CERN openlab

Annual Report
May 2009 – May 2010

Contents

Cover illustration: Successful restart of the LHC with beam circulating on 20 November 2009 in the CERN Control Centre

Second annual report of CERN openlab – Third Phase
Editor: Mélissa Gaillard
Photography : CERN Photo Service Maximilien Brice (cover, p 1, 12), Luigi Gallerani (p 8, 9, 16), Andreas Hirstius (p 21), Michael Hoch (p 22), Vlad Lapadatescu (p 15), Claudia Marcelloni (p 4, p 24), Andrzej Nowak (p 20)
Production: André-Pierre Olivier
ISBN 978-92-9083-345-1

LHC and CERN openlab: moving towards new territory

Since May 2009, the year has been marked by an abundance of anniversaries, crucial achievements, world-records and the move to new territory for both CERN and CERN openlab.

Taking a look at history is a good way to prepare for the future and CERN had the chance this year to celebrate a great number of important anniversaries. It was indeed 50 years since the Proton Synchrotron circulated its first beam and 20 years since LEP, the LHC predecessor, switched on. Thirty years ago, electroweak theory made its appearance on the Nobel map with the award of the physics prize to Sheldon Glashow, Abdus Salam and Steven Weinberg. Five years later, Carlo Rubbia and Simon van der Meer brought the prize to CERN for the first time. Their prize marked the discovery of W and Z particles, communicators of the weak interaction. And 10 years ago, Gerardus 't Hooft and Martinus Veltman received the prize for their pioneering work on elucidating the quantum structure of electroweak interactions. On 3 and 4 December 2009, no fewer than 13 Nobel Prize winners shared their recollections with us, along with people who played key roles in developing CERN's accelerator complex.

Above all, this year coincided as well with the start of a new era for particle physics and the success of a fantastic team effort. Nearly 20 years of hard work by thousands of people have made the Large Hadron Collider a dream come true. On 20 November, the LHC circulated its first beam of 2009, ushering in a remarkably rapid beam-commissioning phase. The first collisions were recorded on 23 November and the world-record beam energy was established on 30 November, which ended the 2009 run on a high note. The first high-energy collisions were successfully achieved on 30 March 2010 at all four points of the ring with the LHC first physics marking the entrance to a new physics era. The new data will give us an unprecedented tool to understand the Universe we live in. There is a real chance over the next two years of discovering supersymmetric particles and possibly giving insights into the composition of about a quarter of the Universe.

CERN openlab also entered the first year of its third three-year phase in 2009. At that time, I welcomed the industrial partners to this new phase and expressed my full support to the CERN openlab which had proven to be a very successful framework for industrial collaboration and an outstanding catalyst for ideas. As outlined in this report, the CERN openlab partnership continued this year to bear fruit and to have a direct and positive impact on the development of computing services that underlie the LHC. Crucial areas have also been tackled with the first results of the Automation and Controls Competence Platform and the start of a new project in the wireless area.

I therefore thank all CERN openlab partners for their continued support of our joint effort. I look forward to the future shared benefits of this unique collaboration, as we have now definitely entered the LHC era, a new age of fundamental physics and challenges

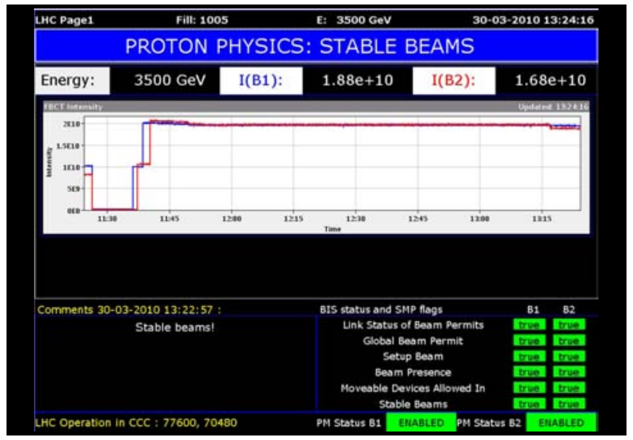
Rolf Heuer
Director-General of CERN

THE CONTEXT

First physics at the LHC

The Large Hadron Collider (LHC) project started more than 20 years ago with the aim of preparing the next major phase in the ongoing quest for a deeper understanding of the fundamental laws of nature. Now that it is completed, 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 100 m underground, it accelerates particles to energies never reached before. Some 5000 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 3D cameras with 100 megapixels, will record these 'mini Big Bangs' at up to 40 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 a bit more than a year following a severe incident due to a faulty magnet interconnect. After the repair and the installation of additional protection systems, the accelerator started operation again on 20 November 2009 with the first circulating beam at the injection energy of 0.45 TeV. Milestones were quickly passed, with twin circulating beams established by 23 November and a world record beam energy of 1.18 TeV being set on 30 November. By the time the LHC switched off for 2009 on 16 December, another record had been set with collisions recorded at 2.36 TeV and significant quantities of data taken. After a short technical stop, beams were again circulating on 28 February 2010, and the first acceleration to 3.5 TeV occurred on 19 March. On 30 March 2010, beams collided in the LHC at 7 TeV, the highest



30 March 2010, CERN Control Centre screen monitoring the two 3.5 TeV beams used for the first high-energy physics collisions

energy ever achieved in a particle accelerator, marking a new world record and the start of the LHC research programme.

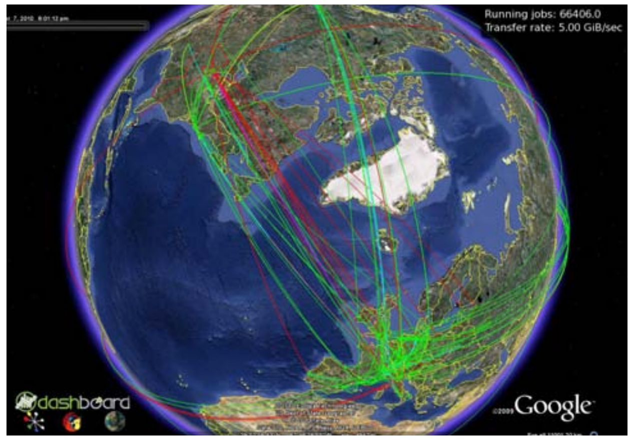
The ALICE, ATLAS, CMS, and LHCb experiments observed and immediately recorded beautiful events in their detectors, giving rise to explosions of joy in the CERN Control Centre, in the experiments' control rooms and around the world, where thousands of scientists and many millions of observers followed live the first collisions of 3.5+3.5 TeV. The accelerator and the experiments' computing systems, as well as the Worldwide LHC Computing Grid operated perfectly, enabling the physicists to share the results instantly and to publish their first scientific papers promptly.

20 November 2009, explosion of joy in the ATLAS detector control room for the successful LHC restart and the first beam splash detected by the experiment

Data taking, sharing and analysing

The new phenomena that scientists hope to find are extremely rare, hidden deeply in already known physics. The LHC has therefore been designed to produce a very high rate of collisions (40 MHz) such that the 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 per second. None of today's computing systems is capable of recording such rates. Sophisticated selection systems, called first level trigger, allow most of the events to be rejected after one millionth of a second, followed by a second level of selection applying more sophisticated criteria such that the remaining data rate per experiment is below one gigabyte per second. Even after such a drastic data reduction, the four big experiments, ALICE, ATLAS, CMS and LHCb, will produce over 15 million gigabytes per year, corresponding to a stack of CDs about 20 km tall.

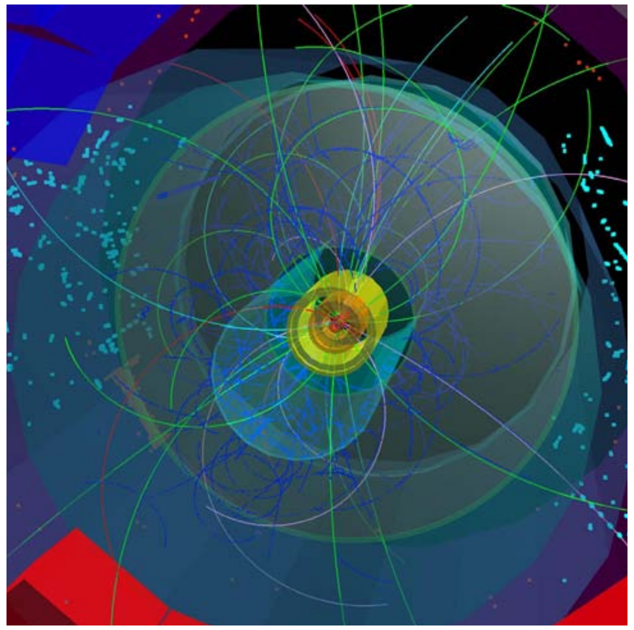
Thousands of scientists around the world have contributed to constructing the sophisticated LHC experiments and they are now eagerly waiting to get their hands on the data to extract the physics during the next fifteen years, the expected lifetime of the LHC. To reach this goal, tens of thousands of 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, connecting and combining the IT power of more than 200 computer centres in 34 countries. CERN is the central 'hub' of the WLCG, the Tier-0 centre. A first copy of all data from the LHC experiments will be held here, where also the first reconstruction and data



WLCG Grid Activity on 7 April 2010

quality assurance take place. This central hub, which provides less than 20% of the total compute capacity, is connected to eleven other major computing centres (Tier-1) and Grid services using dedicated 10 gigabit per second optical private networks. 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 make 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, grouped in Federations, 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.

To rise to such unprecedented computing challenges, new and advanced systems were needed. It required 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, keep collaborating on new solutions with success.



The first high-energy collision recorded by the ALICE detector on 30 March 2010

THE CONCEPT

Catalysing collaboration

The research carried out at CERN often requires technologies which are not available off-the-shelf and need to be specifically developed. Cryogenic systems, superconductivity and vacuum technologies, which are essential to the LHC, were one of those.

In many circumstances, the technologies benefit from being jointly developed and produced with industrial partners. The same applies for Information Technology (IT). This is why, eight years ago, CERN openlab was established as a framework for multilateral projects between CERN and the IT industry.

Within this structure, 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. The openlab partners commit to a three-year programme of work and provide three possible kinds of funding: salaries for young researchers, products and services, and engineering capacity.

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 is now entering the second year of its third phase (2009–2011). It not only capitalises on but also extends the successful work carried out in openlab-II with a similar aim of hosting a few major projects with a particular focus on technologies and services relevant to CERN. The technical activities are organised in four Competence Centres (CC): the Automation and Controls CC, the Database CC, the Networking CC, and the Platform CC.

Besides technical achievements, which are amply described in this report, a number of joint workshops and press events were organised. Disseminating the expertise and knowledge created in openlab is one of the key objectives. Regular training sessions took place throughout the year, including openlab contributions to the CERN School of Computing. The CERN openlab Summer Student Programme with its specialised lectures is another example matching this objective.

In conclusion, it should be highlighted that none of this year's achievements would have been possible without the commitment of our partners. This report is also an opportunity to commend HP, Intel, Oracle, and Siemens for their strong support and engagement in this ambitious partnership which led to remarkable results this year again.

CERN openlab Management Unit
From left to right: Séverine Pizzera, Mélissa Gaillard, Wolfgang von Rüden, Sverre Jarp, François Flückiger

The openlab team: open minds sharing R&D projects

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

A list of the IT and EN department staff most closely involved in the CERN openlab activities is given below. The distributed team structure permits close collaboration with computing experts in the LHC experiments, as well as with engineers and scientists from openlab partners who contribute significant efforts to these activities. Principal liaisons with partners and contributors are listed. In addition, significant contributions are made by students participating in the CERN openlab student programme, both directly to openlab activities (6 students during summer 2009) and more widely to WLCG, EGEE and other Grid and CERN related activities in the IT department (14 students).

CERN openlab Board of Sponsors

Rolf Heuer	CERN (Head of Board)
Wolfgang von Rüden	Head of CERN openlab
Frédéric Hemmer	Head of CERN IT Department
Bill Johnson	HP
Richard Dracott	Intel
Juan F. Rada	Oracle
Thomas Hahn	Siemens

CERN openlab Management Unit

François Flückiger	CERN openlab Manager
Sverre Jarp	Chief Technology Officer
Mélissa Gaillard	Communications Officer
Séverine Pizzera	Administrative Assistant

CERN openlab Fellows and Staff (sponsor indicated)

Zbigniew Baranowski	Fellow (Oracle)
Axel Busch	Technical Student (CERN)
Carlos Garcia Fernandez	Fellow (Oracle)
Milosz Hulbój	Staff (HP)
Ryszard Jurga	Staff (HP)
Omer Khalid	Fellow (Siemens)
Vlad Lapadatescu	Fellow (HP)
Alfio Lazzaro	Fellow (Intel)
Julien Leduc	Fellow (Intel)
Ildefons Magrans de Abril	Fellow (Siemens)
Andrzej Nowak	Staff (Intel)
Svetozar Kapusta	Fellow (Oracle)
Daniel Filipe Rodrigues	Staff (Siemens)
Filippo Tilaro	Fellow (Siemens)

CERN Liaison

Renaud Barillère, Liaison with Siemens
Sverre Jarp, Liaison with Intel
Jean-Michel Jouanigot, Liaison with HP
Mats Möller, Tony Cass, Liaison with Oracle

Other IT and EN Departments Staff contributors to CERN openlab

Brice Copy	ICE Group
Eva Dafonte Pérez	DB Group
Dirk Düllmann	DSS Deputy Group Leader
Maria Girone	ES Deputy Group Leader
Manuel Gonzales Berges	ICE Group
Eric Grancher	DB Deputy Group Leader
Andreas Hirstius	CS Group
Chris Lambert	DES Group
Jose Carlos Luna Duran	CS Group
Alberto Pace	DSS Group Leader
Anton Topurov	DB Group
Artur Wiecek	DB Group

Industry Partner Liaisons with CERN

Dan Ford	HP
Stéphane Laroche	HP
Alberto Soto	HP
Claudio Bellini	Intel
Herbert Cornelius	Intel
Stephan Gillich	Intel
C. Gregory Doherty	Oracle
Monica Marinucci	Oracle
Ulrich Eger	Siemens
Thomas Hahn	Siemens
Juergen Hirte, Juergen Bieber	Siemens

THE RESULTS

The Automation and Controls Competence Centre

Siemens provides a large set of solutions for Automation and Controls and joined the third phase of CERN openlab as a partner, enriching the framework activities portfolio with a new dimension and giving birth to the Automation and Controls Competence Centre. The CERN environment is highly reliant on industrial control systems such as control actuators, remote profibus Input/Output modules, Programmable Logic Controllers (PLCs) and SCADA systems. The collaboration that has been running for one year focuses on security, opening automation tools towards software engineering, and handling large environments.

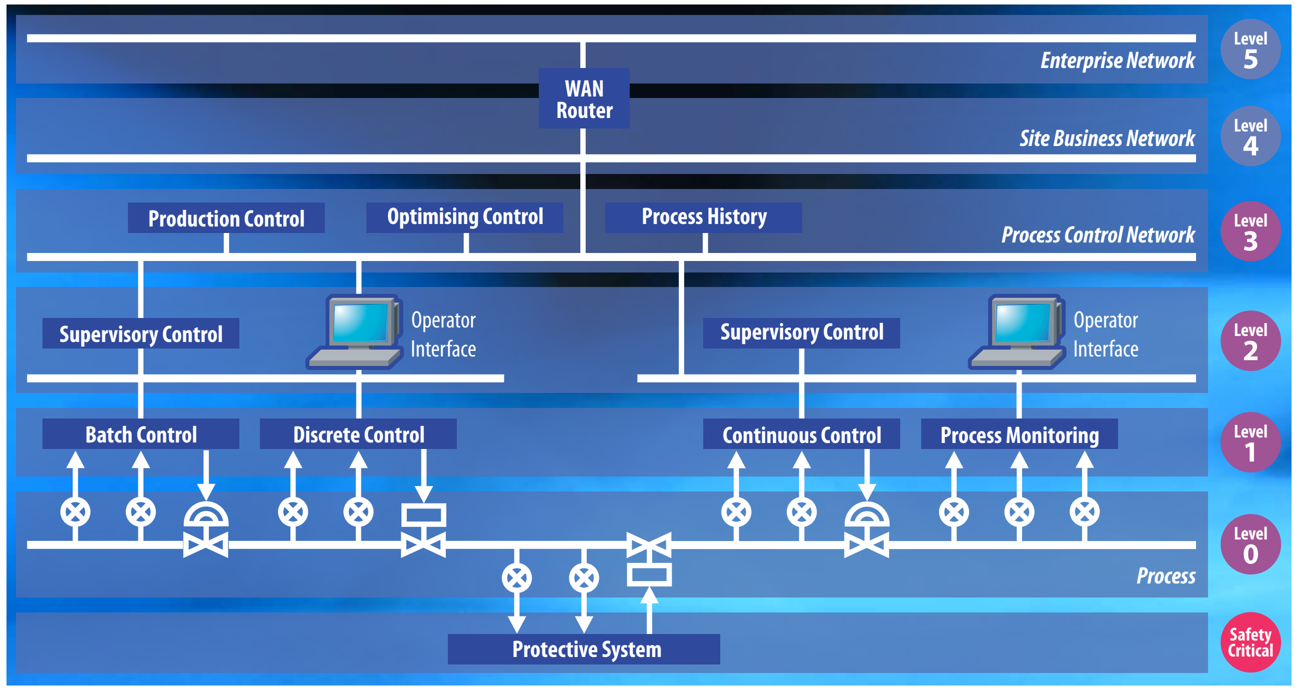
PLCs robustness under test

CERN needs to define approaches for achieving the dual goals of connecting its operational network to the Internet while at the same time keeping its industrial control systems secure from external and internal attacks. With this in mind, the ISA-99 international cyber security standard was adopted as a reference model to define a set of implementation guidelines and a list of security robustness criteria applicable to any network device. Device security represents a key link in the defense-in-depth concept (See Figure ISA reference model for the Distributed Control Systems), as some attacks will inevitably penetrate security boundaries and thus require further protection measures.

Since no solutions based on security standards are available on the market, a wide investigation of the current cyber security testing techniques and systems has been performed in the first phase of the project. The team reviewed the Wurldtech's "Achilles Satellite" product, a powerful testing platform, and successfully used it to analyse and evaluate the effectiveness of different testing techniques. However, to overcome this platform's proprietary aspects and limitations in terms of supported network protocols and attack techniques customisation support, the Test-bench for Robustness of Industrial Equipments ("TRoIE") was designed and implemented. TRoIE aims at discovering possible PLC vulnerabilities through Ethernet communications. Such tests must not be confused with functional testing, where only valid operations are performed to cover all the possible "not malicious, but operating" scenarios. On the contrary, it is necessary to detect possible anomalies arising from an incorrect handling of corrupt communication channels. This approach has already been proven by the valuable findings obtained during the analysis of Siemens S7 PLC ranges. Thanks to this analysis, it was possible to report critical anomalies in the software stack to Siemens and directly contribute to improving the security level and robustness of their PLCs. These initial encouraging results have motivated the team to continue following and expanding this approach for the future of the openlab collaboration.

Software engineering

The process visualisation and control system PVSS is used at CERN for large distributed control systems, some with more than 150 computers. In the first year of collaboration between Siemens/ETM (a subsidiary of Siemens) and CERN, the PVSS project focused on two main areas. First, learning PVSS in depth. Second, starting to use the acquired knowledge for improvements and testing of upcoming features at CERN. To attain these objectives, the team concentrated on four main activities. The Oracle Archiver is a solution in PVSS for storing and retrieving historical data. CERN has been one of the main users of this feature, requiring an excellent performance and stability when faced with large volumes of data and high update frequencies. As a learning task, a number of critical issues and improvements raised by CERN users were solved. The result was the incorporation of code written at openlab in a subsequent patch of PVSS released by ETM.



Users of PVSS at CERN use SVN (a version control system) as a revision control system for their projects. The PVSS development environment did not have SVN integrated, forcing users to control their project revisions using external tools. Within the openlab framework, a SVN plugin for the development environment was prototyped and then used as a showcase to other ETM clients.

The PVSS version reporting tool (PVR) is a new tool that is now available to ETM, and in use at CERN. Typically, users requiring support when using PVSS report their problems through e-mail or issue tracking systems. However, a lot of systematic information is very often required and not sent manually when using this report mechanism. Using the PVSS Version Report Tool instead, a lot of typical information, which can be as simple as the Operating System or PVSS version number or free disk space and project dependencies, is automatically sent. This provides a much more efficient mechanism for reporting.

One of the constant feature requests by CERN users is Web Support in PVSS. ETM provided a Web plugin with the 3.9 release, which has still not been adopted at CERN. In order to test it for possible future use, and provide a testing ground to ETM with the usual CERN extreme requirements, the Web plugin was tested in the openlab scope, tackling both functional and performance aspects of the novel feature. A full report with analysis of issues of concern was sent back to ETM, providing information on aspects to be improved before adoption at CERN.

Step7 Openness and Deployment

The Step7 Openness and Deployment project is sub-divided into two major topics. The first topic focuses on the issues related to the deployment of Step7 in large-scale environments in an automated way. The second one concentrates on the issue of bringing in software engineering concepts and capabilities to enhance the features of the Step7 software stack.

In 2009/2010, deployment of Step7 in large-scale environments was selected as first priority. Step7 installation software is complex and requires a dedicated installer to manage various phases for installation and to dynamically configure the target machine in the post-installation phase. Allowing the system administrator to deploy, maintain, and upgrade installations of Step7 software both remotely and in an automated fashion on a set of target machines was thus a challenging objective.

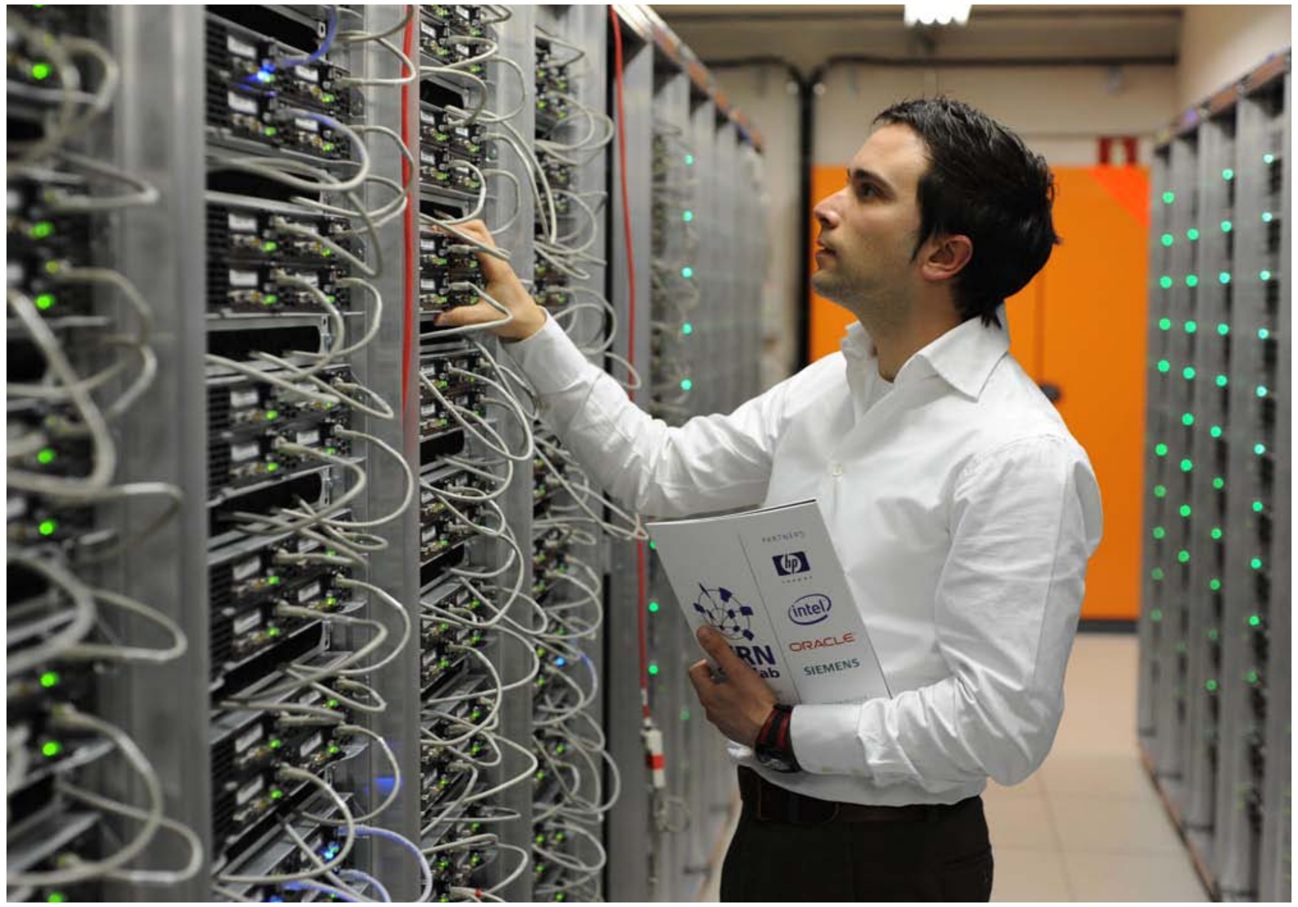
The team first conducted a CERN-wide survey of PLC developers to learn about their deployment scenarios and the possible solutions they might have adopted to achieve this. Secondly, analysis and evaluations of off-the-shelf tools for software inventory management were conducted, and CFEngine was selected as a test candidate to deploy and automate the installation on a set of target machines. Then, various deployment strategies were evaluated spanning both short-term, medium-term and long-term durations to keep in line with the software development life-cycle of version 11 of Step7. The team developed the prototypes for each of these strategies and identified not only pros and cons of each approach, but also highlighted their impact on Step7 architecture and design to fully utilize the benefits of software deployment tools. All these design changes were reported to Siemens. Finally, a strategy using Siemens Installer engine was selected as a way forward to automate the deployment of Step7 software that fits both with Siemens' short-term and long-term goals. The fact that this strategy was approved by Siemens to be part of their next version of Step7 software is one of the fruitful outcomes of this collaboration.

The Database Competence Centre

In the Database Competence Centre, a major focus of work during the period May 2009 to May 2010 has been the database release 11.2, also called 11gR2. This new Oracle release introduces fundamental changes which are of major interest for CERN database usage. Studies were carried out on some of the most innovative areas: Oracle Streams, Oracle Automatic Storage Management, Oracle Active Data Guard and Oracle Advanced Compression. In the past months, new techniques have been developed for Streams deployment, and virtualisation has been a major area of progress, with close collaboration with the Oracle Weblogic-VE team. Finally, significant joint work was done on monitoring.

Oracle Database 11g Release 2

Within the CERN openlab collaboration, several tests have been performed in the Oracle Streams environment which is used for metadata replication from Tier-0 to Tier-1 databases as part of the WLCG framework. These tests were focused on the overall performance of dataflow and new features introduced by Oracle in the 11g release 2 version. The results confirmed that the new concept of Oracle Streams architecture called Combined Capture and Apply provides a ten times more efficient replication than the previous one. After fruitful collaboration with the Oracle Streams project team, it was possible to replicate roughly 40 000 record changes per second, a rate never achieved before. Snapshots, process dumps and test results were sent to Oracle for further improvement of the product. Besides a more than satisfactory data throughput performance, the new release of Oracle Streams provides features increasing the availability of the replication. Intelligent management of multi-destination replication helps to avoid bottlenecks and handles downtimes of replicas. Administration of Oracle Streams is simplified by introducing a new set of packages for replication management and monitoring. For data consistency resolution, Oracle came up with a new feature called Compare and Converge. It enables one to perform comparisons of data objects from primary and replica databases. When inconsistencies are found, the administrator can fix them by running the converging procedure. The results of the performance tests show that the comparison of completely inconsistent tables can result in the processing of two megabytes of data per second on average. When the data are consistent, it is eight times faster (i.e. 16 megabytes per second). The convergence of data has an average speed of four megabytes per second. Despite some limitations, the Compare and Converge package is a promising approach for the resynchronisation of data using direct access to the data files of replicated schemas. It might be useful in the case of unrecoverable failures of media at the primary databases. Therefore, CERN is really looking forward to taking advantage of the Oracle Database 11g Release 2 whose deployment to the production stage is foreseen for next year. In addition, all streams replication monitoring tools, such as Oracle Enterprise Manager and in-house CERN monitoring, have been updated and tested with the latest Oracle database version in order to be ready for the migration.



The new features of Oracle Automatic Storage Management and Oracle ASM Cluster File System (ACFS) have been tested. File system tests were conducted using either local disks (RAID 1 – 500 gigabytes SATA) or SAN storage (three storages over four gigabit Fibre Channel - dual channel with multipathing - 16 SATA 400 gigabytes disks each). A number of file systems were created: ext3 on local disk, ext3 and ext2 on one ASM Dynamic Volume Manager volume, and ACFS shared between two nodes. A number of file operations were tested and the time needed compared: large file creation and deletion, parallel access, small file creation and deletion as well as archive extraction. The results of the tests were shared with Oracle. Oracle ACFS is much faster than ext3 with comparable or less CPU usage for most operations.

Oracle Data Guard is a key technology to achieve high availability. It belongs to Oracle's Maximum Availability Architecture (MAA) best practices. A Data Guard configuration consists of a primary database and of one (or more) standby database running on different resources. If the primary database goes down, the standby database can be activated as the new primary database within a few minutes, thus significantly decreasing the planned and unplanned downtime. In 2009, all critical databases of the LHC experiments deployed a Data Guard setup and some even benefited from activating the standby database as the primary, hence minimizing the impact on the service. Based on the very positive experience with Oracle Data Guard, the LHC experiments are very much looking forward to using the Oracle Active Data Guard feature available in 11g, which was thoroughly tested in openlab. When using Oracle Active Data Guard, the physical standby database can be opened in read-only mode and the recovery process restarted, so the standby will be in sync with the primary and can be used for reporting. This, with the option to take fast backups from the standby, can reduce the load on the primary considerably. The creation of a standby database is easier in 11g thanks to the new Oracle Recovery Manager command which was successfully tested. Real-Time Query performance and long-term stability proved to be very encouraging. The configuration, maintenance and monitoring using the Oracle Data Guard Broker are very promising. Thanks to the ease of deployment, performance and monitoring and its features, Oracle Active Data Guard will replace Oracle Data Guard deployments of all critical LHC databases.

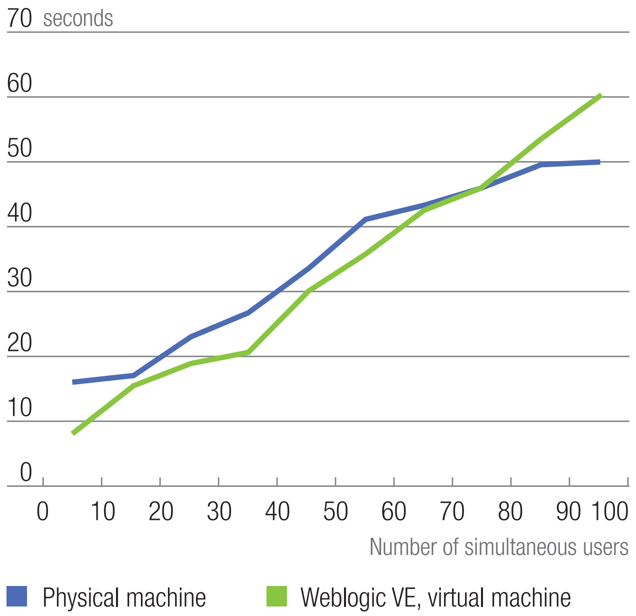
From left to right: Svetozar Kapusta, Carlos Garcia Fernandez, Zbigniew Baranowski, Eric Grancher and Eva Dafonte Perez in the CERN Computer Centre

Worldwide replications using Oracle Streams

Oracle Streams is the main replication technology used for LHC data distribution. From CERN, the data are distributed (using Oracle Streams) to ten sites around the globe (Tier-1 sites), enabling a highly complex replication environment where the ongoing maintenance presents a variety of challenges. One of the main problems is the Streams resynchronisation after a long downtime at one of the destination sites. After five days, the archived log files, where the database changes are logged in, are removed from the primary database. At this point, the defined synchronisation window is exceeded: archived log files recovery from backups is costly and the sustainable replication rate might be surpassed (Streams processes might not be able to recover the backlog generated during the intervention). The unique solution is to do a complete re-instantiation of the replica site. However, the data transfer (schemas and tables) using the Data Pump utility may take days depending on the amount of data to be transferred and the capacity of the destination site. Within the openlab framework, a new procedure to perform a complete Streams re-instantiation of a destination database (out of the replication flow) was developed using transportable tablespaces to copy the data from another replica and minimise the impact of the operations on the source database. When a complete re-instantiation of a replica site in a Hub/Spoke Streams environment is needed, the use of transportable tablespaces (in order to resynchronise the replica database) saves time and speeds up the process thanks to its flexibility.

Virtualisation and monitoring

Within the context of openlab, significant work on virtualisation has been completed and two versions of Oracle VM (2.1.5 and 2.2) have been packaged and made available for automatic installation and central management. The development of these two versions, following input from CERN and following initial openlab evaluations, has been crucial as it enables the use of Oracle VM in large-scale environments; the package can be installed and configured quickly and without human interaction.



Response time of a physical versus virtual machine against the number of simultaneous users using a CERN workload. It shows that the Oracle WebLogic Server on JRockit Virtual Edition solution outperforms or is very close to the physical machine setup. This level of performance was reached with an unreleased version of the Oracle WebLogic Server on JRockit Virtual Edition and may not reflect the performance of the finished product.

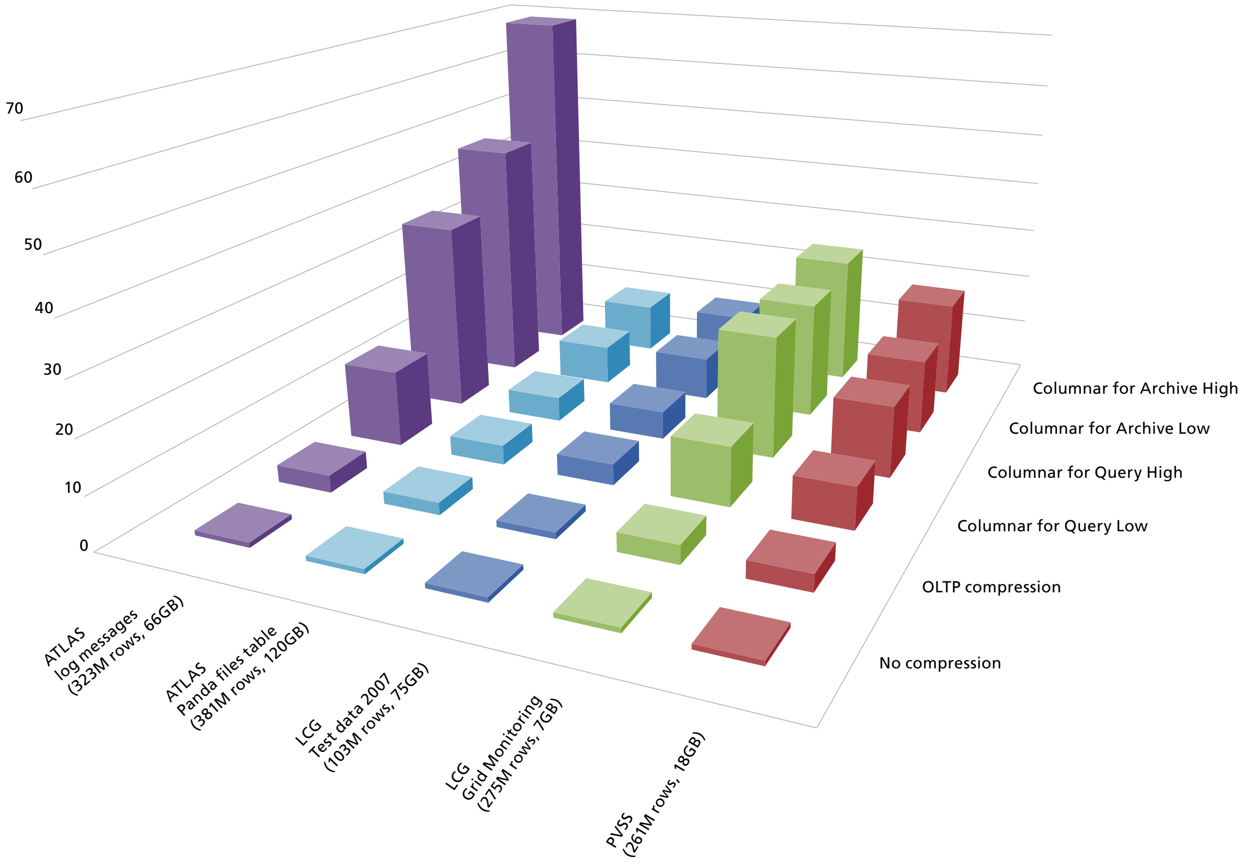
The team has also participated in a series of tests looking at the Oracle WebLogic Server on JRockit Virtual Edition, which is a special version of Oracle WebLogic Server. The deployment benefits for the organisation are impressive as this solution significantly simplifies the maintenance of the middleware solutions and provides cost-effective scalability on demand as it runs without a guest Operating System. It was possible to deploy successfully two Administrative Information Services (AIS) applications taken as models because of their complexity and the intensity of their workload. It gives the team confidence that any AIS application can be deployed. Stress tests were also performed comparing the physical machine and the virtual machine, with an even better performance obtained in the virtual machine in some situations. This level of performance was reached with an unreleased version of the Oracle WebLogic Server on JRockit Virtual Edition and may not reflect the performance of the finished product.

Thanks to the close collaboration between CERN openlab and Oracle's Enterprise Manager team, it was possible to do extensive beta-testing of the newest release -10.2.0.5 which led to a seamless upgrade of the production monitoring system. The team continued to centralise and standardise monitoring of the infrastructure by relying on the new features of Oracle Enterprise Manager instead of legacy monitoring solutions. Focusing on the following features was of great benefit to CERN: User Defined Policies and Metrics were investigated and used for implementing new monitoring rules, which are specially adapted to CERN requirements and complement the out-of-the box policies already used. Oracle Enterprise Manager Beacons is a new feature used to monitor service-based availability. Databases, hosts, listeners and application servers are grouped together to form Systems. Services are represented as a set of user-defined tests configured to run against System's components. These tests run from different locations thus evaluating service availability from the user perspective. The Beacon-based service monitoring is now in pre-production for some major CERN applications like Engineering Data Management System, Administrative Information Services, Software Version Control, etc.

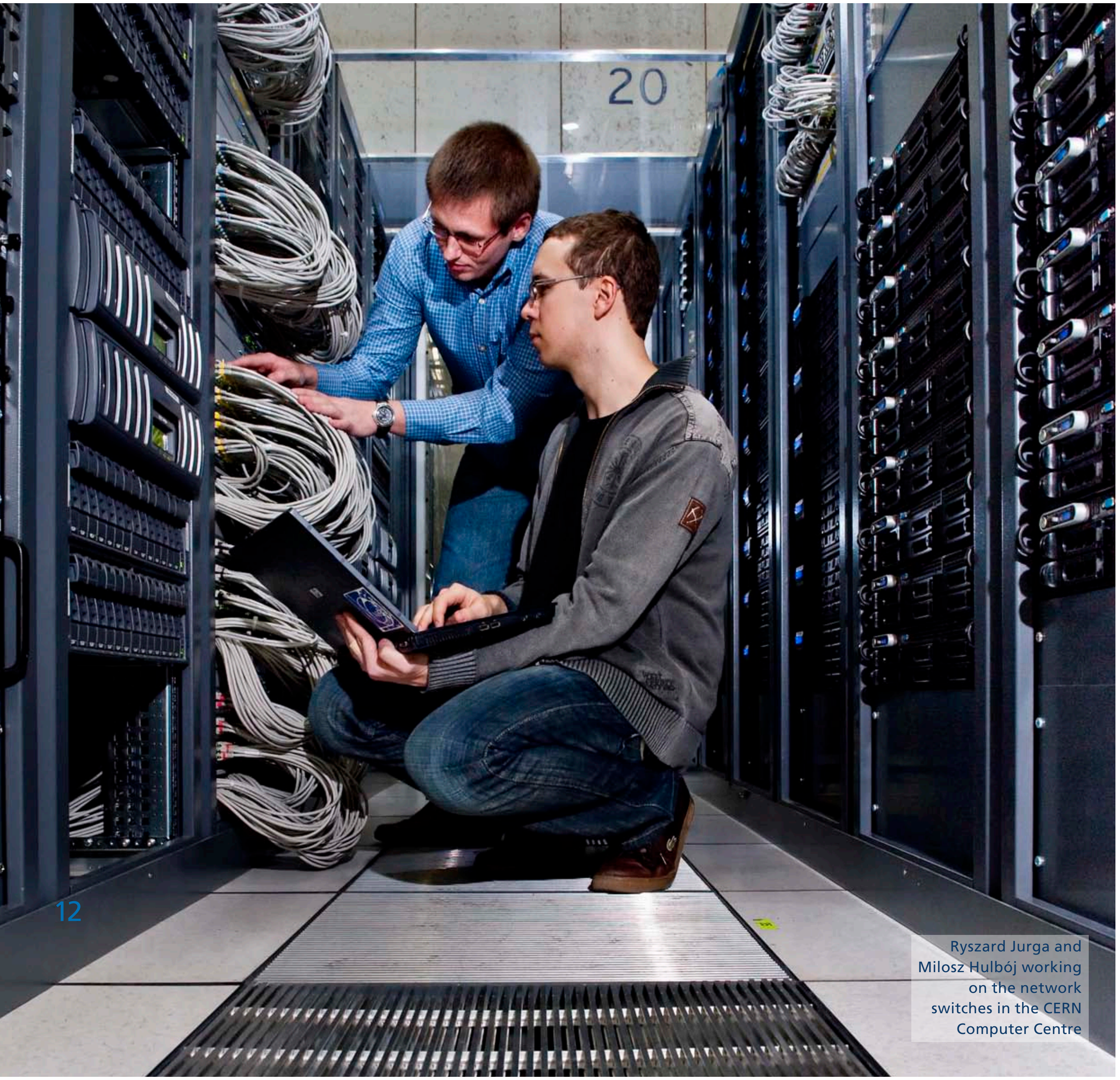
All of these results have been published and were presented at Oracle OpenWorld in San Francisco in October 2009 and the United Kingdom Oracle User Group conference in Birmingham in December 2009.

Test of the Advanced Compression Option with Oracle Exadata

The expected growth of the LHC experiments databases is roughly 20 terabytes per year per experiment. They need to have all data available at all times, not only during the experiment lifetime (10–15 years), but also for some time afterwards, as the data analysis will continue. To meet this need it is necessary to provide an efficient way of accessing and storing the data petabytes which is mostly read-only. The answer to this challenge could be the compression available in Oracle Database 11g Release 2 on the Database machine. Tests were performed to validate the hypothesis. The system used was located in Reading, UK, and accessed remotely from Geneva. It consisted of four nodes and seven storage cells with 12 disks each. The tests focused mainly on OLTP and Hybrid Columnar Compression (EHCC) of large tables for various representative production and test applications used by the physics community, like PVSS, GRID monitoring and test data, file transfer (PANDA) and logging application for the ATLAS experiment. Tests on export datapump compression were also performed. The test results are impressive as the following compression factors were achieved: 2–6X compression factors with OLTP and 10–70X compression factors with EHCC archive high. The EHCC can achieve up to 3X better compression than tar bzip2 compression of the same data exported uncompressed. Oracle Compression offers a win-win solution, especially for OLTP compression as it shrinks the used storage volume while improving performance.



Compression factor for physics database applications using 11gR2 compression. The X axis represents the selected applications and the Y axis represents the different compression types, the Z axis represents the compression factor.



Ryszard Jurga and Milosz Hulbój working on the network switches in the CERN Computer Centre

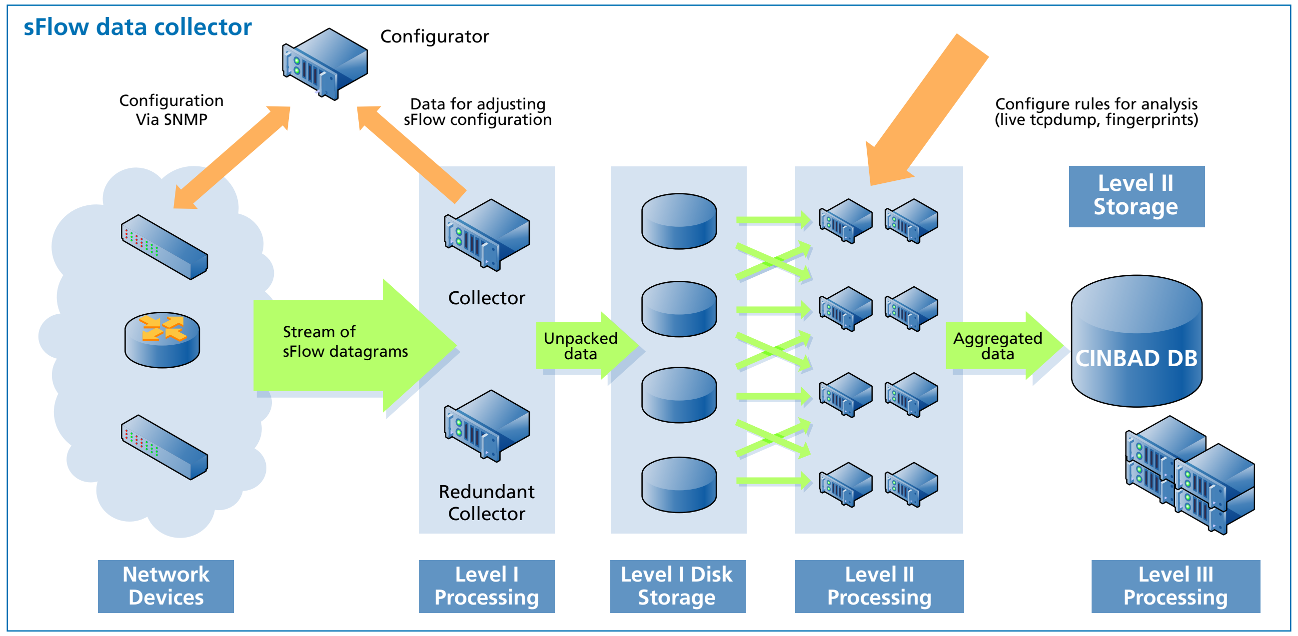
The Networking Competence Centre

The CINBAD (CERN Investigation of Network Behaviour and Anomaly Detection) project was launched in 2007 in collaboration with HP ProCurve. This challenging project's mission is to understand the behaviour of large computer networks in the context of high-performance computing and campus installations such as CERN. CERN's campus network has more than 50 000 active user devices interconnected by 10 000 km of cables and fibres, with more than 2500 switches and routers. The potential 4.8 terabit per second throughput within the network core and 140 gigabit per second connectivity to external networks offer countless possibilities to different network applications. The CINBAD goals are to detect traffic anomalies in such systems, perform trend analysis, automatically take counter-measures and provide post-mortem analysis facilities. The project is divided into three phases: data collection and network management, data analysis and algorithm development, performance and scalability analysis. This research activity is now producing practical results as well as providing crucial information to the CERN security team.

Project

The starting point of the project was to define the requirements and ensure a common framework of precise definitions, for example, what constitutes an anomaly or a trend. The following common denominator emerged: an anomaly is always a deviation of the system from the normal (expected) behaviour (baseline); the normal behaviour (baseline) is never stationary and anomalies are not always easy to define. As a consequence, anomalies are not easy to detect. However, some potential anomaly detectors can be identified. Thus the use of statistical detection methods can be considered. By learning the 'normal behaviour' from network measurements, and continuously updating the 'normal baseline', it is possible to detect new, unknown anomalies. Applying such a method also has some drawbacks as it is still possible to attempt to force a false negative, the selection of suitable input variables is an issue (many anomalies being within 'normal' bounds of the metrics), and finally a false positive can be extremely costly and does not provide a satisfactory anomaly type identification.

With the modern high-speed networks it is impossible to monitor all the packets traversing the links. sFlow is the industry standard for monitoring computer networks by means of random packet sampling. In fact sFlow is derived from the collaboration between HP, the University of Geneva and CERN in 1991. The team decided that the best way to monitor the network on a large scale was to use data statistical analysis by packet sampling. In the first phase of the project the CINBAD team investigated the feasibility of packet sampling in the context of anomaly detection. The results of this research were published at the end of 2007 in a report 'Packet Sampling for Network Monitoring', by Milosz Hulbój and Ryszard Jurga.

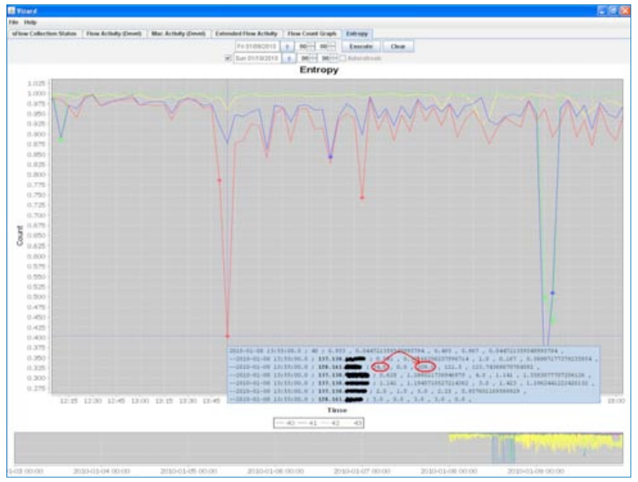


These studies, complemented by an in-depth analysis of the sFlow agents, enabled the CINBAD team to design and implement the sFlow data collector. Given the huge amount of sFlow data (300 000 samples per second) to be collected and analysed, the team decided to benefit from CERN's know-how in data storage and analysis. During the survey on data acquisition, the LHC experiments and Oracle experts were consulted on high-performance data storage, data format and representation and analysis principles. This survey led the team to the design of the multi-stage sFlow collector (see figure below) and to implement it. Since summer 2008, the collection system has been successfully running on a large-scale network, using approximately 1000 HP switches.

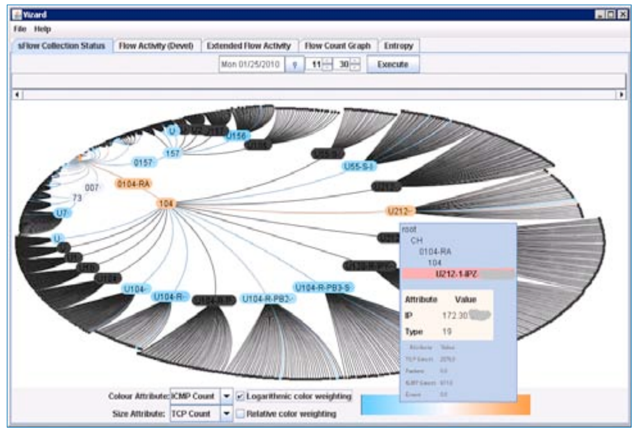
Results

The team developed tools for analysing the stored data based on these data collection tests results. Various data analysis techniques have been tested, among them a statistical data analysis and time series mining and signature based approach. In addition to these tools, the team adapted SNORT (open source network intrusion prevention and detection system) to work with sampled sFlow data. This SNORT setup was complemented by open source traffic rules as well as in-house CINBAD rules. Initial data analysis has enabled the team to detect some misbehaviour and a certain number of anomalies in the CERN network. It appears that most of these security anomalies (malicious software, policy violations) originated from end-user machines.

Over the past few years, the CINBAD team has collected and analysed more than 30 terabytes of data in order to investigate various approaches to enhancements for Network Monitoring and Anomaly Detection. Now, the CINBAD project is turning into the production stage and is delivering more and more new facilities. The current CINBAD toolkit (C-Eye) 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 network operations as well as reducing the time needed to diagnose network problems, since detailed information about the activity of each machine is available. The trend analysis tool provides metrics about evolution of the CERN network. Information about the number of active devices, network infrastructure usage, protocol distribution as well as other network statistics constitutes valuable input to the future design. In addition, the team is delivering a tool that enables one to browse the alerts from the CINBAD anomaly detection engine as well as identify their causes. All the aforementioned tools have already proved to be useful in resolving problems, including printer abusers, spammers and non-legitimate network scans. However, even though the project is entering the final stage, the team is continuously searching for, and identifying, new enhancement possibilities.



CINBAD tools allow for easy identification of the anomalous hosts via analysis of the network parameters' entropy.



CINBAD tools for network operation (C-Eye) are being currently integrated into the network management platform at CERN.

At the end of March, the CINBAD team visited the HP ProCurve group in Roseville, USA. The main goal of this visit was to present and discuss the latest results and transfer the knowledge that was gathered during the course of the project. CINBAD ideas, tools and algorithms were very well received by people from different groups of ProCurve. HP is considering integrating some of them into their products. This, together with the ongoing deployment and integration at CERN is clear evidence of the project's success.

Who knows which way the WIND is blowing?

In February 2010, a new research project which aims at improving the Wi-Fi network deployment and operation in large campus environments was begun with HP ProCurve.

Nowadays, wireless networks are becoming ubiquitous, and the density of wireless networks is increasing. Compared to wired networks, the wireless installations pose many more challenges due to the nature of the radio-wave propagation and limited spectrum available. Research on optimal deployment and operation of large-scale installations is still in its infancy.

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 different areas for access, e.g., conference rooms, auditoriums, long corridors, offices but also warehouse-like buildings and assembly halls. Answering the expectations of such a large and diverse user community is a real challenge.

A new openlab team under the codename WIND (Wireless Infrastructure Network Deployment) will carry out research and provide new algorithms, guidelines and solutions that will support the deployment and operation of the Wi-Fi infrastructure at CERN. Results of the research could then possibly be incorporated into HP ProCurve hardware and software to provide even more robust and efficient networking solutions

Sharp tools are needed to attack this difficult research problem. ProCurve Access Points provide the essential elements of information.

Julien Leduc performing power measurements in the CERN Computer Centre

The Platform Competence Centre

The Platform Competence Centre (PCC) was initially created in the context of CERN openlab-II. CERN openlab-III, in collaboration with Intel, continues to address crucial fields such as thermal optimisation, multi-core scalability, application tuning and benchmarking. It also has a strong emphasis on teaching.

Power and computing efficiency

Generally speaking, mastering energy consumption and thermal aspects in large computing centres is one of the major challenges that the information society now faces. This matter is particularly relevant to CERN, as the organisation has more than 6900 servers in its Computer Centre and another 6000 at the detector sites, processing the enormous amounts of data that are being produced by the LHC. Since the CERN Computer Centre facilities are severely limited both in terms of electrical input and of cooling output, the need for intelligent power optimisation is paramount. In order to ensure that the systems render maximum performance per watt, so that the data centre power consumption is maintained within the 2.9 megawatt limit, production and beta platform power measurements are conducted on a regular basis.

Furthermore, improving the performance of the Computing Centre facilities by even a small amount can be equivalent to saving millions of Swiss Francs in hardware purchases. This is why CERN openlab maintains a keen interest in monitoring the performance of both the most recent and upcoming hardware. Apart from standard performance optimisation activities conducted using various Intel and open source software packages, the team published a paper summarising the performance optimisation strategies currently in use at CERN. This work was initially based on interviews conducted with the representatives of the four major experiments at CERN and two commonly used software frameworks.

Benchmarking and optimisation

In the last year, through its collaboration with Intel, CERN openlab published three benchmark reports made publicly available on the openlab website. It was decided that a standard approach results in a good documentation of findings and effortless comparisons in the future. The first report, published in October 2009, focused on the evaluation of the energy consumption and the performance of Intel's "Nehalem" architecture, represented by the Intel® Xeon® processor 5500 series. The team evaluated three flavours of parts with varying power needs and performance levels: the low power L5520, the mid-range E5540 and the most powerful of the "Nehalem" series, the X5570. Their efficiency was evaluated by measuring their typical power consumption, using standard benchmarks to put stress on the different subsystems in the server. The team also assessed the performance of the processors with the C++ subset of the SPEC2006 benchmarks, dubbed "HEPSPEC06". The tests showed impressive results, with the more recent L5520 delivering a 36% energy efficiency improvement over the previous generation Xeon 5400 "Harpertown" servers and the other 5500 flavours reaching 30%. Improved efficiency was not the only positive point, since the "Nehalem" introduced ''Intel® Turbo Boost Technology'' and reintegrated Hyperthreading Technology, Intel's SMT (Simultaneous Multi Threading) implementation which allows each processor to execute simultaneously two threads per core by sharing the execution pipelines. SMT was thoroughly evaluated and appeared to be promising for the Computer Centre as it enables the throughput of processed jobs to be increased by 15 to 21%, based on the tests. This evaluation involved multiprocessing (using a Monte Carlo based benchmark, «test40») and a multithreaded benchmark («tbb») based on the ALICE High Level Trigger and the Intel Threading Building Blocks as well as a complete real-world framework (from ALICE) and compared the efficiency of different global scheduling policies.

A subsequent report related to the Intel® Xeon® processor 5600 series, codenamed "Westmere", was published in April 2010. The methodology used closely resembled the one established for the "Nehalem" report, but some legacy benchmarks were replaced with modern, real-world multi-threaded code: a parallel prototype of the Geant4 framework processing a simulation workload from the CMS experiment, and a multi-threaded minimisation application, built on the ROOT framework. It was determined that the die-shrunk "Westmere" capitalises on the rich enhancements of the "Nehalem" microarchitecture through an increased core count. With respect to the previous generation, the performance per watt has increased by up to 23%, the overall system performance was between 39% and 61% better, and the benefit of SMT is practically unchanged.

Finally, a third report, also published in April 2010, covered the Intel® Xeon® processor 7500 series, designed for multi-socket "Nehalem-EX" platforms. Using the exact same standard methodology as in the Xeon 5600 evaluation, CERN openlab learned that the "Nehalem-EX" platform provides excellent and close to linear scalability, with many tested applications. Compared to a Xeon 7400 based system, codename "Dunnington", the tested solution excels in many areas. The HEPSPEC06 benchmark yields 3.5x more throughput, and the throughput of other workloads has increased by between 47% and 87%, depending on the application. The Database Competence Centre has also tested the 7500 series and has found impressive scalability for the LHC accelerator Oracle database workload. It enables CERN both to deploy the largest database applications in an optimal way and to consolidate cost-efficiently many of the smaller database applications.

In addition to the regular platform benchmarking activities, research on Solid State Drives (SSD) has been restarted after related developments at Intel. SSD activities take place both in the Platform and the Database Competence Centres, and a preliminary SSD evaluation conducted on a Nehalem server has shown very promising results. A detailed report on the matter is expected in the upcoming year.

Multi-threading and many-core scalability

Another area where CERN openlab contributes is compiler optimisation, where the aim is to improve performance of a wide range of different jobs by influencing the back-end code generator. Tests with the Intel C++ compiler, version 11.1, were performed using both GEANT4 and ROOT benchmarks for both Intel64 and IA64. The project is directly related to the multi-core and many-core revolution, which permits a significant increase in computing power within a constant processor power envelope. The move to multi-core processors has already enabled CERN to benefit from on-going improvements in overall performance, without a corresponding increase in processor power consumption. Although the amount of memory has to be kept constant per core, the power savings compared to a non multi-core scenario have been impressive, which is highly beneficial to CERN. The openlab team has continued to work on establishing how the new multi-core architectures relate to High Energy Physics (HEP) software. As almost all LHC programmes (simulation, reconstruction, data analysis, etc.) are written in-house by high energy physicists, it is crucial to understand which modifications in the code can provide the most benefit from a multi-core or many-core architecture. Many experiments related to this domain were carried out in summer 2009, and have produced interesting results.

In recent months, parallelisation efforts, although not widespread at CERN, have started to bear fruit and are in a large part actively supported by openlab. One such activity is carried out by North-eastern University researchers: PhD student Xin Dong and Prof. Gene Cooperman. It relates to a complete multi-threaded conversion of a serial physics processing framework commonly used in HEP. One of the prototypes resulting from this work had been passed to openlab for testing. Initial examinations showed good scalability on various 8-core systems, prevalent in CERN's computing centre at that time. Other tests, executed on a 24-core Dunnington system provided by Intel, confirmed that such a complex application does not always scale automatically when moving to double-digit core counts The openlab Platform Competence Center team interfaced with the researchers from the USA, as well as with local experts, in order to find ways to make the software more scalable. Very positively, after many months of work by the team's American colleagues, the software now scales up to 32 cores with 93% efficiency – a remarkable and unprecedented result.

As openlab is looking into the many-core future, it had the pleasure of welcoming Tim Mattson, an expert on multithreaded programming from Intel, who visited the team to share his views and gave an IT seminar on 'OpenCL, design patterns and software for heterogeneous many-core platforms' at CERN. The move to a multithreaded paradigm is still in progress at CERN, although significant advances and optimisations have already been made, and there are noticeable improvements introduced by the developing team based on suggestions from openlab. Furthermore, in the light of the availability of 32-core Nehalem-EX systems, openlab is looking forward to expanding its work on scalability thanks to this new hardware, and to gaining even more insight into the scaling behaviour of High Energy Physics frameworks on modern and future many-core architectures.

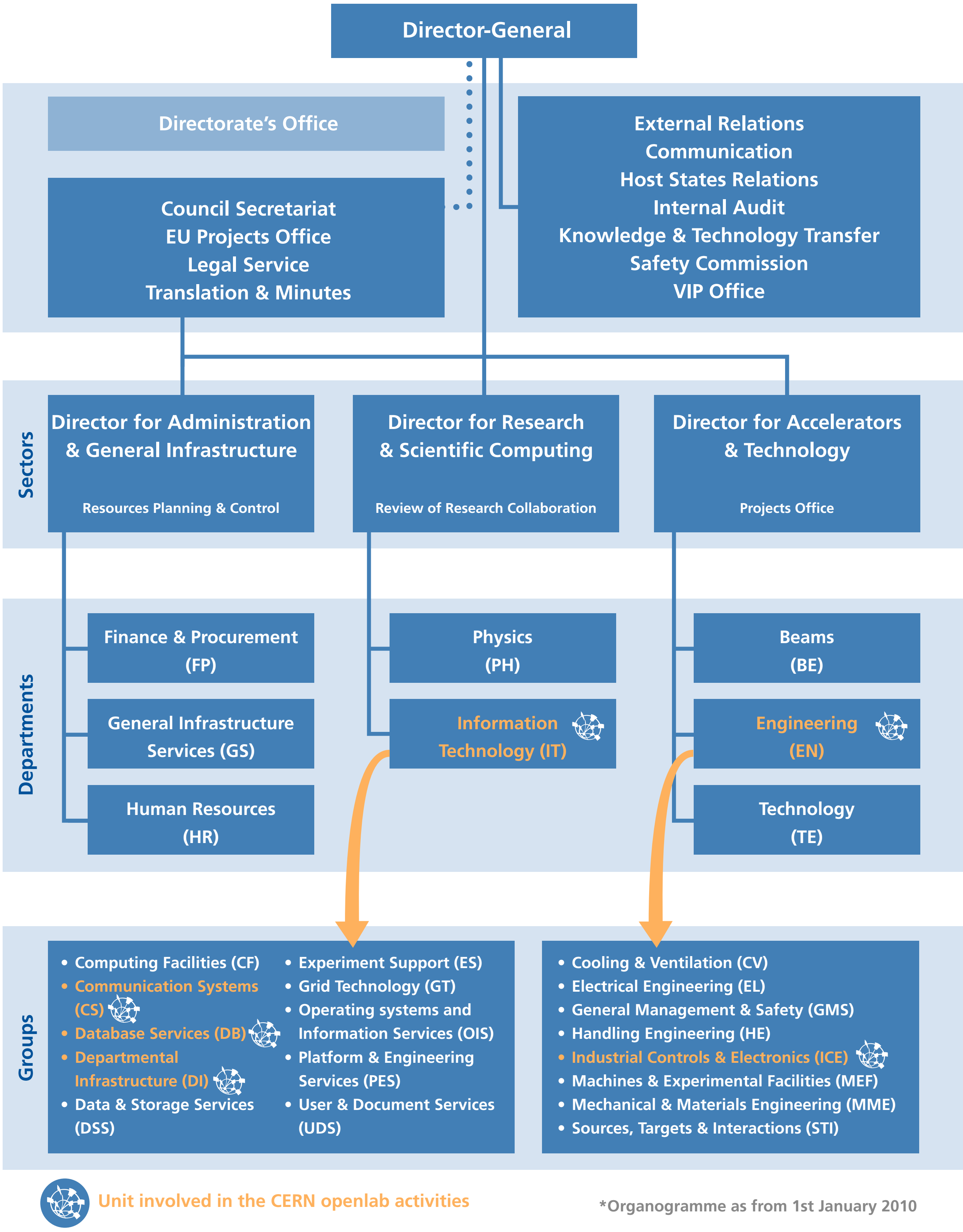
High-speed networking

Last year, several of Intel's 10 gigabit Network Interface Cards (NIC) were evaluated within the CERN openlab framework and in collaboration with other groups in the CERN IT department for their potential use in the IT production environment. The NICs surpassed all expectations and, as a result, there are currently 100 disk servers connected at 10 gigabit in the CERN Computer Centre. In addition, all tape servers, which are used for permanently storing LHC data, will be connected at 10 gigabit from now on. The InfiniBand based High Performance Computing cluster used for CFD calculations and other engineering applications (simulating cooling, temperature distribution for the experiments, etc.) is being replaced by a new cluster based on Intel Nehalem CPUs and Intel NetEffect low-latency 10 gigabit network cards.

The PCC team participated actively in numerous schools and workshops, listed in the dedicated sections of this report, to share and disseminate the knowledge created. For yet another year, CERN openlab has been teaming up with Intel to organise regular training for CERN's programmers. For these, Jeff Arnold, Senior Software Engineer from Intel assisted the regular CERN openlab lecturers. In addition to the regular quarterly courses on computer architecture, performance tuning and multi-threading, three new special courses have been organised for advanced CERN users, and were taught by top-level Intel experts. This activity was warmly welcomed and created good opportunities for Intel, the physics community and CERN openlab to share views and exchange feedback on new technologies and techniques.

The PCC team's continued participation in key conferences, such as Supercomputing and the Intel Developer Forum, ensures awareness of the latest technological developments and allows for valuable meetings with partners. The team also maintains active involvement in the CERN School of Computing, holding lectures and exercises yearly at the School and providing mentoring for lectures at the Inverted CERN School of Computing held in spring 2010.

Positioning CERN openlab activities at CERN



*Organogramme as from 1st January 2010

Education

CERN openlab is a structure designed to create knowledge. This is done through the evaluation of solutions as well as genuine research and development of IT technologies. This knowledge is then disseminated through multiple channels. One of the major ones is the publication of reports and articles, as described in the preceding section, but this may be viewed as passive dissemination. The openlab education programme, which provides active dissemination, is currently implemented through several lines of action.

Workshops or seminars are regularly organised at CERN on advanced topics directly connected to the openlab projects. More than 120 participants attended the tried-and-tested two-day workshop formula on "Multi-threading and Parallelism", "Multi-core and Virtualisation" and "Computer Architecture and Performance Tuning". These workshops have a special feature: they involve a mix of lecturers from both industry and CERN, thus exemplifying the openlab principle of two-way knowledge transfer through active collaboration. Several of the workshops combine hands-off theory with hands-on practice. In addition to the regular quarterly courses, special courses have also been organised this year for advanced CERN users. These classes touched on numerous topics, including future tools and optimisations and were taught by Intel specialists. However, openlab experts also contribute to off-site education activities such as the CERN School of Computing, where eight hours of lectures and exercises were delivered in 2009. These lectures and workshops are listed below.

- **Multi-threading and Parallelism Workshop**, 2-3 June 2009, CERN, J. Arnold/Intel, S. Jarp/CERN, A. Nowak/CERN
- **Multi-core and Virtualisation Workshop**, 25 June 2009, CERN, Experience with multi and many-core at openlab, A. Nowak/CERN
- **Multi-core and Virtualisation Workshop**, 25 June 2009, CERN, Virtualisation from a security perspective, J. C. Luna Duran/CERN
- **CERN School of Computing**, August 2009, Göttingen, Germany. Three lectures: Computer Architecture & Performance Tuning , S. Jarp/CERN, A. Nowak/CERN
- **CERN School of Computing**, August 2009, Göttingen, Germany. Two lectures: Networking Performance, F. Flückiger/CERN
- **1st CERN openlab/Intel Special Workshop 'Inside the Core',** 15-16 September 2009, CERN, C. Dahnken/Intel, A. Semin/Intel
- **CERN openlab/Intel Computer Architecture and Performance Tuning Workshop,** 6-7 October 2009, CERN, J. Arnold/Intel, S. Jarp/CERN, J. Leduc/CERN, A. Nowak/CERN
- **ESC09**, 12-16 October 2009, Bertinoro, Italy. Sverre Jarp/CERN and Andrzej Nowak/CERN delivered four lectures
- **Multi-threading and Parallelism Workshop,** 11-12 November 2009, CERN, J. Arnold/Intel, S. Jarp/CERN, J. Leduc/CERN, A. Nowak/CERN
- **2nd CERN openlab/Intel Special Workshop "Manycore Optimization",** 17-18 February 2010, CERN, L. Akyil/Intel, H. Cornelius/Intel, M. Deilmann/Intel
- **Multi-threading and Parallelism Workshop**, 4-5 May 2010, CERN, J. Arnold/Intel, S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN
- **3rd CERN openlab/Intel Special Workshop**, 6-7 May 2010, CERN, J. Arnold/Intel, K. Knobe/Intel, D. Levinthal/Intel

Mario Deilmann, from Intel Compiler Lab, delivering a lecture during the CERN openlab/Intel Special Workshop focusing on 'Manycore Optimization' in February 2010 at CERN

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. Their stay is financed by an innovative tripartite funding scheme (Industry, Universities and CERN). In 2009, the programme accepted 14 computer science and physics students of 10 nationalities for two months, during the period June to September. The students worked on cutting-edge computing technologies supervised by openlab staff, other groups in the IT Department as well as staff from the WLCG and EGEE Grid projects.

Visits were organised to the CERN Computer Centre, CERN Control Centre, ALICE, the Anti-matter Factory. In addition, the students spent a full day at the Ecole Polytechnique Fédérale de Lausanne (EPFL). A dedicated lecture series for the students was given by CERN and external experts. Several of this year's students were co-funded by Intel and Oracle.

CERN openlab Summer Student Programme Teaching Series, July-August 2009:

- Control System Cyber-Security, S. Lüders/CERN
- Hardware, A. Hirstius/CERN
- Linux OS, P. Kelemen/CERN
- Computer Architecture and Performance, S. Jarp/CERN
- Performance Monitoring, A. Nowak/CERN
- Oracle – How the Database Really Works, B. Kisbye Engsig/ORACLE
- Software Security, S. Lopienski/CERN
- Networking, R. Jurga/CERN
- Worldwide LHC Computing Grid, L. Poncet/CERN
- Grid Middleware, A. Unterkircher/CERN

The CERN openlab summer students 2009 are listed below, with home institute and project topic:

- Manuel CHAMBER GONZALEZ, Spain, Cranfield University, Bedford, UK, (On-line monitoring of job output)
- Pedro CUNHA, Portugal, EPFL, Lausanne, Switzerland, (The development of monitoring tools for the ATLAS Distributed Data Management System)
- Marc D'ARCY, UK, Brunel University, West London, (Performance and threading studies of High Energy Physics software)
- Rune JENSEN, Norway, NTNU-Norwegian University of Science & Technology, Trondheim, (Autovectorization and Compiler Scrutiny)
- Raúl JIMENEZ ESTUPINAN, Spain, Universidad de Las Palmas de Gran Canaria, (The development of a security tool for Oracle databases)
- Hifsa KAZMI, Pakistan, GIK Institute of Engineering Sciences & Technology, (Identification of system changes and recording in the group Change Management system)
- Miguel MARTINEZ PEDREIRA, Spain, University of Oviedo, (The development of harvesting scripts and metadata extraction plugins to collect multimedia from CERN collaborations)
- Raquel MENDEZ GOMEZ, Spain, Universidad Autonoma de Madrid, (The integration of Oracle virtualisation solution, Oracle VM, with existing fabric management system – ELFms)
- Massimo PALADIN, Italy, University of Udine, (The refactoring of the HammerCloud Interface)
- Julien PERROCHET, Germany/Switzerland, EPFL, Lausanne, (The integration of an open-source messaging system into the tools which will be used for monitoring the WLCG Grid Infrastructure)
- Vlad-Stefan PETRE, Romania, Politehnica University of Bucharest, (Investigation and implementation of methods for visualizing data collected by the CINBAD project in a hierarchical way)
- Karol POGONOWSKI, Poland, University of Edinburgh, (The developement of a security monitoring tool called Pakiti)
- Jarno RANTALA, Finland, Tampere University of Technology, (LHC physics simulation using CernVM and BOINC)
- Martin TVERDAL, Norway, NTNU-Norwegian University of Science & Technology, Trondheim, (The development of a plugin to manage Indico webcast requests and proof-of-concept of a chat system)

CERN openlab students at CERN, August 2009

Events and Outreach

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 addition, customer and press visits are organised by openlab partners. These groups are briefed about CERN openlab in a dedicated VIP meeting room known as the CERN openlab openspace. When visiting CERN openlab to discuss common projects with the team, some of our guests also give computing seminars in the IT auditorium.

Media Related Events

AJSPI (Association des Journalistes Scientifiques de la Presse d'Information) Visit, CERN, 18 June 2009
Introduction to CERN openlab as part of the Computer Centre visit.

Bill Thompson (Technology Critic) Visit, CERN, 2 Nov. 2009
Introduction to CERN openlab as part of the Computer Centre visit.

Techradar Visit, CERN, 3 November 2009
Introduction to CERN openlab as part of an interview and the Computer Centre visit.

Intel Press Event, CERN's Globe for Science and Innovation, 4 November 2009
29 press representing 17 countries. Presentation by Sverre Jarp of the results published in the 'Evaluation of Energy Consumption and Performance of Intel's Nehalem Architecture', authored by Axel Busch and Julien Leduc. Tour of the CERN Control Centre and the SM18 magnet testing facility.

Seminars by CERN openlab Guests

'Massive Data Computing in a Connected World', CERN, 5 October 2009
Given by Dr. Pradeep K. Dubey, Director of Throughput Computing, Intel Corporation, as part of his visit.

'The Challenges of Exascale Computing', CERN, 4 December 2009
Given by Dr. Rajeeb Hazra, Director of Supercomputing Architecture and Planning, Intel Corporation, as part of his visit.

'OpenCL, Design Patterns and Software for Heterogeneous Many Core Platforms', CERN, 15 March 2010
Given by Dr. Tim Mattson, Director of Supercomputing Architecture and Planning, Intel Corporation, as part of his visit.

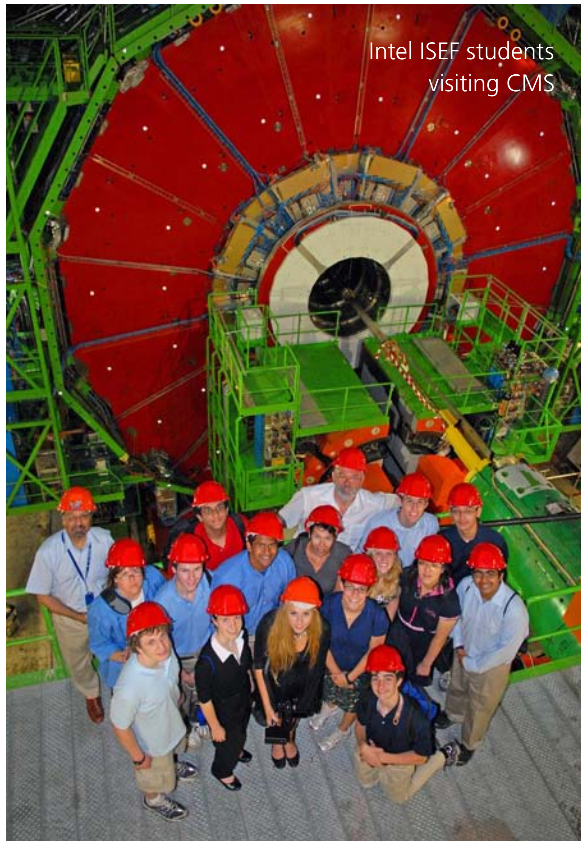
'Cells-as-a-Service', CERN, 24 March 2010
Given by Dr. Patrick Goldsack, Distinguished Technologist, Hewlett-Packard Laboratories, as part of his visit.

'Concurrent Collections (CnC): A new approach to parallel programming', CERN, 7 May 2010
Given by Kathleen Knobe, Scalable Tools Engineer, Intel Corporation, as part of her visit.

'Software Aspects of IEEE Floating-Point Computations for Numerical Applications in High Energy Physics', CERN, 11 May 2010
Given by Jeff Arnold, Senior Software Engineer, Intel Corporation, as part of his visit.

Intel ISEF Students Visit CERN, 28 June to 2 July 2009

During summer 2009, CERN hosted the visit of 13 pre-college students who won 'Best of Category' awards at the Intel International Science and Engineering Fair (ISEF) 2009 in Reno, USA. The young students spent four days at CERN, visited the LHC facility and enjoyed presentations from various prominent scientists. Journalists also came to interview the students and participated in a roundtable on Education.



Intel ISEF students visiting CMS

Publications and Presentations

CERN openlab results have been disseminated in a wide range of international conferences, listed below. For a full record of the presentations, consult the CERN openlab website. In addition, key results from CERN openlab have been the subject of a large number of press articles both in the general and IT-specific press and on the Web. These articles are listed on the CERN openlab website.

Presentations:

- I. Demeure/ENST and X. Gréhant/ENST&CERN, Symmetric Mapping: an Architectural Pattern for Resource Supply in Grids and Clouds, SMTPS, IPDPS, Rome, Italy, May 2009
- M. Hulboj/CERN and R. E. Jurga/CERN, CERN Investigation of Network Behaviour and Anomaly Detection, HEPiX spring session, Umeå University, Sweden, 26 May 2009
- M. Hulboj/CERN and R. E. Jurga/CERN, CERN Investigation of Network Behaviour and Anomaly Detection, Post-C5 Meeting, CERN, 12 June 2009
- M. Girone/CERN, Databases for the CERN LHC-Techniques and Challenges, Oracle OpenWorld 2009, San Francisco, USA, 13 October 2009
- S. Kapusta/CERN, Oracle Advanced Compression Tests, Oracle OpenWorld 2009, San Francisco, USA, 15 October 2009
- M. Guijarro/CERN, Chris Lambert/CERN, Lucia Moreno/CERN, Increasing productivity at CERN with Enterprise Manager Grid Control, Oracle OpenWorld 2009, San Francisco, USA, 11-15 October 2009
- M. Girone/CERN, Analyst Panel Talk, Oracle OpenWorld 2009, San Francisco, USA, 11-15 October 2009
- E. Grancher/CERN, Identifying SQL issues using AWR, cardinality feedback and analytic functions, UK Oracle User Group conference - Tech and E-Business Suite Conference, Birmingham, United Kingdom, 30 November-2 December 2009
- E. Dafonte/CERN, Worldwide distribution of experimental physics data using Oracle Streams, UK Oracle User Group conference – TechandE-BusinessSuiteConference,Birmingham,UnitedKingdom, 30 November-2 December 2009
- A. Lazzaro/CERN, Multicore Computing, ACAT 2010, Jaipur, India, 22-27 February 2010
- A. Lazzaro/CERN, The RooStats Project, ACAT 2010, Jaipur, India, 22-27 February 2010
- S. Jarp/CERN, A.Lazzaro/CERN, Multicore Panel presentations, ACAT 2010, Jaipur, India, 22-27 February 2010
- S. Jarp/CERN, Are we all moving to the Cloud, OPI event, CERN Globe, Geneva, 12 March 2010
- S. Jarp/CERN, More and Moore. CERN openlab tests Intel's latest servers, Intel Press Event, Zurich, 25 March 2010
- S. Jarp/CERN, More and Moore. CERN openlab tests Intel's latest servers, Intel Customer Event, Zurich, 7 April 2010

Publications:

- I. Demeure/ENST and X. Gréhant/ENST&CERN, Symmetric Mapping: an Architectural Pattern for Resource Supply in Grids and Clouds, SMTPS, IPDPS, May 2009
- M. Hulboj/CERN and R. E. Jurga/CERN, CINBAD keeps an eye on the CERN network, CNL, September 2009
- M. Hulboj/CERN and R. E. Jurga/CERN, CERN Investigation of Network Behaviour and Anomaly Detection, recent Advances in Intrusion detection, 12th International Symposium, RAID 2009 Saint Malo, France, September 2009
- F. Tilaro/CERN, The Siemens openlab Team lays Emphasis on Cyber Security Analysis for Industrial Control Systems, CERN openlab Newsletter, Issue 1, January 2010

Posters:

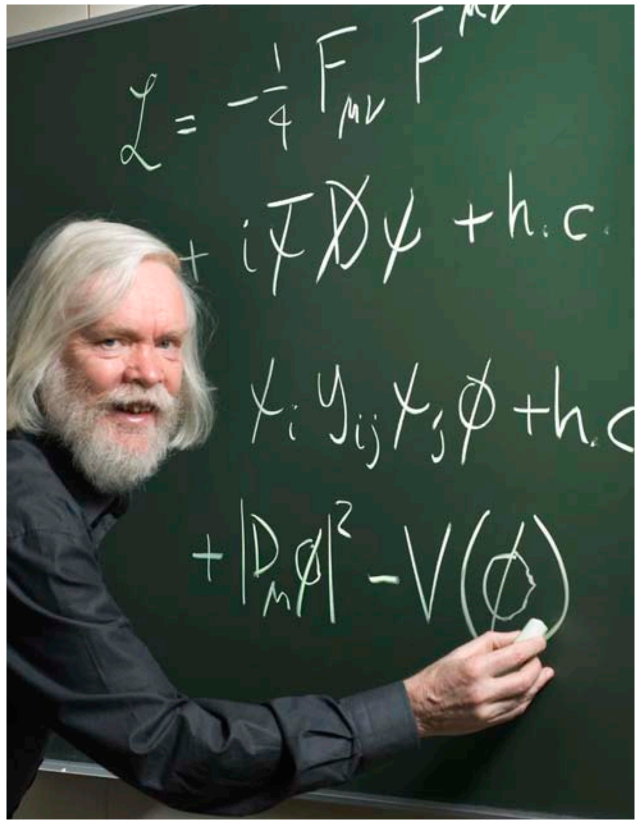
- M. Hulboj/CERN and R. E. Jurga/CERN, CERN Investigation of Network Behaviour and Anomaly Detection,12th International Symposium, RAID 2009 Saint Malo, France, 23 September 2009
- B. Copy/CERN, F. Tilaro/CERN, Standards based measurable security for embedded devices, 12th International Conference on Accelerator and Large Experimental Physics Control Systems (ICALEPCS), Kobe, Japan, 12-16 October 2009
- A. Lazzaro/CERN, Parallelization of likelihood function data analysis software based on RooFit package, ACAT 2010, Jaipur, India, 22-27 February 2010

CERN openlab Reports:

- M. Paladin/Summer Student, A Dynamic Web user Interface for HammerCloud, August 2009
- M. D'Arcy/Summer Student, Performance and Threading Studies of High Energy Physics Software, August 2009
- J. Rantala/Summer Student, LHC Physics Simulation Using CERN-VM and BOINC, August 2009
- H. Kazmi/Summer Student, Grid Control and Change Management, August 2009
- R. Jimenez Estupiñan/Summer Student, Development of a Security Tool for Oracle Databases, August 2009
- P. A. Cunha/Summer Student, Monitoring System of Data Subscriptions for ATLAS on the Grid, September 2009
- J. Perrochet/Summer Student, Nagios Messaging with ActiveMQ, September 2009
- R. Erlend Jensen/Summer Student, Compiler Comparisons using Performance Counters, September 2009
- R. Gaspar Aparicio/CERN, 11gR2 New Features test, September 2009
- A. Busch/CERN, J. Leduc/CERN, Evaluation of Energy Consumption and Performance of Intel's Nehalem Architecture, October 2009
- C. Garcia Fernandez/ CERN, Oracle WebLogic Server - Virtual Edition Testing at CERN openlab, December 2009
- A. Nowak/CERN, Strategies Employed for LHC Software Performance Studies, March 2010
- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, Evaluation of the Intel Westmere-EP server processor, April 2010
- S. Jarp/CERN, A. Lazzaro/CERN, J. Leduc/CERN, A. Nowak/CERN, Evaluation of the Intel Nehalem-EX server processor, April 2010

THE FUTURE

On the way to a long run



"Humanity is about to look deeper inside matter than has ever been possible before. We have many theories of what we might find, but only experiments can tell us which, if any, are right. Why do particles weigh? What is the dark matter that fills the Universe? What was the origin of the matter in the Universe? The answers provided by the discoveries of the LHC will revolutionise our understanding how the Universe works, and how it has evolved."

John Ellis, CERN Theory Group
March 2010

CERN will run the LHC for 18–24 months with the objective of delivering sufficient data to the experiments to make significant advances across a wide range of physics channels. As soon as the LHC experiments have 're-discovered' the known Standard Model particles, a necessary precursor to looking for new physics, they will start the systematic search for the Higgs boson. With the amount of data expected, called one inverse femtobarn by physicists, the combined analysis of ATLAS and CMS will be able to explore a wide mass range, with a possible chance of discovery if the Higgs has a mass near 160 GeV.

For supersymmetry, ATLAS and CMS will each have enough data to double today's sensitivity to certain new discoveries. Experiments are sensitive to some supersymmetric particles with masses up to 400 GeV. One inverse femtobarn at the LHC pushes the discovery range up to 800 GeV.

Even at the more exotic end of the LHC's potential discovery spectrum, this LHC run will extend the current reach by a factor of two. LHC experiments will be sensitive to new massive particles indicating the presence of extra dimensions up to masses of 2 TeV, where today's reach is around 1 TeV.

Following this run, the LHC will shut down for routine maintenance, and to complete the consolidation work needed to reach the LHC's design energy of 14 TeV. Traditionally, CERN has operated its accelerators on an annual cycle, running for seven to eight months with a four to five month shutdown each year. Being a cryogenic machine operating at very low temperature, the LHC takes about a month to be brought up to room temperature and also one month to cool down again. A four month shutdown, as part of an annual cycle, no longer makes sense for such a machine, so CERN has decided to move to a longer cycle with longer periods of operation followed by longer shutdown periods when needed.

The longer shutdown periods are also the right time to perform more significant changes at the experiments, both on the detectors and in the computing systems. Therefore, the planned shutdown period matches well the timing of openlab, where we are preparing new solutions now to be readily available for a possible implementation at the end of the present running period of the LHC.

About the partners

CERN openlab partners

HP, the world's largest technology company, simplifies the technology experience for consumers and businesses with a portfolio that spans printing, personal computing, software, services and IT infrastructure. More information about HP is available at www.hp.com.

Intel, the world leader in silicon innovation, develops technologies, products and initiatives to continually advance how people work and live. Additional information about Intel is available at www.intel.com/pressroom.

For 30 years, **Oracle** has been helping customers manage critical information. The company's goal is to make sure that customers spend less money on their systems while getting the most up-to-date and accurate information from them. Oracle does this by simplifying or outsourcing IT infrastructure to reduce costs, and by integrating disparate systems to create a single, global view of the customer's business. www.oracle.com.

Siemens AG (Berlin and Munich) is a global powerhouse in electronics and electrical engineering, operating in the industry, energy and healthcare sectors. The company has around 430,000 employees (in continuing operations) working to develop and manufacture products, design and install complex systems and projects, and tailor a wide range of solutions for individual requirements. For over 160 years, Siemens has stood for technological excellence, innovation, quality, reliability and internationality. www.siemens.com.

For more information

Executive Contact:
Wolfgang von Rüden, Head of CERN openlab,
Wolfgang.von.Rueden@cern.ch

Technical Contact:
François Flückiger, Manager of CERN openlab,
Francois.Fluckiger@cern.ch

www.cern.ch/openlab

CERN, the European Organization for Nuclear Research, is the world's leading laboratory for particle physics. It has its headquarters in Geneva. At present, its Member States are Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Italy, Netherlands, Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland and the United Kingdom. India, Israel, Japan, the Russian Federation, the United States of America, Turkey, the European Commission and UNESCO have Observer status.

www.cern.ch · www.youtube.com/cern · www.twitter.com/cern

CERN openlab

www.cern.ch/openlab

PARTNERS

hp invent

intel

ORACLE

SIEMENS