultra-high vacuum pipes in opposite directions and are brought to collision in four well-defined points, recreating the conditions that existed a fraction of a second after the Big Bang. Four very large detectors, comparable to huge high-resolution 100 megapixels 3D cameras, record these 'mini Big Bangs' up to 600 million times per second once the machine is running at its full potential.

The LHC saw its first beam in September 2008, but stopped operating for slightly over a year following a severe incident caused by a faulty magnet interconnect. After the repair and the installation of additional protection systems, the accelerator started operation again on 20 November 2009. Milestones were passed extremely quickly, and the first world record beam energy was set on 30 November, promptly followed by many others. On 30 March 2010, beams collided in the LHC with an energy of 7 TeV, marking the start of the LHC research programme. The ALICE, ATLAS, CMS, and LHCb experiments immediately observed and recorded events in their detectors. In April 2012, the Large Hadron Collider (LHC) experiments could start taking data at the new collision energy of 8 TeV (up from 7 TeV in 2011).

By 4th July, the ATLAS and CMS collaborations had crunched enough data to announce that they had both observed a new particle, consistent with the long-sought Higgs boson,



The Grid never sleeps: WLCG Grid activity on 1 January 2013, just after midnight, with almost 250,000 jobs running.

a major announcement covered by more than 17,000 press articles in 108 countries and an estimate of a billion viewers worldwide. The discovery was made possible by the excellent performance of the LHC, the experiments' computing systems, and the Worldwide LHC Computing Grid (WLCG). By the end of the year, each experiment was in the position to report that the significance of its observation stood close to the 7 sigma level, well beyond the 5 required for a discovery, and that the new particle's properties appeared to be consistent with those of a Standard Model Higgs boson. They were both careful to say, however, that further analysis of the data, and a probable combination of both experiments' data in 2013, will be required before some key properties of the new particle, such as its spin, can be determined conclusively.

2012 was a year particularly rich in announcements. Indeed, the LHC not only performed proton-proton runs this year but it collided protons with lead ions for the first time in a test run for further lead-proton collisions in 2013, and rounded off the year in style with a new performance milestone, nearly doubling the number of proton bunches in the machine, just as the three-year proton run came to an end. Furthermore, the LHCb experiment discovered two excited states for the Λb beauty particle, and measured one of the rarest processes so far observed in particle physics, the decay of a Bs (pronounced B-sub-s) meson into two muons. ALICE performed detailed studies of the quark-gluon plasma, the matter of the primordial universe, and measurements from the TOTEM experiment are giving insights on the structure of the proton and provide input to the analyses of the other LHC experiments. In addition, exciting results came from the antimatter hall with ALPHA making the first spectroscopic measurements of antihydrogen.

Computing is at the heart of these discoveries. The new phenomena that scientists hope to find are extremely rare, hidden deep in already known physics. The LHC has therefore been designed to produce a very high rate of collisions (600 million per second) such that rare events can be found within a reasonable time. The amount of raw data produced per second, once the machine is up to its full potential, is in the order of one million gigabytes. None of today's computing systems are capable of recording such rates, nor would it make sense. Sophisticated selection systems, called first level triggers, allow most of the events to be rejected after one millionth of a second, followed by a higher level of selection applying more sophisticated criteria. This drops the data rate per experiment to below one gigabyte per second. Even after such a drastic data reduction, the four big experiments, ALICE, ATLAS, CMS and LHCb, produced over 25 petabytes this year, the equivalent of 5.3 million DVD movies, which would take a thousand years to watch.

To store, share and analyse these data, tens of thousands of computers distributed worldwide are being harnessed in a distributed computing network called the Worldwide LHC Computing Grid (WLCG). This ambitious project supports the offline computing needs of the LHC experiments by connecting and combining the IT power of more than 170 computer centres in 36 countries.

CERN is the central 'hub' of the WLCG, the Tier-0 centre, providing approximately 15% of the total computing capacity. This is where a first copy of all data from the LHC experiments is held and where the first reconstruction and data quality checks are performed. In 2012, CERN signed a contract with the Wigner Research Centre for Physics in Budapest for an extension of the CERN data centre. The Wigner Centre for Physics will substantially extend the capabilities of the WLCG Tier-0, and will mitigate as well the risk of having all of Tier-0 in one location. State-of-the-art networking solutions are now in place to connect CERN and the Wigner Centre with multiple 100 gigabit per second links, and the latter will be operational by end of 2013.

The Tier-0 is connected to 11 other major computing centres (Tier-1) using dedicated optical fibre links working at multiples of 10 gigabits per second. These sites are large computer centres with sufficient storage capacity and round-the-clock support. They provide distribution networks, processing of raw data, data analysis, and storage facilities. The Tier-1 sites then make the data available to Tier-2 centres, each consisting of one or several collaborating computing facilities, which can store sufficient data and provide adequate computing power for specific analysis tasks. The Tier-2 sites cover most of the globe. Individual scientists access these facilities through local (also sometimes referred to as Tier-3) computing resources, which can consist of clusters in a University Department or even individual PCs.

The WLCG runs more than 250,000 jobs at any moment, about 1.5 million jobs per day, (corresponding to a single computer running for more than 600 years), and moves 20 petabytes of data (33 million files) per month. With data available to a community of 10,000 scientists within hours, the speed of analysis has been dramatic, with results ready for publication within weeks. To rise to such unprecedented computing challenges, new and advanced systems were needed requiring the joint forces of science and industry to expand technological boundaries. CERN openlab partners contributed in a tangible way to their development and, as shown in this report, continue to collaborate in various domains on new solutions with success.

The Concept

First CERN openlab IV annual Board of Sponsors meeting, in the presence of the CERN Director-General, the sponsors and the CERN openlab team members.

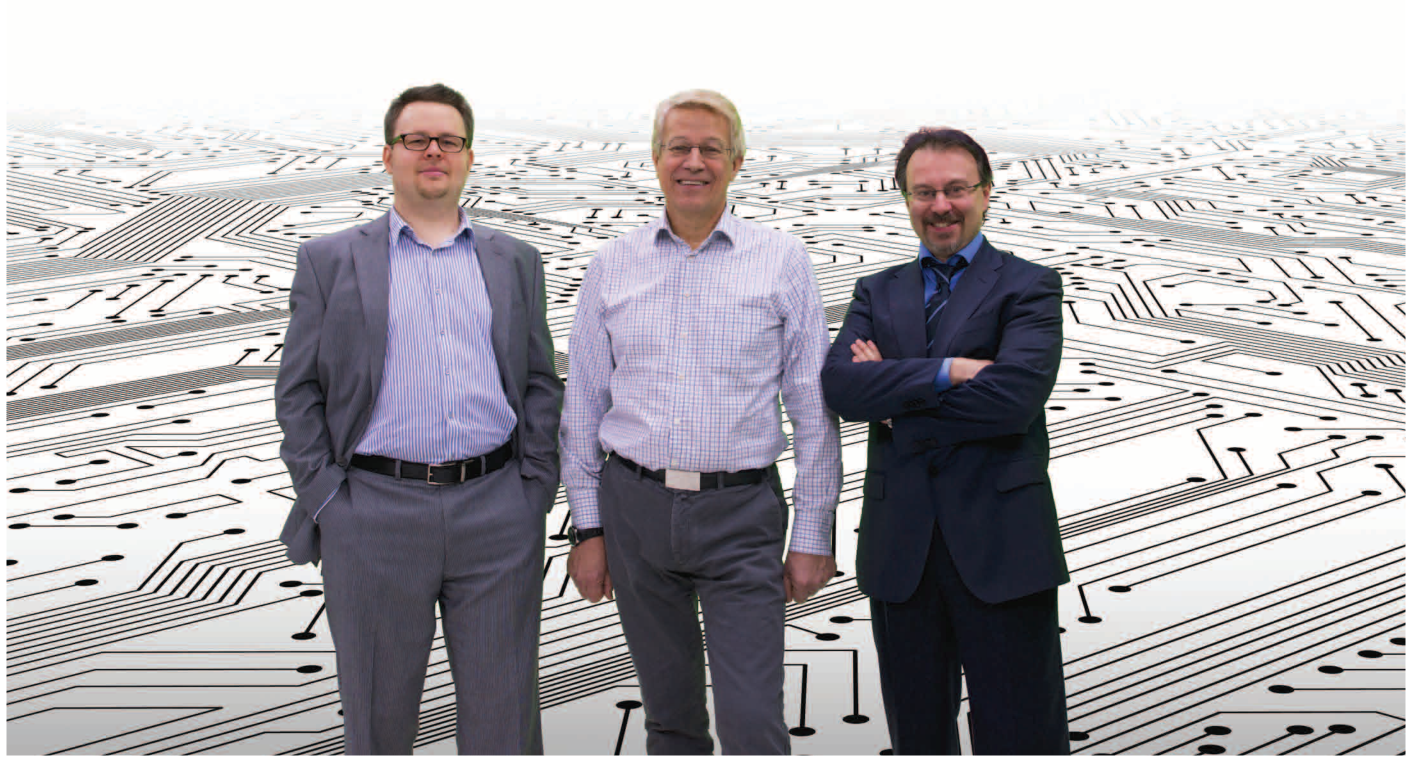
Catalysing collaboration

CERN openlab is a framework for multilateral, multi-year projects between CERN and the IT industry.

Within the CERN openlab framework, CERN provides access to its complex IT infrastructure and its engineering experience, in some cases even extended to collaborating institutes worldwide. Testing in CERN's demanding environment provides the partners with valuable feedback on their products while allowing 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.

Sponsorship can be at the associate, contributor, or partner level. Each type of sponsorship represents a different level of investment. The sponsors engage a combination of cash and in-kind contributions, the cash being used to hire young IT specialists dedicated to the projects. The associate status formalises a one-year collaboration based on independent and autonomous projects that do not require a presence on CERN site. The contributor status is a one-year collaboration based on tactical projects which includes a contribution to hire a young IT specialist supervised by CERN staff to work on the common project. The partners commit to a three-year programme of work and provide three kinds of resources: salaries for young researchers, products and services, and engineering capacity.

The successful CERN openlab concept was formulated in 2001 and stayed basically unchanged throughout the last decade. CERN openlab has been organised into successive three-year phases. In openlab-I (2003–2005), the focus was on the development of an advanced prototype called opencluster. CERN openlab-II (2006–2008) addressed a range of domains from platforms, databases and Grid, security and networking with HP, Intel and Oracle as partners and EDS, an HP company, as a contributor. The combined knowledge and dedication of the engineers from CERN and the companies have produced exceptional results leading to significant innovation in many areas. CERN openlab III (2009-2011) not only capitalised on but also extended the successful work carried out in openlab II with the aim of hosting several major



From left to right, CERN openlab CTO Office group picture: Andrzej Nowak, Sverre Jarp, Alberto Di Meglio. Missing on the picture: Bernd Panzer-Steindel.

projects with a particular focus on technologies and services relevant to CERN and its partners. This annual report covers the first year of the CERN openlab fourth phase (2012-2014).

The current phase is addressing new topics crucial to the CERN scientific programme, such as cloud computing, business analytics, the next generation of hardware, and security for the myriads of networks devices. The technical activities are organised in four Competence Centres (CC): the Automation and Controls CC with Siemens as a partner, the Database CC with Oracle as a partner, the Networking CC with HP as a partner, and the Platform CC with Intel as a partner and Huawei as a contributor for 2012.

Each CERN openlab team is supervised by the CERN staff ensuring the liaison with its sponsor company. At the monthly minor review meetings, the teams are updated about the progress of the on-going projects, which fosters exchanges and ensures timely follow-up. At the bi-annual major review meetings, the sponsors meet with the teams who present their last results, and consider possible synergies. At the occasion of its annual meeting, the board receives information and exchanges views on the progress and medium term plans of CERN openlab.

The CERN openlab team is formed of three complementary groups of people: the young engineers hired by CERN and funded by the partners, technical experts from partner companies involved in the openlab projects, and CERN management and technical experts working partly or fully on the joint activities. A list of the IT and EN departments people 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 partners who contribute significant efforts to these activities. Principal liaisons with partners and contributors are listed on page 13. In addition, valuable contributions are made by students participating in the CERN openlab student programme, either directly to openlab activities or more widely to WLCG, and other Grid and CERN related activities in the IT department.

CERN openlab Management

Rolf Heuer	CERN DG, Chair of CERN openlab Board of Sponsors
Frédéric Hemmer	Head of CERN IT Department
Bob Jones	Head of CERN openlab
Kristina Gunne	Administrative Officer

CERN openlab CTO Office

- Alberto Di Meglio
- Sverre Jarp
- Andrzej Nowak
- Bernd Panzer-Steindel

CERN openlab Communication Office

- Mélissa Gaillard
- Morag Hickman

CERN openlab Fellows and Staff *(Sponsor Indicated)*

Milosz Hulbój	Staff (HP)
Vlad Lăpădătescu	Fellow (HP)
Dan Savu	Staff (HP)
Stefan Stancu	Fellow (HP)
Mirela-Madalina Botezatu	Fellow (Intel)
Julien Leduc	Fellow (Intel)
Yngve Sneen Lindal	Technical student (Intel)
Andrzej Nowak	Staff (Intel)
Liviu Vâlsan	Staff (Intel)
Ignacio Coterillo Coz	Fellow (Oracle)
Andrei Dumitru	Fellow (Oracle)
Maaike Limper	Fellow (Oracle)
Manuel Martin Márquez	Fellow (Oracle)
Lorena Lobato Pardavila	Fellow (Oracle)
Marcin Bogusz	Fellow (Siemens)
Pavel Fiala	Fellow (Siemens)
Filippo Tilaro	Staff (Siemens)
Maitane Zotes Resines	Fellow (Huawei)

CERN openlab Liaisons

- Jean-Michel Jouanigot, Liaison with HP
- Alberto Pace, Liaison with Huawei
- Sverre Jarp, Liaison with Intel
- Tony Cass, Liaison with Oracle
- Manuel González Berges, Liaison with Siemens

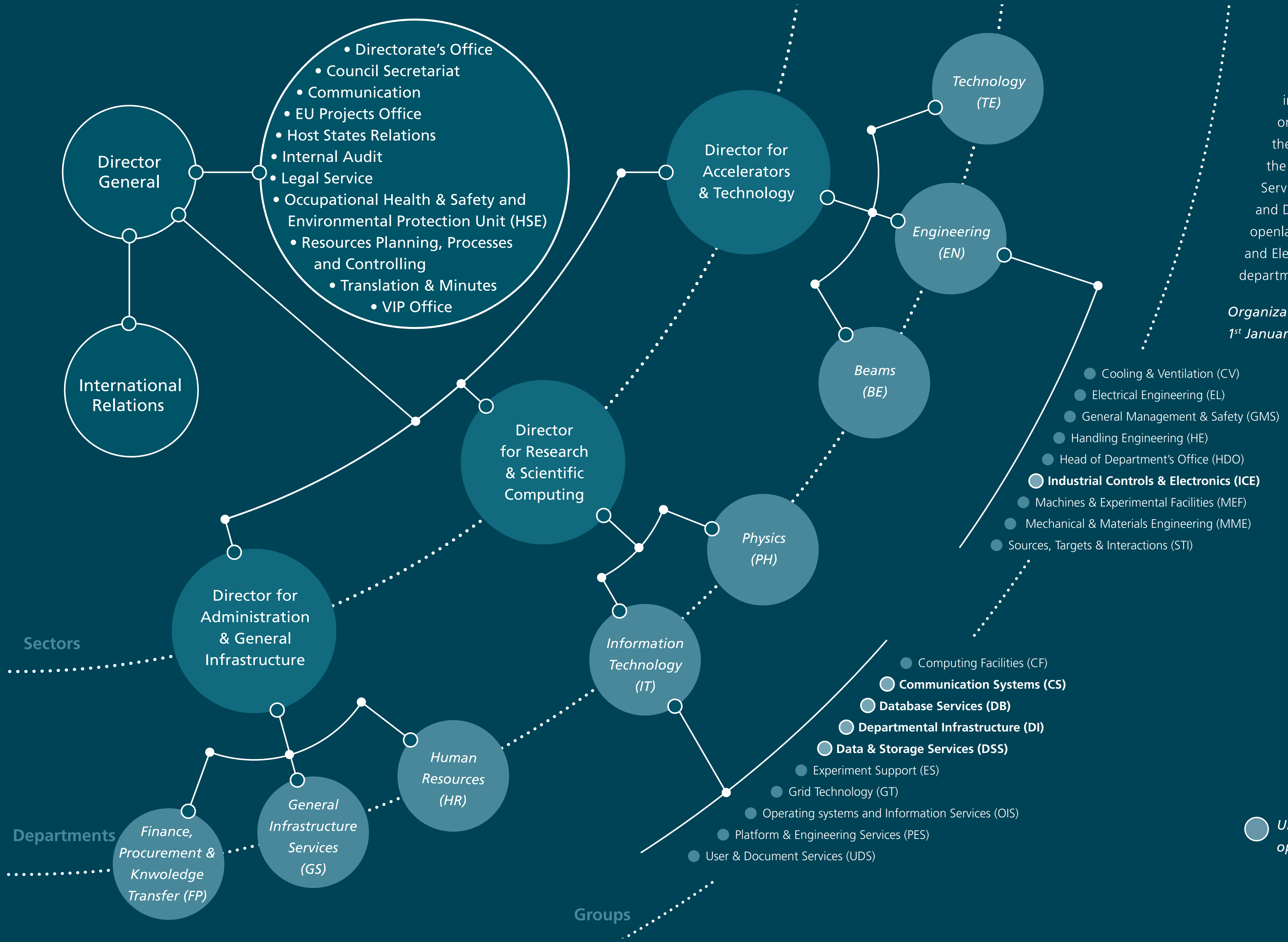
Other IT and EN Departments Staff contributors to CERN openlab

Sébastien Ceuterickx	CS Group
David Gutiérrez Rueda	CS Group
José Carlos Luna Duran	CS Group
Eva Dafonte Pérez	DB Group
Zbigniew Baranowski	DB Group
Eric Grancher	DB Group
Dirk Düllmann	DSS Group
Brice Copy	ICE Group
Philippe Gayet	ICE Group
Piotr Golonka	ICE Group
Fernando Varela Rodríguez	ICE Group

Industry Partner Liaisons with CERN

Charles Clarck	HP
Dan Ford	HP
Bill Johnson	HP
Stéphane Laroche	HP
Amol Mahajani	HP
Rex Pugh	HP
James Prescott Hughes	Huawei
Davis Wu	Huawei
Claudio Bellini	Intel
Herbert Cornelius	Intel
Marie-Christine Sawley	Intel
Stephan Gillich	Intel
C. Gregory Doherty	Oracle
Monica Marinucci	Oracle
Thomas Hahn	Siemens
Guenther Zoffmann	Siemens (ETM)

Positioning CERN openlab activities at CERN



The people involved in the CERN openlab activities are not concentrated in a single group or department at CERN but on the contrary, they span multiple units. In the Information Technology (IT) department, the Communication Systems (CS), DB Database Services (DB), Data and Storage Services (DSS), and DI (Departmental Infrastructure) groups host openlab activities, as does the Industrial Controls and Electronics (ICE) group in the Engineering (EN) department.

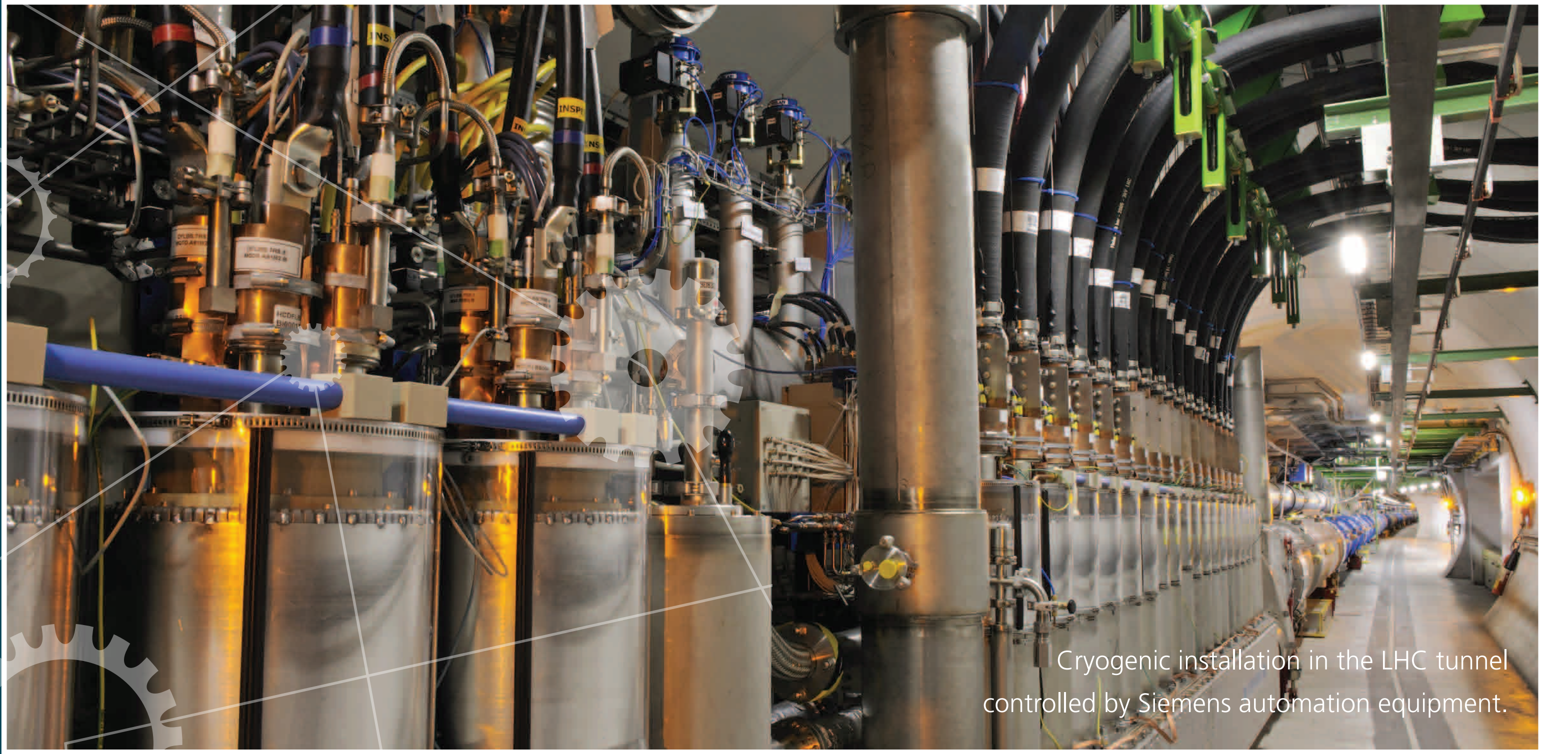
Organization chart as from 1st January 2012

Unit involved in the CERN openlab activities

The Results

Automation and Controls Comptence Centre

Control Systems are at the heart of all CERN facilities, accelerators, experiments and infrastructure.



Cryogenic installation in the LHC tunnel controlled by Siemens automation equipment.

During the last years there has been an increased tendency to use commercial products and benefit from the solutions available in industry. The partnership with Siemens has played an important role in this process. Indeed, CERN systems are extremely demanding and to address such extreme needs, Siemens worked on new solutions which were also beneficial to other clients and business sectors. In the CERN openlab framework, several projects have been carried out by the Automation and Controls Competence Centre (ACCC) team on the different layers of the control system (cf. figure: CERN industrial control systems architecture). At the Process Management layer, the focus has been on security while at the Supervision layer, the joint work has covered data archiving, deployment and remote access.

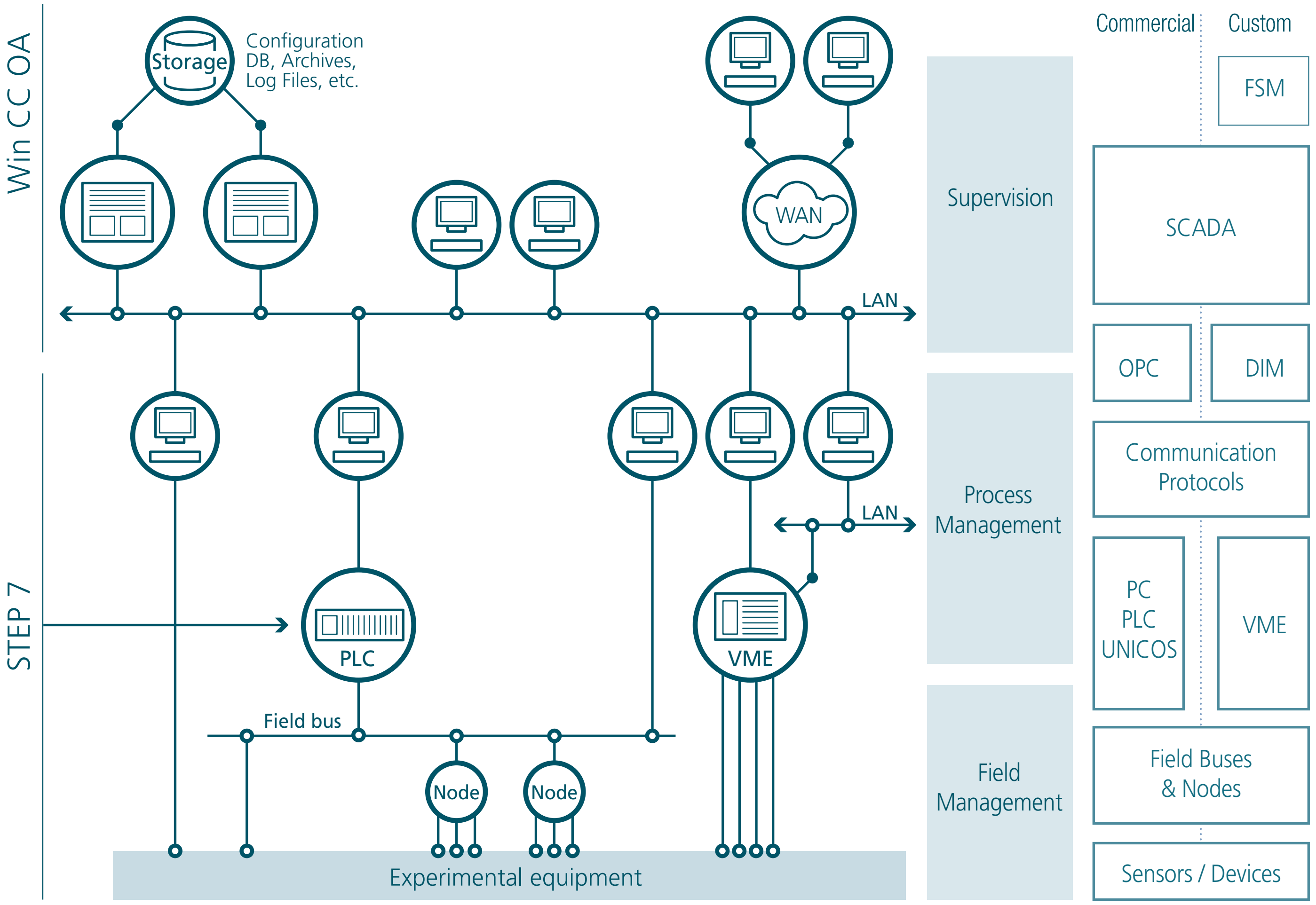
Industrial Security

The increasing number of recent cyber-attacks demonstrates the current vulnerability of hardened Industrial Control Systems (ICSs). This affects the essential services and the critical infrastructure of any manufacturing, power plant and facility relying on an industrial process. Even the most trivial breach into the industrial process control can effect the entire system and leave companies and governments exposed to significant risks.

Historically, ICSs were kept separated from other corporate and business systems, so even though these networks may not have been effectively secured, they were completely isolated for health and safety reasons. This is not true anymore: the continuous integration of the IT business systems with ICSs on one hand provides economical and functional benefits for the companies, but on the other hand exposes the industrial infrastructure to external cyber-attacks. For this reason, it is necessary to adapt the current ICSs technologies to face with the changing threats and vulnerabilities of the cyber world. One of the joint activities carried out by CERN and Siemens within the openlab collaboration focuses on evaluating and improving the ICSs' security. Unfortunately to achieve this, it is not possible to directly apply relative mature IT solutions, because they have been differently designed and sometimes could even interfere with the normal and correct operation of the industrial system.

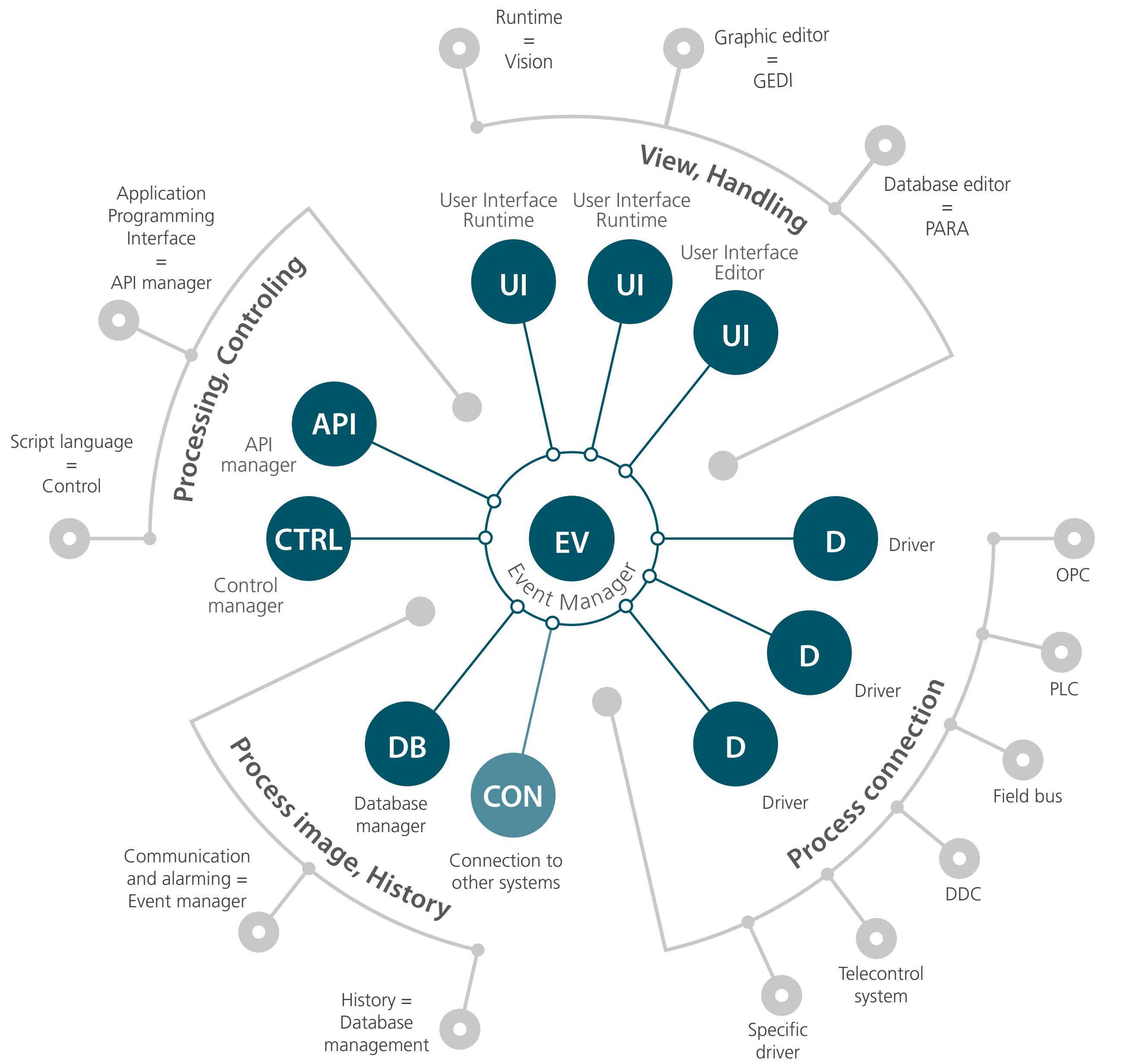
ACCC team picture, from left to right: Pavel Fiala, Manuel Gonzalez Berges, Marcin Tadeusz Bogusz, Piotr Golonka, Filippo Maria Tilaro, Fernando Varela Rodriguez. Missing on the picture: Brice Copy.

CERN industrial control systems architecture



CERN and Siemens have joined their efforts to design and implement an extensible test-bench, specially developed to evaluate the industrial devices robustness by taking into account their specifications and requirements. Complex systems, like CERN experiments, make use of a wide variety of communication protocols to interconnect the different applications and software modules, which, if left unpatched, expose the entire system to risk. This increases the complexity of the testing phase as it is necessary to cover all of them and ensure that they not only function as intended under ideal conditions, but that they also contain no vulnerabilities and are resilient against possible attacks. The benefits of the research and security testing activities at CERN extend beyond the goal of reducing the possibility of an attack but try to enhance the overall system reliability and to improve the robustness of Siemens products.

WinCC OA architecture



Relational Database Archiver

During the past year, the ACCC team kept working on the Oracle archiving module for SIMATIC WinCC Open Architecture (OA) version 4, which is the future unreleased revision of the Supervisory Control and Data Acquisition (SCADA) system currently used by CERN. The next major release is based on the architecture designed to handle Siemens products requiring data archiving. An outcome of the CERN openlab Automation and Controls Competence Centre (ACCC) activities is the so-called Oracle Storage plugin – a software module responsible for handling Oracle database archiving. The module, part of the new architecture, was partially written in 2011; and has been re-tested and adjusted to reflect updated requirements this year.

The collaboration with Siemens also includes know-how exchanges as CERN is running one of the most demanding and complex WinCC OA installations in the world and has unique experience with high-throughput data archiving. The initial work on the new logging architecture's integration into WinCC OA 4.0 has recently been completed while the design phase of the relational database schema for Oracle has started. The work on the new system was carried out in parallel with the work for improving the current Archive Manager (version 3.X) – a part of the system responsible for storage and retrieval of alarms and process value changes from Oracle databases, which are used for data archiving at CERN. The improvements were included in the latest release of WinCC OA 3.11 and comprised better disk data buffering in the event of loss of connection with the database, as well as improved retrieval of data trend graphs by performing grouped queries.

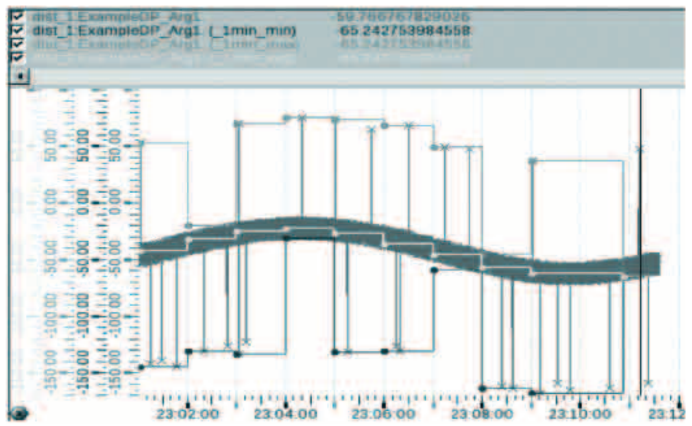
A validation of new features in WinCC OA 3.11 has also been performed within the scope of openlab, as CERN prepares to upgrade its SCADA system to the latest available version during the Long Shutdown in 2013/2014. The new functionality of data aggregations on the database side, called RDB Compressions, has been especially welcomed by the CERN WinCC OA community. The new mechanism helps to reduce the amount of historical data by applying various functions over user-defined time intervals and has been thoroughly tested at CERN, which has led to significant performance improvements.

Centralised Deployment Tool

The WinCC Open Architecture (OA) product from ETM (Siemens) is especially well suited for the control and supervision of very large and highly distributed systems like the controls of the LHC experiments and many accelerator services. These control systems consist of many inter-cooperating WinCC OA applications, which are distributed over many computers. The unprecedented size of these control systems, as well as their long life-cycle require the setting-up of an efficient maintenance strategy. The WinCC OA Centralized Deployment Tool is largely inspired by an existing CERN solution and will aim at:

- Easing the initial set up of the control systems.
- Providing an overview of the different software packages forming each of the application in the control system.
- Providing a powerful way to push upgrades onto sets of applications in a centralised fashion with very limited resources.

Example of RDB compressions and other aggregations over one minute time interval

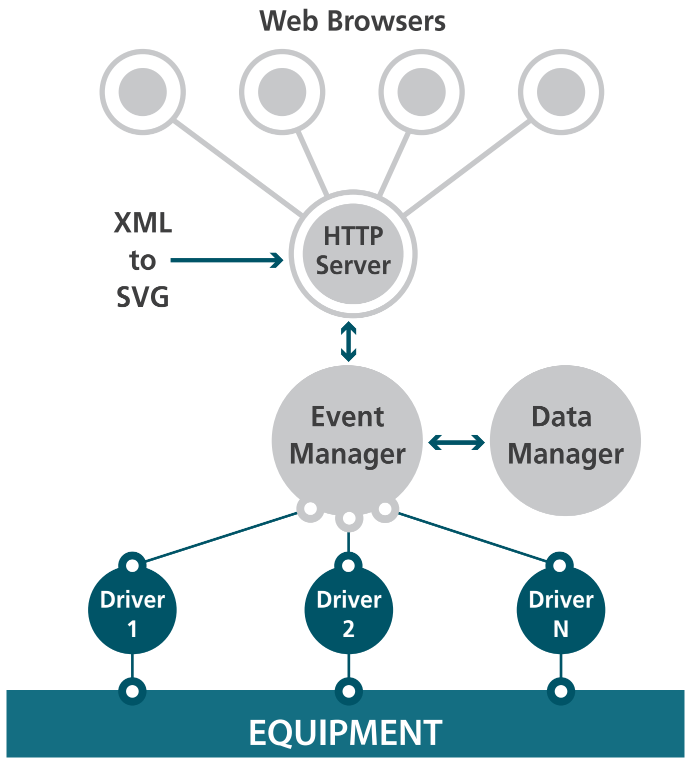


Although the final specification is still under discussion with the different stakeholders (CERN, ETM and some other Siemens branches), different work packages have already been identified and the initial CERN contribution to the project in the context of openlab has been agreed on. In particular, the Central Deployment Tool will make intensive use of the so-called ASCII manager, which is a standard component of WinCC OA, to set up the real-time database of the controls applications. An efficient and robust operation of the ASCII manager is of vital importance for the success of the Central Deployment Tool. In the past few months, the team focused on benchmarking the performance of the process. Extending the current functionality of the ASCII manager by providing support for the XML file format or by implementing pre-validation of the database files prior to their import is also considered. This would facilitate the interface with the enterprise environment (e.g. application generation tools, centralised maintenance tools). Although the work performed is driven by the requirements imposed by the Central Deployment Tool, every WinCC OA user would benefit from the improvements made to the ASCII Manager.

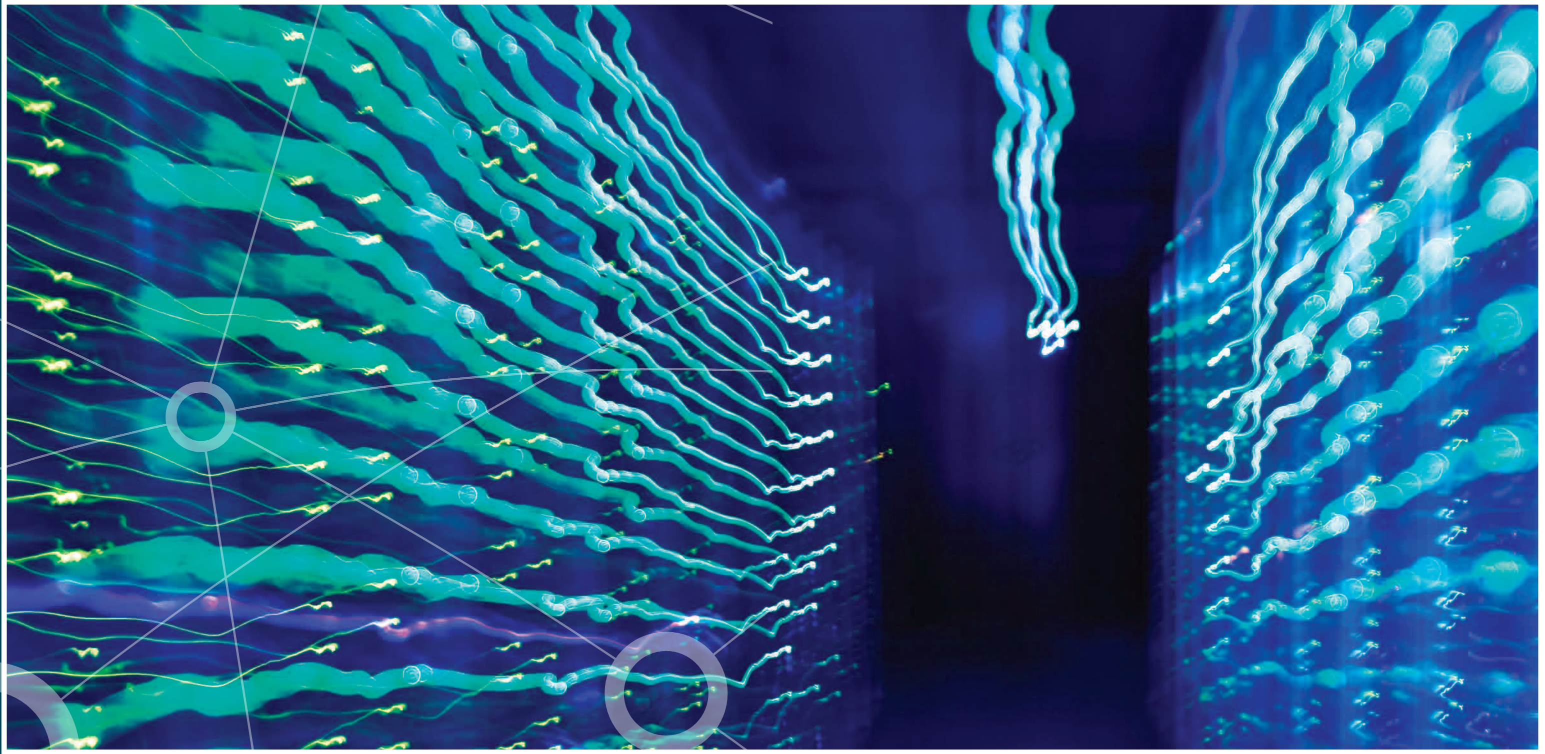
Ultralight Client

Ultralight Client, a new tool providing remote access to control systems, was extensively tested in the CERN environment. The performance and compatibility with the most popular web browsers was checked and the test reports greatly contributed to the improvement of the tool before its final release. Ultralight client integrates Web technologies (e.g. HTTP, SVG, javascript) within the controls environment to allow the system to be monitored and operated by a large number of remote users. This is handled in a user friendly way and installing software on the users' computers is not required. The remote interventions of the stand-by service to minimise the impact of failures on the operation of CERN's facilities is an important use case for this tool.

Ultralight Client processes



The Results



Database Competence Centre

Working with Oracle on key technology for accelerator operations, physics and administration.

In January 2012, the start of the fourth phase of CERN openlab saw an increase in the number of Oracle sponsored fellows. In addition to expanding existing openlab studies of database technology, database monitoring and replication and virtualisation, two new work areas were added to the CERN openlab Database Competence Centre (DCC) programme of work in openlab IV: an investigation of the feasibility of using a relational database for physics data analysis and a study of how CERN could take advantage of Oracle analytics capabilities to extract knowledge from and gain further insight into the large amounts of data collected on the LHC operation and other CERN systems.

Discovering the Higgs boson in an Oracle database

Maaike Limper, an experienced physicist from the ATLAS experiment, joined the DCC team at the beginning of 2012 to investigate the feasibility of using a relational database for physics data analysis, the first of the two new topics. She quickly made progress porting a C++ based analysis using CERN's ROOT framework (http://root.cern.ch) to a mixture of SQL (Structured Query Language) and PL/SQL (Procedural Language / Structured Query Language) operating on data stored in Oracle. The analysis chosen was a search for events where a standard model Higgs boson is produced in association with a Z^0 boson, a particularly interesting Higgs decay mode.

The first step, of course, consisted in making sure that the analysis in an Oracle database produces results that are absolutely identical to those produced with the ROOT based analysis. Maaike successfully demonstrated that the results are actually identical, at least for a simple analysis, as can be seen from the plots on page 25 and as highlighted and presented during the DCC team visit to Oracle's headquarters in March.

DCC team picture, from left to right: Tony Cass, Maaike Limper, Zbigniew Baranowski, Ignacio Coterillo Coz, Manuel Martin Márquez, Lorena Lobato Pardavila, Andrei Dumitru, Eva Dafonte Perez, Eric Grancher.

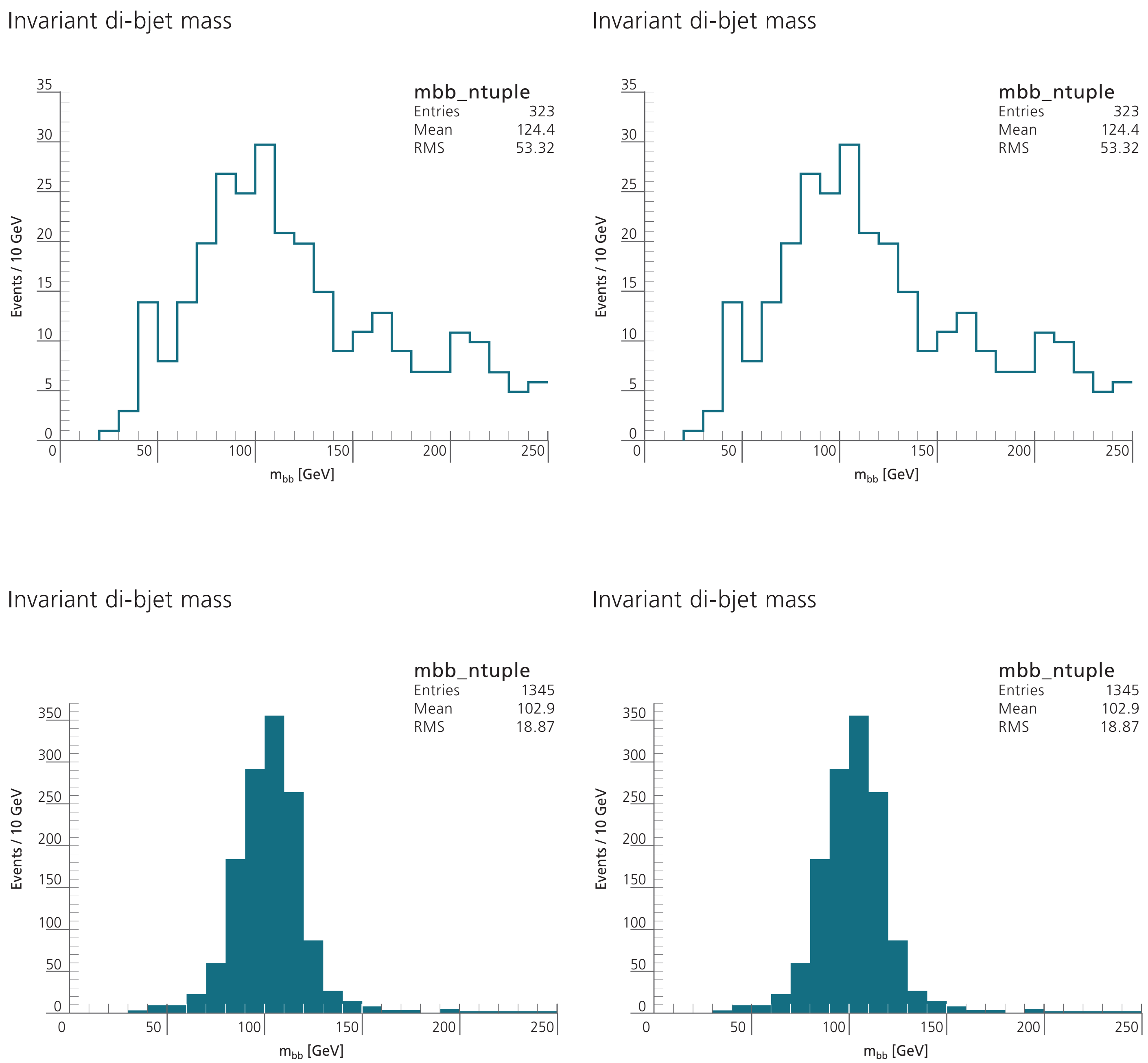
Further time and effort was then needed to understand how to efficiently invoke C++ algorithms that were too complex for rewriting in PL/SQL, and to optimise the internal database processing. By the end of the year, although the database version of the analysis was slower when running on a Monte Carlo generated set of 30,000 signal events, it was more efficient when run against a set of some 1,650,000 background events. The reason for this is that the background sample has fewer interesting events and these are selected very efficiently by the database.

In principle, this is good news for the database version since a "real-world" analysis of the data would be over a much larger sample of events with an even smaller percentage of events tagged as containing a Z^0->ll decay. However, in the real world, the ROOT based analyses are performed in parallel around the world at sites participating in the Worldwide LHC Computing Grid (WLCG), with many thousands of analyses running at any one time. As foreseen when preparing the programme of work, prior to the start of the fourth phase of CERN openlab, access to Exadata systems is now required to see how the database would support multiple different analyses in parallel. The DCC team looks forward to being able to make these tests in 2013.

Extracting knowledge from operations data of LHC with Oracle Analytics

Whilst the physics output of the LHC experiments are stored outside the relational databases, Oracle databases nevertheless play a key role in the operation of the LHC accelerator. Logging data from the thousands of different LHC components has grown from 60 terabytes (TB) and 2.2 trillion records at the time of the last openlab annual report to 165 TB and 4.8 trillion records today. Greg Doherty, the DCC Development Executive Sponsor at Oracle has long commented that active analysis and mining of this huge dataset might help improve the operation of the accelerator by identifying patterns of events that indicate potential problems.

Plots showing results from an Oracle-based analysis (left) identical to those from a ROOT based analysis (right)



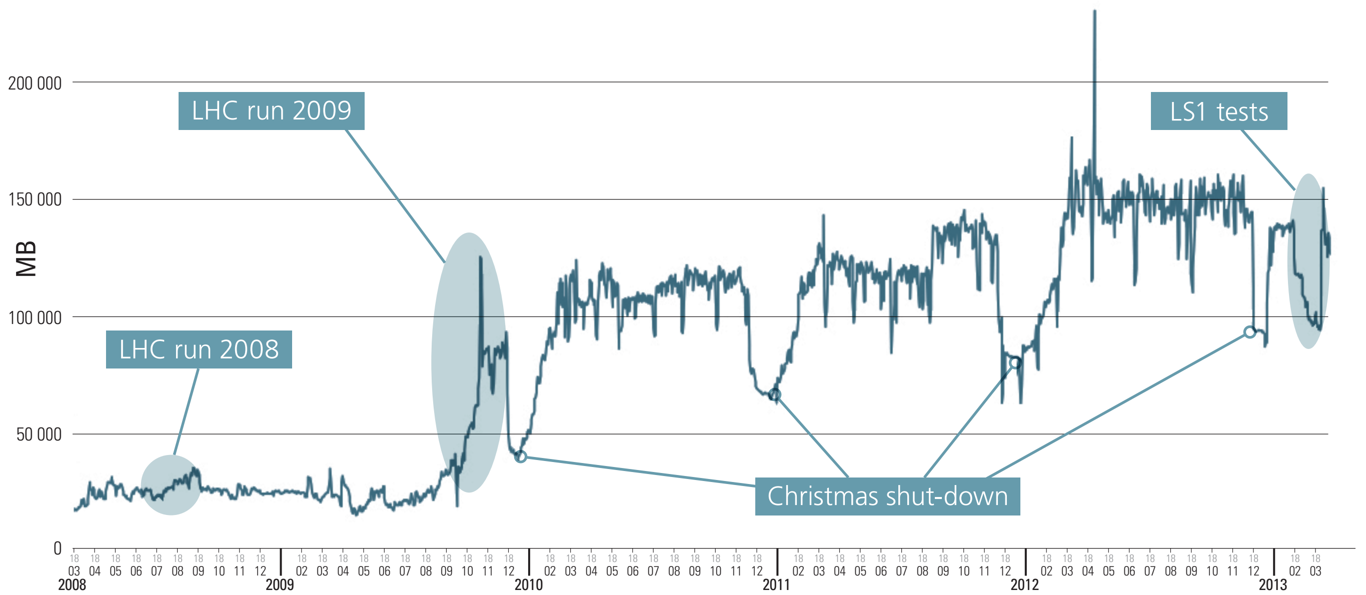
Following up on these ideas was the reason for the creation of the second new work area in the DCC for openlab IV. Manuel Martin Marquez joined the team over the summer and was responsible for organising a successful Data Analytics workshop at CERN in November. The workshop brought together speakers from CERN, Oracle and another openlab partner, Siemens, to address specific use cases at CERN. Presentations on Oracle's Real-Time Decisions (RTD) and Oracle Endeca products attracted much interest and a regular Forum has been established to continue work in this area.

Database technology, replication and virtualisation

The well-established lines of the collaboration in CERN openlab for core database technology, database replication, and in the monitoring and virtualisation areas, have of course continued in openlab IV. Two of these areas, database technology and virtualisation, converge with the development of multi-tenancy database technology for the next generation Oracle database. This potentially offers a very attractive and efficient option to replace databases in virtual machines and the database team at CERN, including Andrei Dumitru from CERN openlab, extensively tested this feature in the framework of the Oracle Database 12c Beta Testing programme. The next step will be to exploit pluggable databases, and other 12c features, in a production environment.

Deploying technology in production which had been tested first in the CERN openlab framework is exactly what has been done in the area of database replication. Oracle Active Data Guard was evaluated in openlab during the early stages of the 11g database release. When CERN databases migrated to 11g in early 2012, Active Data Guard was deployed to replace the Streams based replication of the databases supporting the online operations of the LHC experiments. Active Data Guard replicas have also been added to offload some primary databases, notably in the accelerator operations area where the secondary databases can be used to support the data analytics work mentioned earlier without any risk of affecting production accelerator operations.

CERN Accelerator Logging Service daily storage



The logging service stores data using Oracle RAC databases, of close to one million pre-defined signals coming from heterogeneous sources, and it provides access to logged data for more than 700 registered individuals, more than 100 registered custom applications from around CERN, and even offsite access for purposes such as the CNGS experiments in Gran Sasso Italy.

In the replication area, the DCC welcomed a new team member, Lorena Lobato Pardavila, who has been working on various GoldenGate features, notably an improved monitoring plugin for Enterprise Manager and technology that has the potential to address issues identified in earlier openlab work. The DCC team hopes to join a forthcoming beta programme and to invest more effort in this area during 2013.

The final new team member, Ignacio Coterillo Coz, has been working in the virtualisation area, continuing the DCC work with the Oracle VM team, as well as the joint openlab project with Intel to evaluate the impact of Single-Root I/O Virtualisation (SR-IOV) for virtualised database workloads. Here, the ability of Oracle's Real Application Testing technology to record and replay workloads means that the DCC team can measure directly the impact of SR-IOV on the performance of CERN's applications rather than having to make assumptions with we extrapolations from test loads. These tests show that use of SR-IOV can reduce the time taken to execute I/O intensive queries significantly (by up to a factor of three in some cases) which is very positive as the team looks to further exploit virtualisation to improve service flexibility and overall reliability.

In summary, it has been another fruitful year for the DCC. Oracle's extended support in openlab IV has enabled the team to launch interesting investigations for physics data analysis and data mining, as well as to continue evaluating exciting new developments, giving direct feedback to the Oracle developers. Although the work in the latter area cannot be disclosed, the solutions tested by the DCC team are promising and very much in line with CERN's ideas for future service needs. Working on these in the openlab context has been again extremely beneficial to both CERN and Oracle and more will come in 2013.

The Results

Networking Competence Centre

The decade long CERN openlab collaboration with HP comprises several long-term successful projects addressing the challenges of modern computer networks.

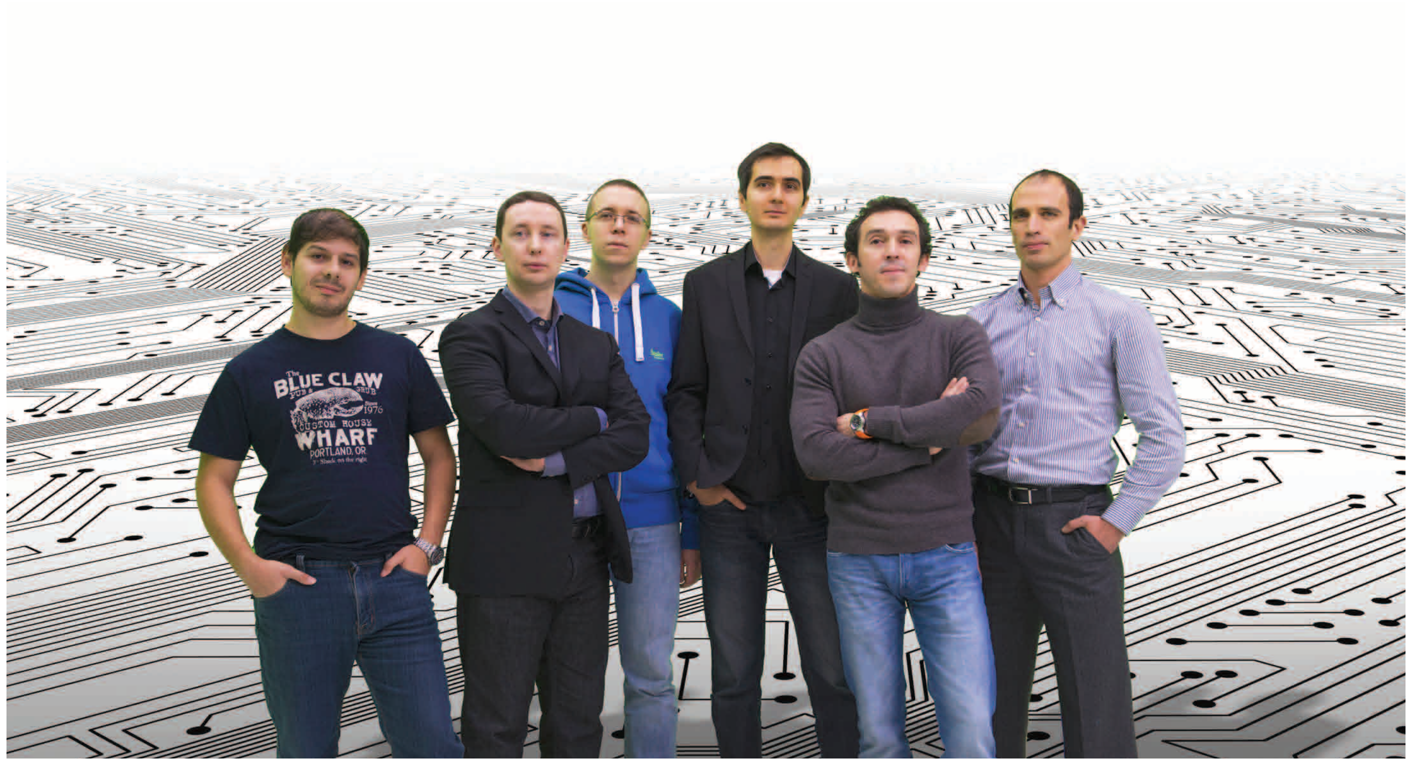
In 2012, the WIND (Wireless Infrastructure Network Deployment) project was launched to carry out a research activity and provide new algorithms, guidelines and solutions to support the deployment and operation of the Wi-Fi infrastructure at CERN. This research activity is now producing extremely concrete results and provides CERN's Wireless Service with unique monitoring and diagnostic information. HP has also expressed interest in incorporating the developed technologies in their products.

Given the broad dynamic of Software Defined Networking (SDN), the Networking Competence Centre (NCC) grasped the potential of this technology for scaling out specific sections of CERN's communication infrastructure.
HP's expertise and pioneering work in SDN combined with CERN's challenges created an ideal opportunity to start a new joint project for pursuing this research topic: ViSION (Virtual Services in OpenFlow Networks).

Wi-Fi challenges

Compared to wired networks, the wireless installations pose many more challenges due to the nature of the radio-wave propagation, limited spectrum available and unpredictable user behaviour. In order to deploy a wireless network one has to perform site surveys and spend a lot of time in the detailed frequency planning. Once deployed, it is not always easy to assess much about the current health of the wireless network from the end-user point of view. There are no clear procedures which help pinpointing potential problems. Usually one does not know about a problem until signalled by the user himself. Even when signalled, it is actually difficult to find the root cause. In many cases, the diagnosis is part of a random trial-and-error process that is usually concluded by an on-site visit, which is time consuming. Running a large yet flexible and efficient wireless service is a complex process that requires much work.

The demand for wireless access is growing in all organisations and so are the user needs. BYOD (Bring Your Own Device) is on a rise and it is not



NCC team picture, from left to right: Vlad Lăpădătescu, Sébastien Ceuterickx, Milosz Hulbój, Dan Savu, David Gutierrez Rueda, Stefan Stancu. Missing on the picture: Jean-Michel Jouanigot

uncommon for one user to have several wireless devices. Faster standards (802.11n) created a paradox. It was expected that the higher transmission rates would alleviate some of the previously known problems. Instead, it raised the user expectations, who now treat wireless as a viable alternative to cabled networking. Furthermore, the fact that end users are often unaware of the complexity of wireless deployment and operation adds to the difficulty. In the beginning the WIND team investigated the existing research papers, commercial solutions and future standards in order to identify the most promising directions to explore. The results of this research were published in 2010 in a report "Wireless Control and Optimisation", by Milosz Hulbój and Vlad Lăpădătescu.

CERN provides a particularly challenging environment for wireless networks. The demand for wireless access is growing and so are the user needs. There is a wide range of deployment areas: conference rooms, auditoria, long corridors, offices, warehouse-like buildings, assembly halls and underground installations. As of November 2012, there are over 750 HP base stations installed in over 170 different buildings. The biggest locations host a dense environment of more than 60 base stations. Daily, WIND observes over 10,000 distinct wireless stations with 60% out of them being registered as CERN devices. On average, there are 150 personal base stations operating in parallel to the official infrastructure (referred to as rogue access points).

WIND project findings

Experience from previous HPN-openlab network monitoring and anomaly detection project (CINBAD - CERN Investigation of Network Behavior Anomaly Detection) allowed the WIND team to build an efficient data collection system. Most of the data comes from the HP access points as it is currently not technically possible to collect data from wireless clients without installing special software on them. In order to infer about the performance and "wellness" of a wireless network, the WIND team assessed different metrics. Laboratory tests and real-life observations have demonstrated that the amount of time that a radio is using the wireless medium (either transmitting or

WIND researchers, Milosz Hulbój and Vlad Lăpădătescu, during their visit at HP headquarters in Montreal, Canada.

receiving) allows for inferring about the network state and for identifying a broad range of problems. The team calls this metric a "channel time". After being presented with encouraging results obtained with channel time estimates, HP decided to implement a more accurate way of measuring the channel time in their base stations. It allowed for more accurate identification of numerous wireless problems: co-channel interference, adjacent channel interference, hidden nodes, etc.

During the course of the project, WIND created multiple prototype perspectives for visualising different aspects of wireless network. Each of them was then carefully evaluated together with Sébastien Ceuterickx, the Wireless Service Manager, in order to select the most useful ones. At the same time, the WIND team analysed the most common reported wireless problems and the way in which the Wireless Service approached the problem. WIND was able to codify a part of the problem solving workflow into a decision system process. The rules were codified using the Drools system which allows for easy modification and extension by any user. Such a rule based tool allowed for creating a mechanism that automatically discovers the problems (both station and radio related) within CERN's infrastructure.

The current WIND toolkit enables one to visualise detailed information on the actual state of the network at any point in time. This new facility improves day-to-day operations and reduces the time needed to diagnose network problems. All of these tools have already proven to be useful in resolving problems, including: radio misconfiguration, hidden nodes, station driver problems, co-channel interference, non-802.11 interferers (microwave ovens and wireless headphones) and many more. The WIND team focuses now on transforming the software created into a reliable service that could be used without detailed training by CERN's network engineers to resolve Wi-Fi problems.

At the beginning of November, the WIND team visited the HPN wireless group once more in Montréal, Canada. The main goal of this visit was to present and discuss the latest results and to transfer the knowledge that was gathered during the course of the project. WIND ideas, tools and algorithms were very well received by engineers and product managers. HPN is considering integrating some of these ideas into their portfolio of products. This, together with the on-going deployment and integration at CERN is a clear evidence of the project's achievements.

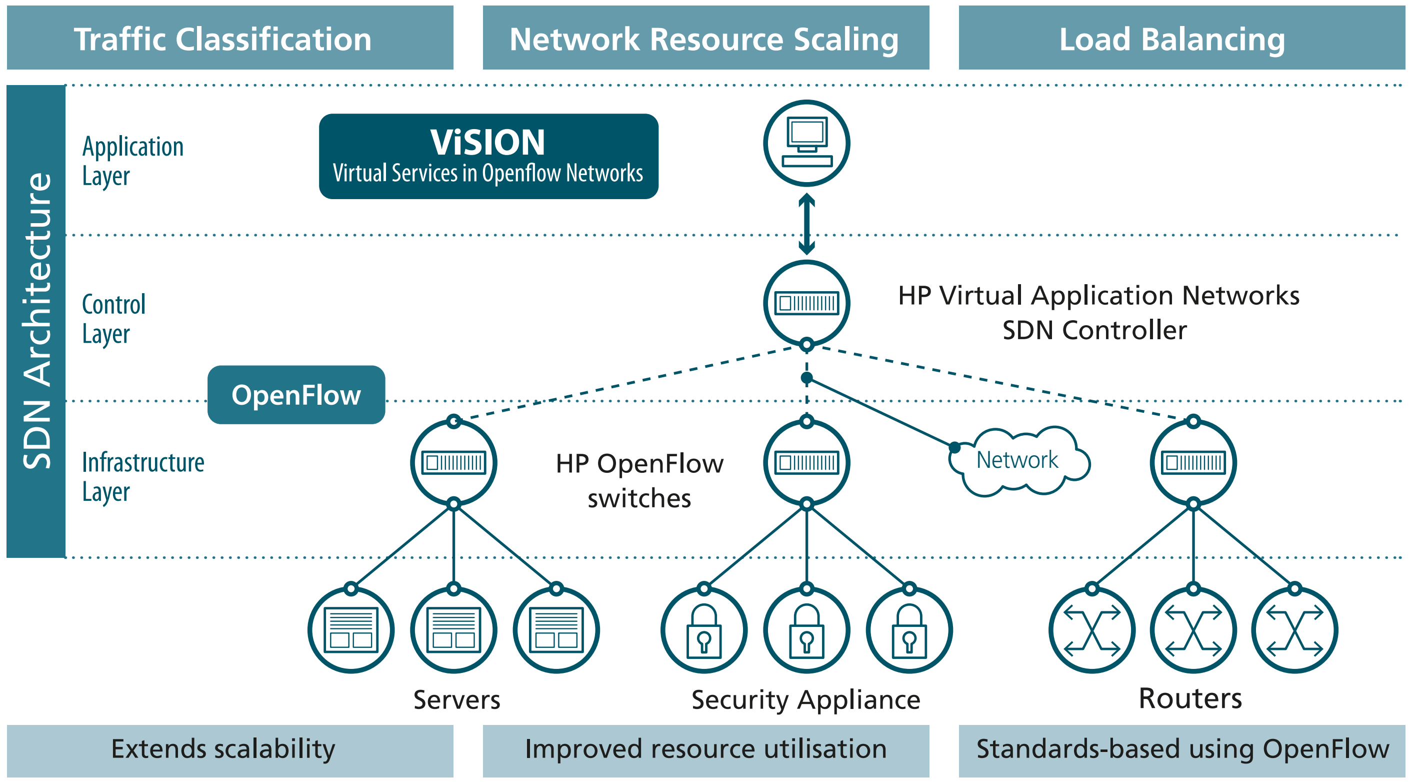
The ViSION project

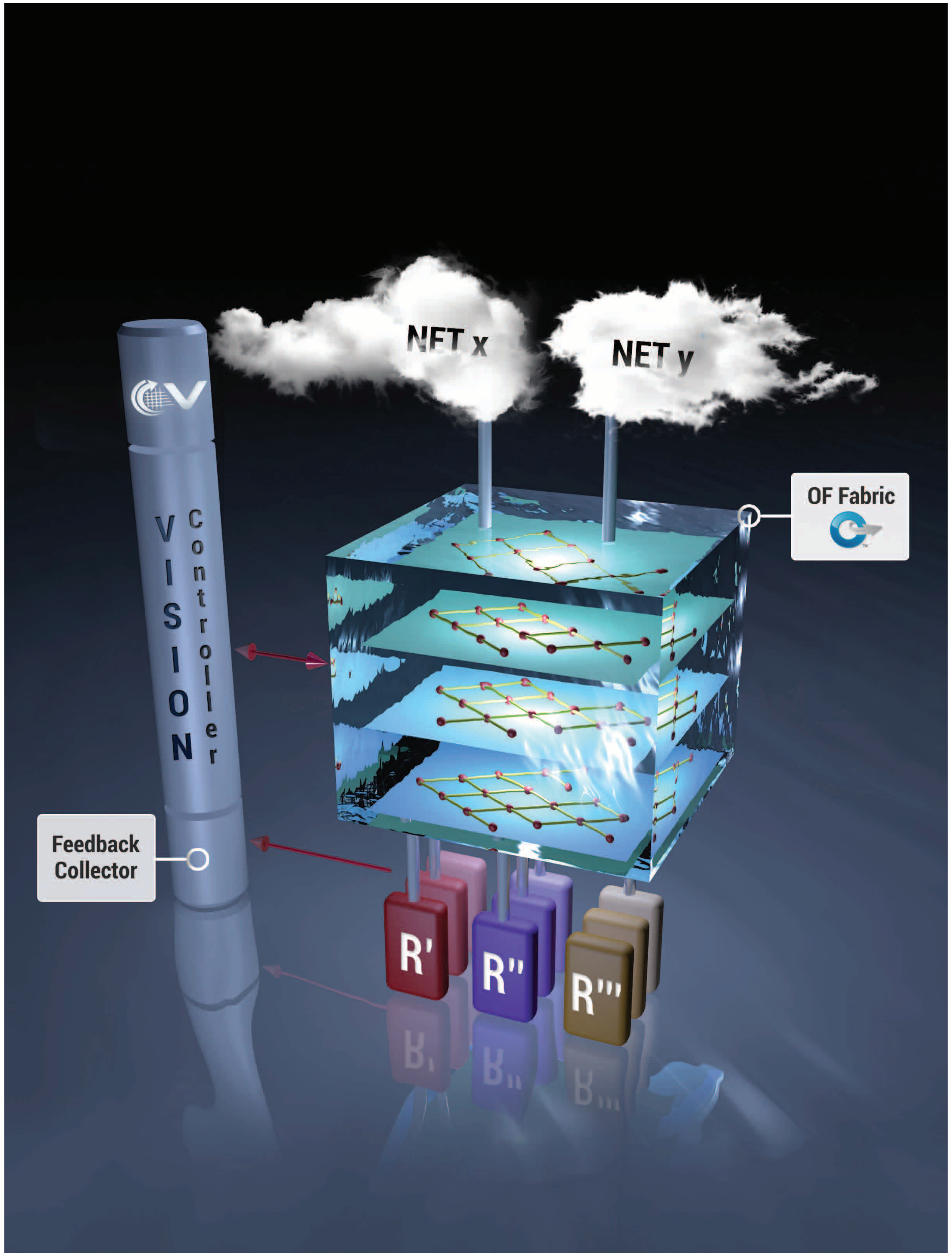
ViSION is a three-year project that started in February 2012 in collaboration with HP Networking. The name stands for Virtual Services in OpenFlow networks, and its target is to optimise network resource utilisation using a Software Defined Networking approach based on the emerging OpenFlow technology. Software Defined Networking is a recent technology that exports the control/decision logic from the networking equipment to an external software controller, which has a global view of the network and can make educated traffic engineering decisions. This significantly differs from the traditional networking paradigm, where each node takes its own decisions. OpenFlow is the main protocol enabling SDN. It has been created by the need to accelerate innovation in the networking field, which has fallen behind software server virtualisation. Initially promoted in academic environments the protocol is now supported by the Open Networking Foundation, comprising the majority of networking manufacturers.

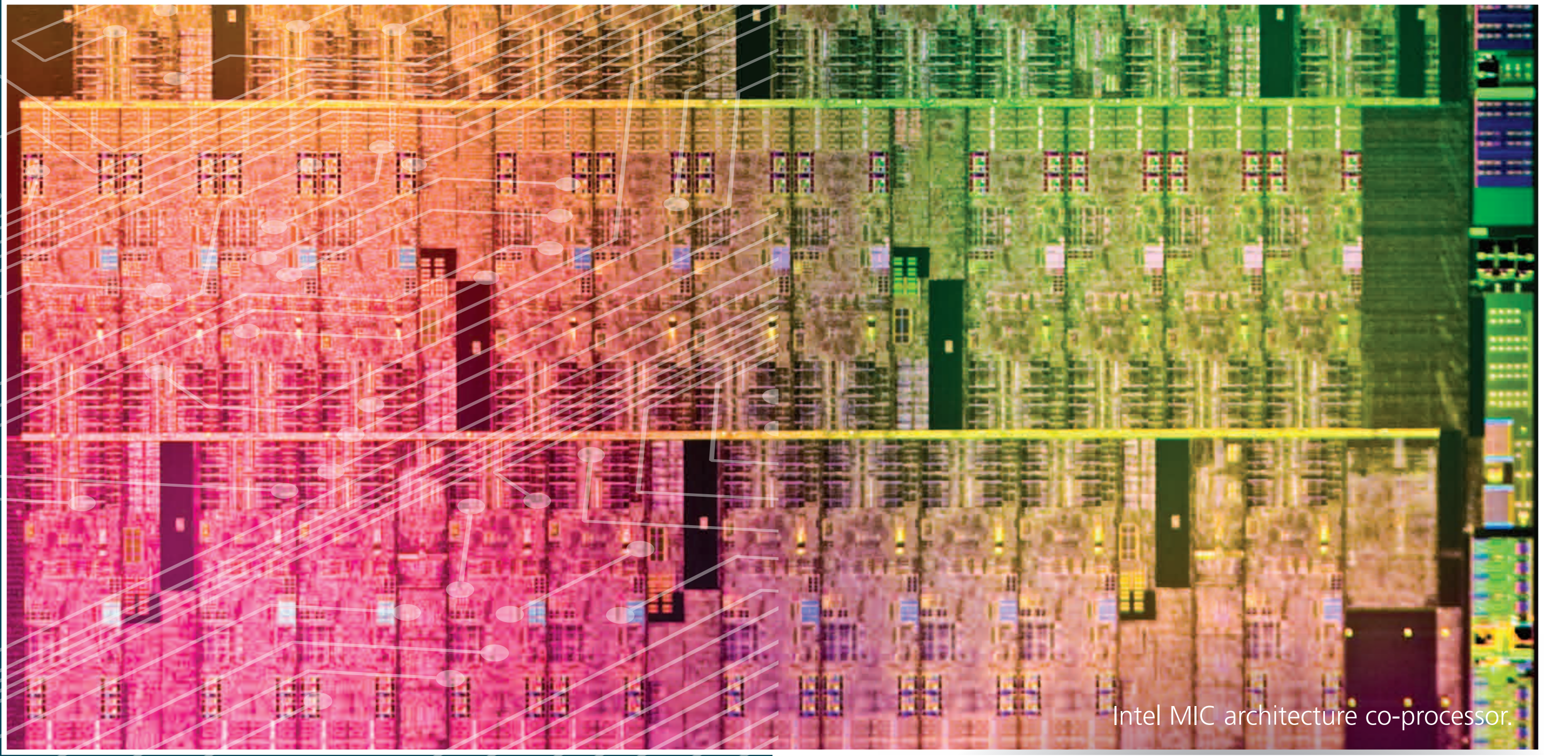
The goal of the ViSION project is to develop an application aware traffic orchestrator using OpenFlow. The orchestrator will distribute traffic over a set of network resources, performing both classification (different types of resources) and load sharing (similar resources). To make the goal achievable, the ViSION team has early access to the recently announced HP SDN APIs and the HP SDN controller. Currently, approximately one year into the project, the initial study and design phase was completed and work has started to develop a prototype.

A particular case of direct interest to CERN in order to cope with the increase of Internet traffic is scaling out its firewall system capacity by using the orchestrator to distribute traffic over multiple units. In light of HP's new Virtual Application Networks SDN solution, HP has a direct interest in developing techniques for application specific traffic orchestration. Furthermore, ViSION will provide HP with valuable feedback on their SDN framework and OpenFlow implementation in their products.

ViSION Software Architecture







Intel MIC architecture co-processor.

The Results

Platform Competence Centre

Evaluating and improving new architectures for tomorrow together with Intel and Huawei.

Huawei joined the Platform Competence Centre (PCC) this year as a contributor, while the work with the Intel partner continued on a high note. Amongst the highlights, for the PCC team members working with Intel, were the move to Intel® Xeon® E5-2600 processor-based servers, the start-up of a new, very successful workshop dedicated to Numerical Computing and a comprehensive report issued on software methodologies for vectorisation and parallelisation. On the Huawei side, the PCC team members have been investigating the applicability of new storage techniques and architectures for the processing of high energy physics data from the Large Hadron Collider experiments.

Several changes in the staffing of the PCC team occurred during the year. Mirela-Madalina Botezatu joined the Intel PCC team as a technical student in February. Thomas Bach, a CERN technical student, joined the PCC in October in order to compose his Master thesis. Julien Leduc, who had come to the end of his three-year openlab fellowship, accepted a Staff position in the Data Storage and Services (DSS) group in the IT Department. Liviu Vâlsan, who was recruited to replace Julien in his CERN openlab position, came with a broad portfolio of skills from the ATLAS online computing group and was able to integrate smoothly into the PCC team thanks to his extensive experience. Finally, Alfio Lazzaro, who had been a CERN openlab fellow for almost three years, specialising in software parallelisation techniques, accepted a position with CRAY at the Swiss Supercomputer Centre in Lugano (CSCS) at the end of his CERN openlab contract. On the Huawei PCC side, Maitane Zotes Resines was hired to work on the joint project, supervised at CERN by Alberto Pace and Dirk Düllmann, while joined on her daily work by Huawei engineers visiting CERN for periods of several months.

Benchmarking and evaluation

The team successfully benchmarked both the Intel Sandy Bridge (Intel Xeon processor E5-2600 family) dual socket server and the four-socket one (with E5-4600 processors). Both reports highlighted the improvements compared to previous generations of a compatible nature and were published

From left to right, CERN openlab PCC-Intel team members: Andrzej Nowak, Liviu Vâlsan, Sverre Jarp, Mirela-Madalina Botezatu, Julien Leduc. CERN openlab PCC-Huawei team members: Maitane Zotes Resines, Dirk Düllmann, Alberto Pace.

on the CERN openlab new website. With the emergence of the Intel® Xeon Phi™ as a mainstream accelerated compute platform, the PCC also focused on other power efficient compute solutions.

The CERN openlab cluster was enhanced with 80 Intel Xeon E5-2600 processor-based servers (at various frequencies) allowing the PCC team to offer state-of-the-art equipment to the workshop participants and collaborators seeking to do research and run benchmarks on the most modern hardware available. These Xeon processor systems are the first generation to offer the Advanced Vector eXtensions with 256-bit wide vectors – a doubling compared to previous systems. The PCC team also maintained a set of individual servers, ranging from low power Intel® Atom™ Boards to "enterprise" servers.

In November Intel announced the availability to the Xeon Phi co-processor, previously known as the "MIC" (Many Integrated Cores) accelerator. The openlab team has been an active partner in its development for many years and proposed multiple improvements, ranging from enhanced double-precision support in the vector-based instruction set, Linux as the operating system on the card hosting the co-processor, and the use of open-source software as broadly as possible. The team evaluated multiple versions of the Phi processor (pre-alpha, alpha, and beta) and obtained interesting results for the applications that vectorised well. This was true, for instance, for the MLfit application mentioned on page 37, but also a trackfitting prototype developed for heavy-ion experiments, such as ALICE. During the year, the PCC team gradually opened up the access to several physics groups who wanted to gain experience on this new number-crunching platform.

Software

A major part of the PCC efforts are dedicated to software, in particular the evaluation of new software versions or new methodologies. The team spent considerable time evaluating the Intel® Cluster Studio XE 2013 Software Suite and reported multiple issues to Intel. The team maintained regular contacts with the SFT (Software design for experiments) group in CERN Physics (PH) Department as a complementary effort to the PCC workshops and teaching activities. In particular, "Concurrency Forums" were held every other week to review strategies for enhancing and rewriting portions of the LHC software frameworks with a goal to enhance the exposed parallelisation. The excellent contacts with key people in Intel SSG helped drive the overall effort.

A constructive dialogue was established with the Intel Exascale Lab in Paris where the researchers assist in optimising the OpenGate medical framework that is built on top of Geant4, a toolkit for the simulation of the passage of particles through matter. Given that the PCC has considerable experience with the Multithreaded Geant4 having participated in its development, the teams met both in Versailles and at CERN to discuss how to leverage this new opportunity and, in addition, how to interact with the software developers of Geant5, a prototype that promises to exploit both parallelisation and vectorisation. A statement of work is planned for the joint activities in 2013 and 2014.

Conferences and publications

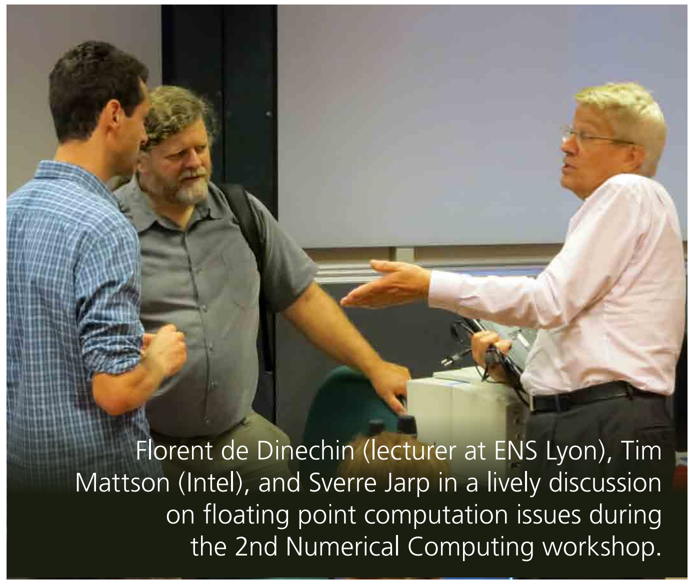
Andrzej Nowak presented a paper at CHEP2012 in New York entitled "The future of commodity computing and many-core versus the interests of HEP software", which was well received by the High Energy Physics community and generated a peak of interest in cutting-edge computing technologies, such as the Intel Xeon Phi. A poster, "Many-core experience with HEP software at CERN openlab" was also selected for lightning talks during the conference. Sverre Jarp gave two presentations at the International Supercomputer Conference 2012 in Hamburg in June. The PCC was also invited to deliver a keynote at the Intel European Research and Innovation Conference 2012 in Barcelona, where Andrzej Nowak presented the work of CERN and CERN openlab to over 400 attendees. Dirk Düllmann presented the "CERN Cloud Storage Evaluation", related to the work done with Huawei, at the HEPiX Fall 2012 conference, in Beijing.

One of the PCC highlights of the year was the publication of an extensive whitepaper on software methodologies for vectorisation and parallelisation using MLfit, a maximum likelihood data analysis application. The report was the answer to a direct request from Intel's Software and Services Group (SSG) to perform an evaluation of the multiple software technologies that now exist. The key software methods used were OpenMP, Intel Threading Building Blocks (TBB), Intel Cilk Plus, and the auto-vectorisation capability of the Intel compiler (Composer XE). Somewhat surprisingly, the Message Passing Interface (MPI) was successfully added, although the focus of the report was on single-node rather than multi-node performance optimisation. The paper concluded that the best implementation in terms of both ease of implementation and the resulting performance was a combination of the Intel Cilk Plus Array Notation for vectorisation and a hybrid TBB and MPI approach for parallelisation. A data mining study on the relationship between Compiler Flags and Performance Events was also published by the team. The idea behind the study was to see if it is possible to quickly identify the performance bottlenecks which exist in a given code and determine the compiler flags that are likely to alleviate a performance issue without compromising the accuracy and reproducibility of the results.

Liviu Valsan and Mirela Botezatu installing one of the new Intel Knights Corner co-processors.

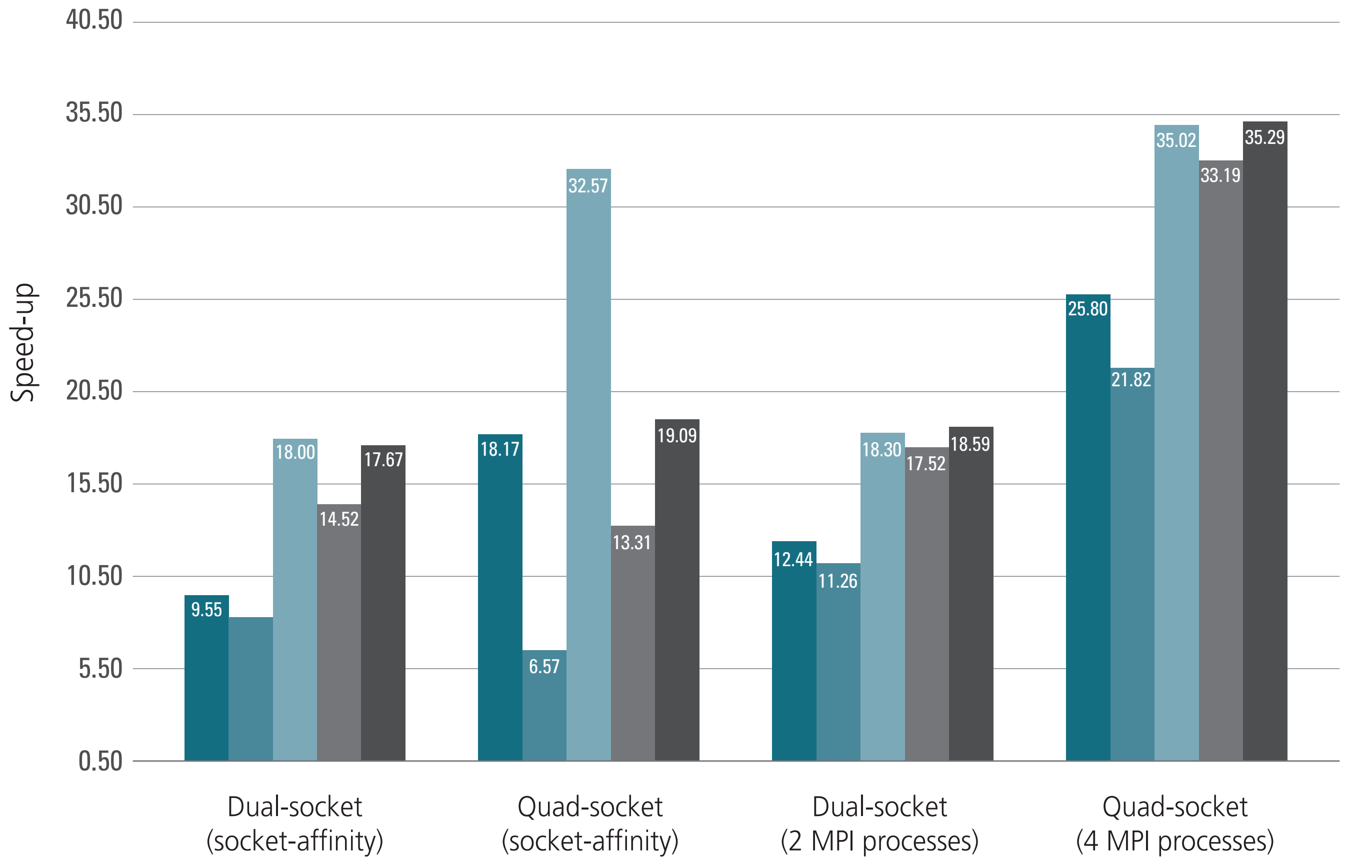
Workshops and teaching efforts

The openlab/PCC team has been running its educational programme comprising regular workshops with a focus on performance tuning and parallelisation since 2007. In this record year, with ten multi-day tutorials offered by the PCC, it was decided to expand the portfolio by proposing a two-day hands-on workshop on Numerical Computing (using IEEE-754 floating-point arithmetic). The first instance was organised in February 2012 with teachers from CERN, ENS in Lyon, Intel SSG, and Intel Research. The workshop was very quickly oversubscribed, and hence was repeated in September,



Florent de Dinechin (lecturer at ENS Lyon), Tim Mattson (Intel), and Sverre Jarp in a lively discussion on floating point computation issues during the 2nd Numerical Computing workshop.

Comparison of several implementations for the dual-socket and quad-socket systems



- OpenMP, fine-grained parallelism, w/o block splitting
- Intel Cilk Plus, fine-grained parallelism, w/o block splitting
- OpenMP, coarse-grained parallelism, w/ block splitting
- Intel Cilk Plus, coarse-grained parallelism, w/ block splitting
- Intel TBB, coarse-grained parallelism, w/ block splitting

The speed-up values were obtained running the implementations with full-loaded systems with respect to reference sequential execution time. Socket-affinity and hybrid MPI parallelisation were used in the tests.

also with a high attendance. In addition, the team was asked by Intel to repeat the workshop in CASPUR in Rome. The two standard two-day workshops were run as usual in the spring, but were collapsed into a three-day workshop, named "Parallelism, Compilers and Performance" in the autumn. This is a formula that will be pursued in the future, as it gives the opportunity to teach a wider portfolio of topics. In addition to these courses, several special workshops were also organised. The first one, held in early July, focused on advanced VTune Amplifier usage, as well as more lectures on TBB, both taught by well-known Intel experts. Finally, a workshop was co-organised with the ATLAS experiment on the Gooda performance analyser. In addition to the presenter David Levinthal from Google/US, the participants enjoyed the presence of several other renowned performance experts: Stéphane Eranian from Google/Grenoble, Patrick Demichel from HP/Grenoble, and Ahmad Yasin from Intel/Israel. The PCC was also invited to give a half-day tutorial at the IEEE ISPA Symposium in Madrid in July, and taught at the CERN School of Computing (CSC) in August as well as the INFN ESC12 School in October.

Upcoming collaboration

The PCC's efforts to expand its collaboration with Intel on additional planes were particularly successful in 2012. In late 2011, CERN openlab, LHCb, Intel, and a group of Irish universities co-authored an EU FP7 project proposal focused on next-generation data acquisition capabilities. This year the ICE-DIP project application was scored very highly by evaluators, particularly because of the cutting edge science proposed – only less than 10% of the proposals are actually funded. As a result, CERN and Intel will benefit from over €1.25 million of funding to hire five bright PhD candidates, who will work with the CERN experiments on world-class cutting-edge technology, a shining example of joint industrial and public sector collaboration.

Huawei's contribution

The year 2012 also brought a new contributor into the Platform Competency Centre as Huawei joined covering the rapidly expanding area of cloud storage. Since CERN is faced with rapidly expanding requirements in storage scalability and performance for physics and infrastructure data, these new technologies and implementations are expected to complement the existing service offerings at CERN and may play a role in lowering the service cost by consolidating services through novel deployment approaches for very large-scale disk storage. Huawei on the other side is expanding their product offering with new highly scalable cloud storage products aimed at providing redundant, large-scale cloud storage closely following the S3 interfaces at a reduced cost of ownership.

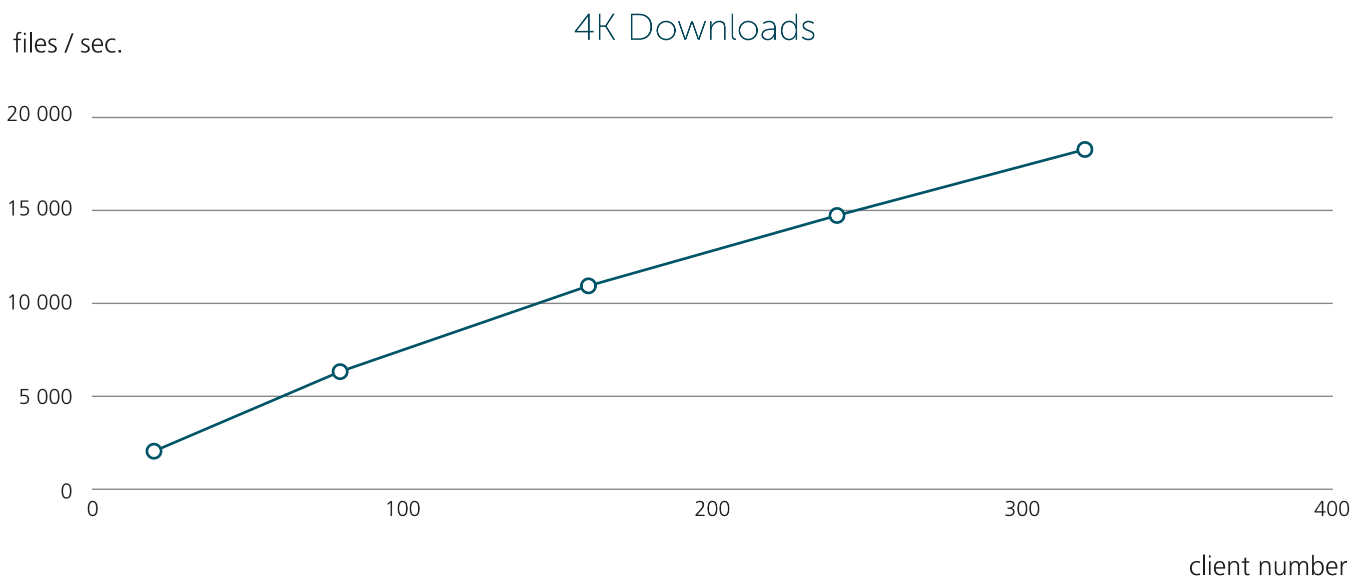
In order to evaluate the benefits of cloud storage in the environment of science data storage and analysis, a comprehensive test plan has been agreed and executed throughout the year focusing on scalability and performance. Hardware basis for the evaluation is a storage system consisting of 384 hard drives with associated low power processors interconnected in a highly redundant cluster which is connected via two 10 gigabit network connections to client nodes in the CERN computer centre.

Huawei's cloud storage system in the CERN data centre.

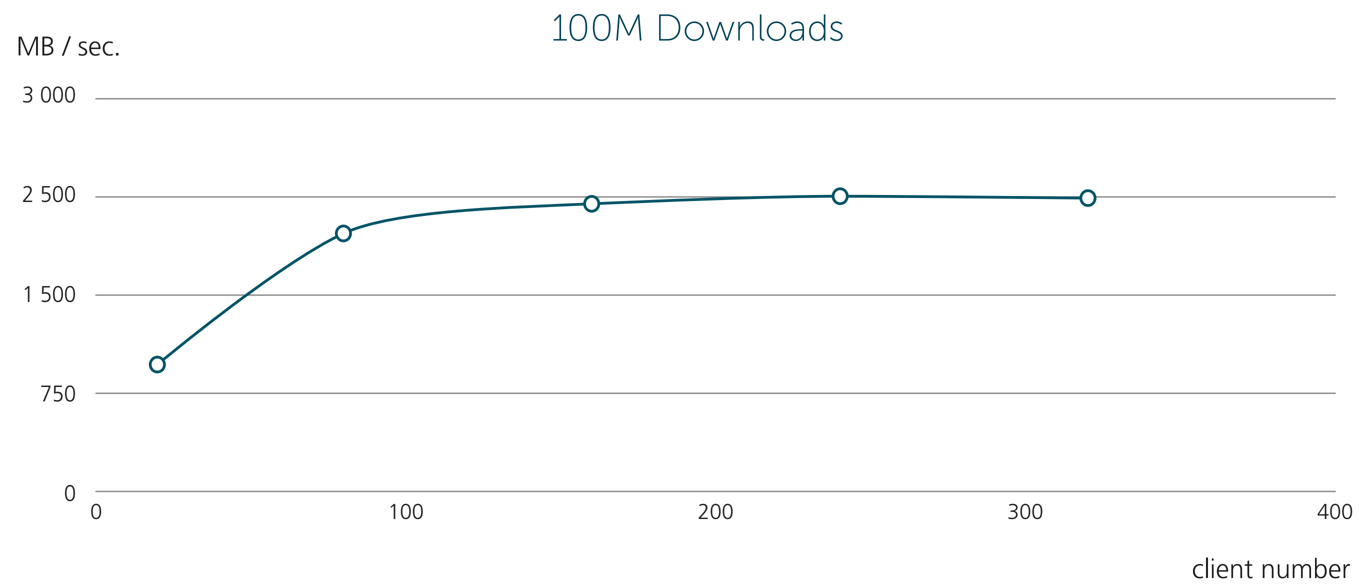
After the rapid commissioning and system testing at CERN by Huawei storage engineers in January, the joint tests quickly confirmed that the system was compliant to a S3 protocol test suite, which had been developed at CERN. The team then proceeded to evaluate the performance and scalability of individual product components such as the distributed meta-data storage, which are of crucial importance for data analysis use cases characterised by rapid access to a large number of different files. The project team confirmed the expected scalability up to 18,000 file downloads and 1,400 file uploads per second. A few smaller issues have been quickly identified and resolved via weekly conference calls with the main product architect and the help of on-site Huawei engineers.

In collaboration with the ROOT team at CERN, direct access to S3 based storage from production experiment applications could be demonstrated with satisfying first performance results, taking into account that cloud storage is not usually optimised for random access use cases. In a final testing phase the number of storage clients has been ramped up until the available network bandwidth of 20 gigabit per second was saturated serving 336 concurrent clients and the project team could confirm close to linear scalability with increasing number of clients and storage resources. CERN and Huawei are now considering scenarios for how to leverage the results achieved as part of a future joint programme.

Huawei's cloud storage system scales up to 18000 files per second.



21 clients reach bandwidth limit of 18 gigabit per second.



Education

Numerical Computing Workshop organised at CERN on 7-8 February 2012.

Building human capital

Knowledge is created through the evaluation of solutions as well as genuine research and development of IT technologies. It is disseminated through multiple channels.

The openlab education programme, which provides active dissemination, is currently implemented through several lines of actions. Workshops or seminars are regularly organised at CERN on advanced topics directly connected to the openlab projects. More than 390 participants attended the 11 topical workshops promoting the work of the competence centres. Most of these workshops have a special feature: they involve a mix of lecturers from both industry and CERN, thus exemplifying the CERN openlab principle of two-way knowledge transfer through active collaboration. Several of the workshops combine hands-off theory with hands-on practice.

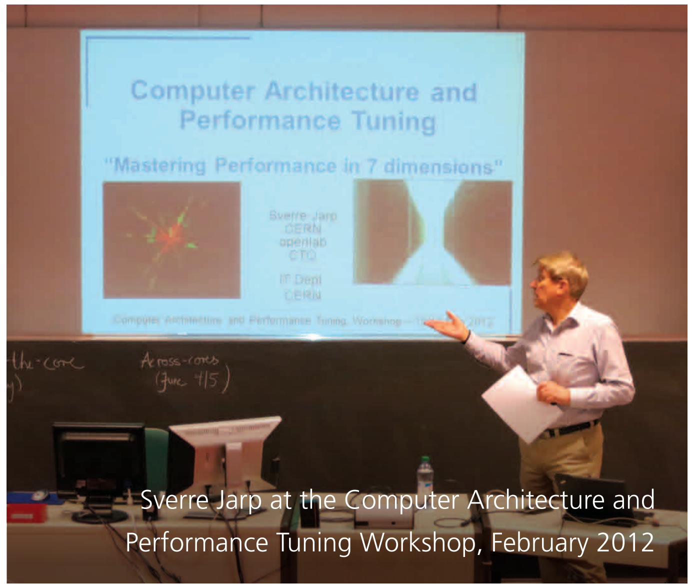
Special courses have also been organised this year for advanced CERN users, in addition to the regular quarterly courses. These classes touched on numerous topics, including data analytics, future tools and optimisations and were taught by Intel, Oracle and Siemens specialists. CERN openlab experts also contribute to off-site education activities such as the CERN School of Computing, where seven hours of lectures and exercises were delivered in 2012. These lectures and workshops are listed on page 44.

These direct training activities are complemented by the CERN openlab Student Programme, which itself is a genuine educational undertaking. This programme was launched in 2002 to enable undergraduate, Masters and Ph.D. students to get hands-on experience with Grid technology and other advanced openlab-related topics. A total of 162 young computer scientists have participated in this programme. In 2012, the programme accepted 15 computer science and physics students of 12 nationalities for two months, during the period June to September.

The students worked on cutting-edge computing technologies supervised by CERN openlab staff, other groups in the IT Department as well as staff from WLCG. Visits were organised to the CERN Data Centre, the CERN Control Centre, the ATLAS experiment, the LHC magnets test hall, the anti-matter factory and the PS-LINAC-LEIR accelerators.

In addition, the students toured Zurich and were given talks at ETHZ, Google, and Open Systems. A dedicated lecture series for the students was given on site by CERN experts. Several of this year's students were co-funded by CERN and Huawei, Intel, or Oracle. Full details are given on page 45.

CERN openlab results have been disseminated in a wide range of international conferences. These publications, presentations and reports can be consulted on the openlab website, as well as a large number of press articles coming out in the general press, IT-specific press and on the Web. The full list of presentations, publications, posters, and reports for 2012 is available on pages 46 and 47 of this annual report.



Sverre Jarp at the Computer Architecture and Performance Tuning Workshop, February 2012

CERN openlab Topical Workshops

- **CERN openlab/Intel - Winter 2012 Numerical Computing Workshop,** 7-8 February 2012, CERN, J. Arnold/Intel, M. Corden/Intel, F. Dinechin/ENS Lyon, J. Gustafson/Intel, V. Innocente/CERN, L. Moneta/CERN
- **CERN openlab/Intel Winter 2012 Computer Architecture and Performance Tuning Workshop,** 15-16 February 2012, CERN, J. Arnold/Intel, S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A.Nowak/CERN
- **CERN openlab/Intel Spring 2012 Multi-threading and Parallelism Workshop,** 4-5 June 2012, CERN, J. Arnold/Intel, S. Jarp/CERN, A. Lazzaro/CERN, A. Nowak/CERN
- **Intel/CERN openlab Workshop on Advanced Parallelism,** 3-4 July 2012, CERN, L. Akyil/Intel, H. Pabst/Intel, S. Jarp/CERN, A. Lazzaro/CERN, A. Nowak/CERN, L. Valsan/CERN
- **CERN School of Computing**, 13-24 August 2012, Uppsala, Sweden. Four lectures: Understanding Scalable Hardware, S. Jarp/CERN, Software that Scales the Hardware, S. Jarp/CERN, Key Aspects of Multi-threading, A. Nowak/CERN, Performance Optimization, A. Nowak/CERN
- **CERN openlab and Intel Numerical Computing Workshop**, 25-26 September 2012, CERN, J. Arnold/Intel, M. Corden/Intel, F. Dinechin/ENS Lyon, T. Mattson/Intel
- **Autumn School on Numerical Computing,** 1-2 October 2012, CASPUR consortium, Rome, Italy, J. Arnold/Intel, M. Corden/Intel, F. Dinechin/Ecole Normale Supérieure de Lyon (ENS), T. Mattson/Intel, S. Jarp/CERN, A. Nowak/CERN
- **ATLAS/CERN openlab/Google Tutorial on Advanced Performance Monitoring,** 17-18 October 2012, CERN, D. Levinthal/Google
- **CERN openlab/Intel Workshop on Parallelism, Compilers** and Performance, 30 October – 1 November 2012, CERN, J. Arnold/Intel, H. Pabst/Intel, S. Jarp/CERN, A. Lazzaro/CERN, A. Nowak/CERN, L. Valsan/CERN
- **CERN openlab Workshop on Data Analytics,** 16 November 2012, CERN, P. Andrade/CERN, J. Breysse/Oracle, G. Calabrese/Oracle, T. Cass/CERN, D. Cressy/Oracle, S. Jensen/CERN, M. Marquez/CERN, P. Merle/Oracle, M. Ragogna/Oracle, D. Santiago/Oracle, A. Voitier/CERN, H-G. Zimmermann/Siemens, R. Zimmerman/Oracle
- **Fourth International Workshop for Future Challenges in Tracking and Trigger Concepts,** 28-30 November 2012, CERN, Three lectures: A study of Compiler Flags and Performance Events, M. Botezatu/CERN, MIC (Xeon Phi) and openlab's experience, A. Nowak/CERN, Recent benchmarking with Sandy Bridge, L. Valsan/CERN

CERN openlab Summer Student Programme

CERN openlab Summer Student Programme Teaching Series, July-August 2012

- General Security, S. Lüders/CERN
- Software Security, S. Lopienski/CERN
- Web Application Security, S. Lopienski/CERN
- Linux Kernel Development, P. Sakkos/CERN
- Physics Computing, H. Meinhard/CERN
- Worldwide LHC Computing Grid Overview, F. Furano/CERN
- Control System Cyber-Security, S. Lüders/CERN
- Oracle Databases at CERN, E. Grancher/CERN
- Overview of the Grid Middleware, D. Smith/CERN
- Invenio Technology, T. Simko/CERN
- Size and Complexity of the CERN Network, E. Martelli/CERN
- Data Reliability at CERN and Ideas on How to Improve it, A. Pace/CERN

CERN openlab Summer Students 2012, with Nation State, Home Institute and Project Topic

- A. Andronidis, Greece, Aristotle University, Greece, "Puppet Firewall Module and Landb Integration"
- J. N. F. Batista, Portugal, University of Porto, Portugal, "Enhanced Web Interfaces for Administering Invenio Digital Library"
- A. Boulmier, Switzerland, HEPIA, Switzerland, "Media Structure Modifications in Invenio"
- D. Bukkapatnam, India, Birla Institute of Technology and Science, India, "Particle Navigation Algorithm Optimized for Use on GPU Architectures"
- Z. Chen, China, BarcelonaTech (UPC), Spain, "FTS Parameter Transfer Optimization"
- A. Gupta, India, JSS Academy of Technical Education, India, "Experimentation of Message Passing System in Oracle for the CASTOR Project"
- G. Jahn, Germany, Georg-August University, Physical Institute, Germany, "An Interactive Shell and a Python Module for DMLite"
- D. Jovanoski, University American College, Republic of Macedonia , "Detect Unpatched Web Applications"
- K. Kapusta, Poland, AGH University of Science and Technology, Poland, "Log Files Management"
- F. P. A. F. Pinto, Portugal, University of Porto, Spain, "A graphical Visualizer for Benchmark Data"
- J. D. Sadlo, Poland, Warsaw University of Technology, Poland, "Implementation of RotaTool, Action Management, DadEdit2: Three Tools to Automate the Work of the Database Administrator, the End User and the Application Developer, based on Oracle and Apex"
- J. K. Schingler, Canada, University of Maryland, USA, "Open Learning and Citizen Science"
- A. C. Siqueira, Brazil, Federal University of Campina Grande, Brazil, "Indico - Offline Event Storage"
- O. Smorholm, Norway, University of Bergen, Norway, "Vectorization of Software used for the Simulation of Physics"
- M. Umair, Pakistan, Royal Institute of Technology (KTH), Sweden, "Building a Benchmark Code and Result Repository Web Portal for CERN Storage"


A fraction of the CERN openlab summer students 2012 with Bob Jones, head of CERN openlab.

CERN openlab Presentations and Publications

Presentations

- A. Nowak/CERN, Computing with Accelerators - Opportunities for QCD studies, CERN PH-TH seminar, 24 February 2012
- A. Nowak/CERN, SSG Tools, Intel SSG visit at CERN openlab, 29 February 2012
- S. Jarp/CERN, J. Leduc/CERN, A. Nowak/CERN, An Evaluation of the Intel Xeon E5 Processor Family, Zurich Launch Event, 8 March 2012
- A. Nowak/CERN, Current and Future Developments in Commodity Computing and Many-core, Forum on Concurrent Programming Models and Frameworks, CERN, 25 April 2012
- A. Nowak/CERN, An Overview of Intel MIC - Technology, Hardware and Software, Many-core Architectures for LHCb, CERN, 25 April 2012
- A. Nowak/CERN, Many-core Experience with HEP Software at CERN openlab, CHEP 2012, New York City, USA, 22 May 2012
- A. Nowak/CERN, The Future of Commodity Computing and Many-core versus the Interests of HEP Software, CHEP 2012, New York City, USA, 24 May 2012
- B. Jones/CERN, The LHC Computing Model and its Evolution, IVI Summer Summit, National University of Ireland, Maynooth, 5 June 2012
- B. Jones/CERN, Oracle Global Research Strategy Meeting, 6 June 2012
- A. Nowak/CERN, Multithreaded Geant4 on KNC - Math and Profiling Overview, Intel-CERN technical meeting, CERN, 7 June 2012
- A. Nowak/CERN, The Growth of Commodity Computing and HEP Software, ATLAS Software and Computing Week, CERN, 11 June 2012
- B. Jones/CERN, The Worldwide LHC Computing Grid and its Evolution, ECIS 2012, Barcelona, Spain, 11 June 2012
- S. Jarp/CERN, A. Lazarro/CERN, A. Nowak/CERN, L. Valsan/CERN, Hybrid Parallelization of Maximum Likelihood Fitting with MPI and OpenMP, International Supercomputing Conference (ISC), Hamburg, Germany, 19 June 2012
- S. Jarp/CERN, The Extreme Scale: Data Analytics at LHC (the Large Hadron Collider), International Supercomputing Conference, (ISC), Hamburg, Germany, 20 June 2012
- S. Jarp/CERN, Clouds in Science - Preparatory Steps for LHC Computing, International Supercomputing Conference (ISC), Hamburg, Germany, 21 June 2012
- S. Jarp/CERN, A. Lazzaro/CERN, A. Nowak/CERN, L. Valsan/CERN, Hybrid Parallelization of Maximum Likelihood Fitting with MPI and OpenMP, HPC - Europa2, Amsterdam, Netherlands, 26 June 2012
- A. Nowak/CERN, The Performance Dimensions of PC Servers, The 10th IEEE International Symposium on Parallel and Distributed Processing with Applications (ISPA '12), Madrid, Spain, 10 July 2012
- A. Nowak/CERN, Performance Monitoring at CERN openlab, Open Seminar, Intel Exascale Lab, Paris, France, 20 July 2012
- A. Nowak/CERN, An Overview of Intel MIC - Technology, Hardware and Software, ATLAS Future Software Technologies Forum, CERN, 22 August 2012
- J. Batista/CERN, S. Jarp/CERN, B. Jones/CERN, A. Nowak/CERN, S. Stancu/CERN, F. Tilaro/CERN, M. Zotes Resines/CERN, The CERN openlab projects, IT Technical Forum, CERN, 31 August 2012
- T. Cass/CERN, Enabling and Supporting Data-Driven Research, Oracle OpenWorld, San Francisco, USA, 30 September – 4 October 2012
- K. Dziedziniewicz/CERN, E. Grancher/CERN, Efficient Database Cloning with Direct NFS and Clonedb, Oracle OpenWorld, San Francisco, USA, 30 Sept. – 4 Oct. 2012
- E. Grancher/CERN, D. Wojcik/CERN, CERN's MySQL "as a Service" Deployment with Oracle VM: Empowering Users, Oracle OpenWorld, San Francisco, USA, 30 September – 4 October 2012
- D. Wojcik/CERN, Implementing Oracle Enterprise Manager 12c to Manage Oracle VM, Oracle OpenWorld, San Francisco, USA, 30 Sept. – 4 Oct. 2012
- S. Skorupinski/CERN, Reliable Data Protection with Oracle RMAN and Oracle Data Guard, Oracle OpenWorld, San Francisco, USA, 30 Sept. – 4 Oct. 2012
- D. Savu/CERN, S. Stancu/CERN, CERN openlab Vision Project, Interop 2012, New York City, USA, 5 Oct. 2012
- D. Düllmann/CERN, CERN Cloud Storage Evaluation, HEPiX Fall 2012, Institute of High Energy Physics, Beijing, China, 16 October 2012
- A. Nowak/CERN, Collaboration, Big Data and the Search for the Higgs Boson, Intel European Research and Innovation Conference 2012, Barcelona, Spain, 23 Oct. 2012
- S. Jarp/CERN, Solving the Riddles of the Universe through Big Data, Big Data Warehousing and Business Intelligence 2012 Conference, Sydney, Australia, 26 Nov. 2012
- M. Limper/CERN, How to discover the Higgs Boson in an Oracle database, UK Oracle User Group (UKOUG) Conference 2012, Birmingham, UK, 4 December 2012
- L. Canali/CERN, M. Blaszczyk/CERN, Active Data Guard at CERN, UK Oracle User Group (UKOUG) Conference 2012, Birmingham, UK, 4 December 2012
- B. Jones/CERN, CERN: Big Science Meets Big Data, Big Data Analytics 2012, London, UK, 6 December 2012

Publications

- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, Many-core Experience with HEP Software at CERN openlab, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, May 2012
- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/ CERN, The Future of Commodity Computing and Many-core versus the Interests of HEP Software, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, May 2012
- F. Chatal/CERN, L. Gallerani/CERN, M. Piorkowski/CERN, A. Topurov/CERN, Management of Virtualized Infrastructure for Physics Databases, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, 21-25 May 2012
- E. Grancher/CERN, M. Piorkowski/CERN, A. Topurov/ CERN, Characterisation of HEP Database Applications, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, May 2012

Posters

- A. Nowak/CERN, Many-core Experience with HEP Software at CERN openlab, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, 21-25 May 2012
- A. Topurov/CERN, Management of Virtualized Infrastructure for Physics Databases, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, 21-25 May 2012
- E. Grancher/CERN, M. Piorkowski/CERN, A. Topurov/ CERN, Characterisation of HEP Database Applications, Computing in High Energy and Nuclear Physics (CHEP), New York City, USA, 21-25 May 2012

CERN openlab Reports

- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, Evaluation of the Intel Sandy Bridge-EP server processor, March 2012
- M. Botezatu/CERN, Benchmarking ROOT Whitepaper: Comparison of ICC Compiled Version vs GCC Compiled one. ROOT on Different Architectures, March 2012
- M. Zotes Resines/CERN, Initial Benchmark Results from the Huawei Storage System, March 2012
- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, L. Valsan/CERN, Parallelization of MLfit benchmark using OpenMP and MPI, July 2012
- D. Jovanoski/Summer Student, Detect Unpatched Web Applications, August 2012
- Z. Chen/Summer Student, FTS Parameter Transfer Optimization, August 2012
- A. Gupta/Summer Student, Experimentation of Message Passing System in Oracle, for the CASTOR Project, Aug. 2012
- A. C. Siqueira/Summer Student, Indico - Offline Event Storage, August 2012
- A. Boulmier/Summer Student, Media Structure Modifications in Invenio, August 2012
- J. Sadlo/Summer Student, Implementation of Rota Tool, DAD Edit and Frame for Account Actions Management in Oracle APEX, August 2012
- A. Andronidis/Summer Student, Puppet Firewall Module and Integration, August 2012
- G. Jahn/Summer Student, An Interactive Shell and a Python Module for DMLite, August 2012
- K. Kapusta/Summer Student, Log Files Management, August 2012
- J. Batista/Summer Student, Enhanced Web Interfaces for Administering Invenio Digital Library, August 2012
- F. Pinto/Summer Student, A Graphical Visualizer for Benchmark Data, September 2012
- S. Jarp/CERN, A. Lazzaro/CERN, A. Nowak/CERN, L. Valsan/CERN, Comparison of Software Technologies for Vectorization and Parallelization, September 2012
- M. Botezatu/CERN, A Study on Compiler Flags and Performance Events, November 2012
- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, L. Valsan/CERN, Evaluation of the Intel 4 socket Sandy Bridge-EP server processor, December 2012

Events and outreach

The 12 pre-college students who won the CERN Special Award at the Intel International Science and Engineering Fair (ISEF) 2012, in Pittsburgh, USA, accompanied by CERN representatives, Jan Iven and Wolfgang von Rüden.

Creating and disseminating knowledge

The CERN openlab framework gives CERN and its sponsors a means to share a vision of 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 partners, through joint workshops and events, as well as to disseminate this to a wider audience, including partner clients, the press and the general public.

Top delegations from governments and industry regularly tour CERN. In 2012, 169 protocol visits were organised at CERN. The CERN openlab concept as well as the CERN openlab sponsors' projects are systematically presented to the guests visiting the CERN IT department.

In addition, customer and press visits are organised by CERN openlab partners. These groups are briefed about CERN openlab in a dedicated VIP meeting room known as the CERN openlab openspace.

Our Sponsors' top management come to visit CERN openlab to discuss common projects with the team. As an example, we enjoyed this year the visit of Christos Georgiopoulos, Vice President of the Software and Services Group and General Manager of the Developer Relations Division at Intel, and of David Perlmutter, Executive Vice President, General Manager Architecture Group and Chief Product Officer at Intel.

As part of their visit, our guests also give computing seminars in the IT auditorium, which are sometimes webcasted (www.cern.ch/webcast) and made accessible later on via the CERN Document Server (CDS) and the CERN openlab website. Seven of these talks were given in 2012 on topics as diverse as cloud storage, high-throughput x86 computing, computer 'speed' definition, the Blue Brain project, parallel programming, exploitation of the multicore parallelism and auto-tuning potential, and embedded specialisers to turn patterns-based designs into optimised parallel code.

Rolf Heuer and Loïc le Guisquet cut cake to celebrate 30 years of collaboration between CERN and Oracle.

Intel Visit to CERN, Christos Georgiopoulos, Vice President of the Software and Services Group and General Manager of the Developer Relations Division, Intel, February 2012

Intel ISEF students Visit to CERN, Award winners: Saumil Bandyopadhyay, Benjamin Craig Bartlett, Linn Bieske, Anirudh Dasarathy, Valerie S. Ding, Viola Mocz, Mingu Kim, Akash Krishnan, Henry Wanjune Lin, Christopher Stephen Nielsen, Nicholas Benjamin Schiefer, Connor Everett Tom. 17 June to 22 June 2012

Google students Visit to CERN, 4 July 2012
Introduction to CERN openlab and visit of the CERN facilities.

HP Danish Customers Visit to CERN, 11 September 2012
Introduction to CERN openlab and to the joint projects with HP as part of the CERN facilities' visit.

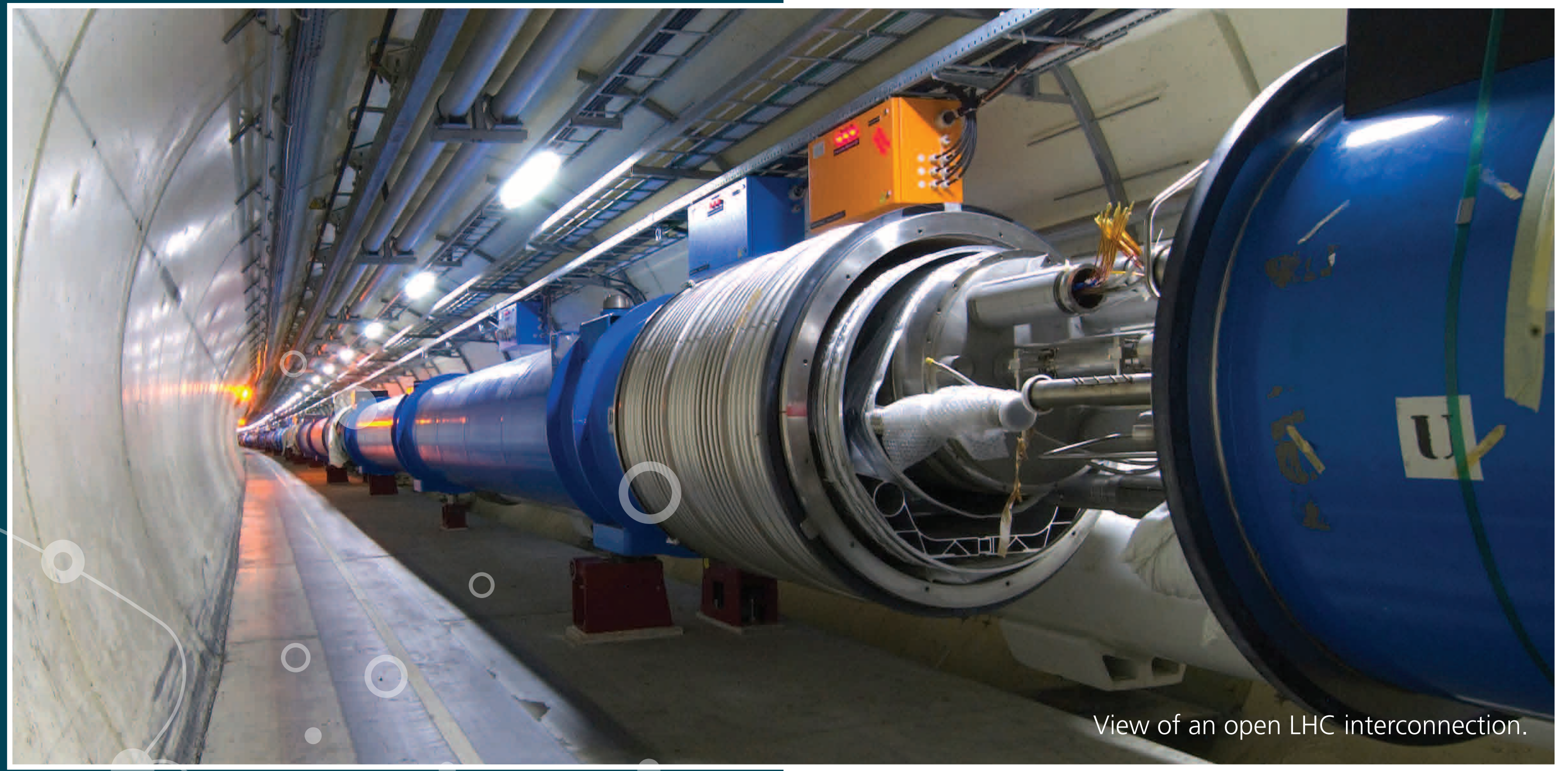
Intel Staff Visit to CERN, 9 October 2012
Introduction to CERN openlab and to the joint projects with Intel as part of the CERN facilities' visit.

Intel Visit to CERN, Claudio Bellini, Intel Partner Liaison with CERN, David Perlmutter, Executive Vice President, General Manager Architecture Group and Chief Product Officer, Intel, Neil Green, Director Worldwide Sales and Operations Group, Intel, Stephan Gillich, Director HPC and Workstation EMEA, Intel, 19 October 2012

Celebration of the 30th Birthday of the CERN-Oracle collaboration, CERN, 1 February 2013
The last session of the `IT requirements for the next generation research infrastructures' workshop was dedicated to the celebration of the 30th anniversary of the CERN-Oracle relationship. In recognition to the longstanding collaboration which started in 1982 and to the fruitful contribution of Oracle to the CERN scientific programme, Rolf Heuer, CERN director general, handed over an award to Loic Le Guisquet, Executive Vice President, Oracle Europe, Middle East, and Africa. Oracle has been part of CERN openlab since 2003.

Seminars by CERN openlab guests

- **An Overview of the CERN openlab – Huawei Collaboration,** CERN, 17 January 2012. Given by James Prescott Hughes, Cloud Storage Chief Architect, Huawei, as part of his visit.
- **High-Throughput x86 Computing (TFLOPS on a Chip),** CERN, 6 February 2012. Given by Herbert Cornelius, WW HPC Solution Architect, Intel, as part of his visit.
- **Defining Computer 'Speed' - An Unsolved Challenge,** CERN, 9 February 2012. Given by John Gustafson, Director, Intel Labs Santa Clara, California, as part of his visit.
- **The Blue Brain Project: Following CERN's Example,** CERN, 15 May 2012. Given by Henry Markram, Founder & Director - Brain Mind Institute, EPFL, Lausanne, Switzerland, as part of his visit.
- **Writing Parallel Programs that Work,** CERN, 8 June 2012. Given by Paul Petersen, Senior Principal Engineer, Software and Solutions Group (SSG), Intel, as part of his visit.
- **Dealing with BIG Data - Exploiting the Potential of Multicore Parallelism and Auto-Tuning,** CERN, 23 July 2012. Given by Victor Pankratius, Scientist at MIT, as part of his visit.
- **SEJITS: Embedded Specializers to Turn Patterns-based Designs into Optimized Parallel Code,** CERN, 28 September 2012. Given by Tim Mattson, Principal Engineer, Intel, as part of his visit.



View of an open LHC interconnection.

The future

Exciting times ahead

Work for the LHC's first long shutdown and new CERN openlab projects get under way.

The LHC's performance has been fantastic and since the first physics run in 2009 it has provided physicists with a huge quantity of data to analyse. After a short run in 2013, the LHC will be shut down for two years to prepare the machine for operation at higher energy. During this first long shutdown called 'LS1', it will be upgraded as well as renovated.

To cover the entire 27 km circumference of the LHC, no less than 10,170 high-current splices between the superconducting magnets will be consolidated. The teams will open the 1,695 interconnections between each of the cryostats of the main magnets and will repair and consolidate around 500 interconnections simultaneously. Sensitive electronic equipment protection will be optimised too by relocating the equipment or by adding shielding.

However the work will not be confined to the LHC. For example, the whole tunnel ventilation system of the Proton Synchrotron (PS) will be dismantled and replaced. Meanwhile, at the Super Proton Synchrotron (SPS), about one hundred kilometres of radiation-damaged cables used in the instrumentation and control systems will be removed or changed. CERN will also take advantage of LS1 to improve the installations connected with the experiments, the accelerators, the electronics, and the computing systems, with a view to a spectacular resumption of its main activities after the shutdown.

Therefore, the planned shutdown period matches well the timing of CERN openlab, where new solutions are being prepared now to be readily available for a possible implementation before the restart of the LHC at the end of the CERN openlab Fourth phase and of LS1. In this exciting context and based on the successful outcome of the benchmark performance evaluation, Huawei and CERN decided to upgrade their engagement from contributor to full partnership for the three upcoming years, while Yandex is to join as an associate to carry out research and development in the field of Advanced Event Processing.

First annual report of CERN openlab phase IV
Editor: Mélissa Gaillard
Photography: CERN openlab Platform Competence Centre – Alfio Lazzaro and Andrzej Nowak (p 37, p 38, p 43, p 44), CERN Photo Club - Fons Rademakers (p 11), CERN Photo Service – Maximilien Brice and Samuel Morier-Genoud (p 5, p 12, p 18, p 24, p 30, p 36, p 50, p 53), CMS Experiment (p 7), Sylvain Fasel (p 29), Mélissa Gaillard (p 45), Huawei (p 40), HP (p 31, p 33), Intel (p 35), Veronika McQuade (p 23), Shutterstock (p 1, p 2, p 3, p 4, p 6, p 10, p 16, p 22, p 28, p 34, p 42, p 48, p 52, p 56), Siemens (p 17), Society for Science & the Public (p 49)
Production: CERN Graphic Design Service – Fabienne Marcastel, Freelance – André-Pierre Olivier
ISBN 978-92-9083-385-7

CERN openlab

www.cern.ch/openlab