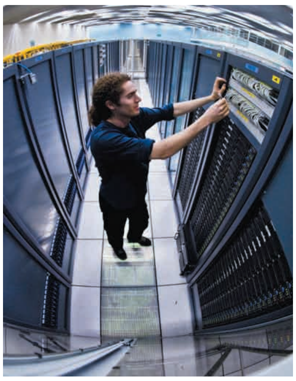
CERN openlab

Annual Report
May 2007 – May 2008

Table of Contents



Cover illustration: CERN openlab Ph.D. student Xavier Gréhant in the CERN Computer Centre

Second annual report of CERN openlab - Second Phase
Editor: François Grey
Photography: CERN photo service
Production: André-Pierre OLIVIER
ISBN 978-92-9083-312-3

Creating synergy

The Large Hadron Collider (LHC) brings together some 9000 researchers from around the world. Approved by the CERN Council in 1996, it will begin operation in 2008 and has an expected operational lifetime of around 15 years. The expe-rimental collaborations running the LHC's detectors, ALICE, ATLAS, CMS, LHCb and TOTEM are poised to bring new and profound insights into the workings of our Universe. Three main ingredients are necessary for this research at the frontier of knowledge – the LHC machine to accelerate and collide particles of matter, the detectors to observe the collisions, and a vast global computing infrastructure, the Worldwide LHC Computing Grid, to analyze the resulting data.

The LHC is now fully installed in its 27 km tunnel, and commissioning is well under way. As a result, the LHC is on course for start-up in summer 2008, although we will not be able to fix the date for certain before the whole machine is cold and magnet electrical tests are positive. We expect to see circulating beams by the summer, provided no major difficulties are encountered during the final stages of the commissioning.

Installation of the detectors is approaching its conclusion, too, and attention is turning towards physics analysis, including testing of the full data chain from the detectors through the Grid to data storage. Already, some experiments are taking data routinely with cosmic rays, and baseline Grid services are in regular daily operation.

As outlined in this report, the CERN openlab partnership continues to have a direct and positive impact on the development of the Grid and computing services that underlie the LHC. This happens through close technical collaboration between CERN openlab and the LHC Computing Grid project, as well as with the related European multi-science Grid project EGEE.
The synergy that CERN openlab creates with leading IT companies is vital, as the long-term future of Grid computing will surely depend on a broad commercial adoption of the technology, just as the hardware currently in CERN's Computer Centre now relies heavily on advances made for the commodity market. I therefore thank all CERN openlab partners and contributors for their continued support of our joint effort.

I anticipate many future shared benefits of this unique collaboration, and look forward to welcoming the openlab partners and contributors to the inauguration ceremony that will take place for the launch of the LHC in October.

Robert Aymar
Director General of CERN

THE CONTEXT

The LHC – Ready to go

At a brief ceremony deep under the French countryside in November 2007, CERN Director General Robert Aymar sealed the last interconnect between the main magnet systems in the Large Hadron Collider (LHC). This was a major milestone in commissioning the LHC, the world's most powerful particle accelerator.

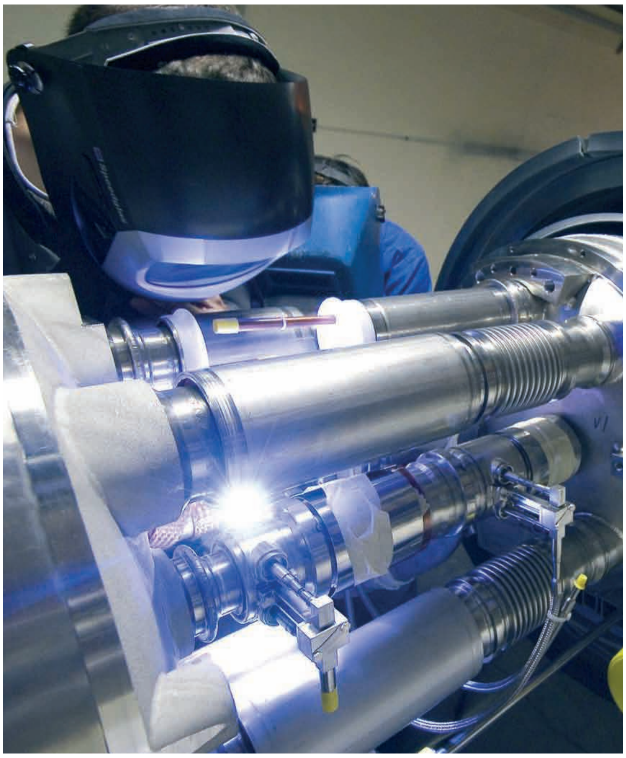
The LHC's superconducting main magnets will operate at a temperature of just 1.9 degrees above absolute zero (–271.3°C), colder than outer space. To cool the magnets, over 10,000 tonnes of liquid nitrogen and 130 tonnes of liquid helium will be deployed through a cryogenic system including more than 40,000 leak-tight welds. A two year programme of work was required to connect all the main dipole and quadrupole magnets in the LHC, a complex task that includes both electrical and fluid connections. During 2007 and early 2008, the final components of the LHC experiments at CERN are slotting into place, the last pieces of a giant jigsaw puzzle. At ALICE, ATLAS, CMS and LHCb the remaining large pieces of equipment have been carefully lowered into the caverns in preparation for the start up in 2008 of the LHC.

The four big experiments are crammed full of some of the most complex scientific apparatus in the world. Manœuvring the pieces, sometimes weighing hundreds of tonnes each and measuring several metres in diameter, is a formidable challenge. The huge parts have to be manipulated in very confined spaces, often with only a few centimetres of clearance on either side of the cavern shaft. Consider, for example, the installation of the world's largest silicon tracking detector which was successfully completed in December 2007 for the CMS experiment. With a total surface area of 205 square metres, about the same as a singles tennis court, the CMS Silicon Strip Tracking Detector is by far the largest semiconductor silicon detector ever constructed. Its silicon sensors are patterned to provide a total of 10 million individual sensing strips, each of which is read out by one of 80,000 custom designed microelectronics chips. Data are then transported via 40,000 optical fibres into the CMS data acquisition system.

Or consider the so-called small wheels of the ATLAS muon spectrometer, the second of which was lowered 100 metres into its underground experimental cavern in February 2008. Each of these wheels is in fact huge: 9.3 metres in diameter and weighing 100 tonnes, including massive shielding elements. The wheels are covered with sensitive detectors to identify and measure the momentum of particles that will be created in the LHC collisions. The entire muon spectrometer system contains a detector area equal to three football fields, including 1.2 million independent electronic channels.

In all respects, the engineering challenges involved in commissioning the LHC have been enormous. And in 2008, twelve years after the project was first approved by CERN Council, the LHC is ready to go.

Small wheels of the ATLAS muon spectrometer



Connecting the superconducting dipole magnets of the LHC

The Grid – Primed for data

The LHC Computing Grid project (LCG) supports the offline computing needs of the LHC experiments in a globally distributed manner. In 2007, the number of sites contributing to LCG continued to grow, reaching over 200 computing centres in 35 countries, organized in a Tier-1 of 11 major data centres, and a Tier-2 consisting of 38 federations of smaller centres.

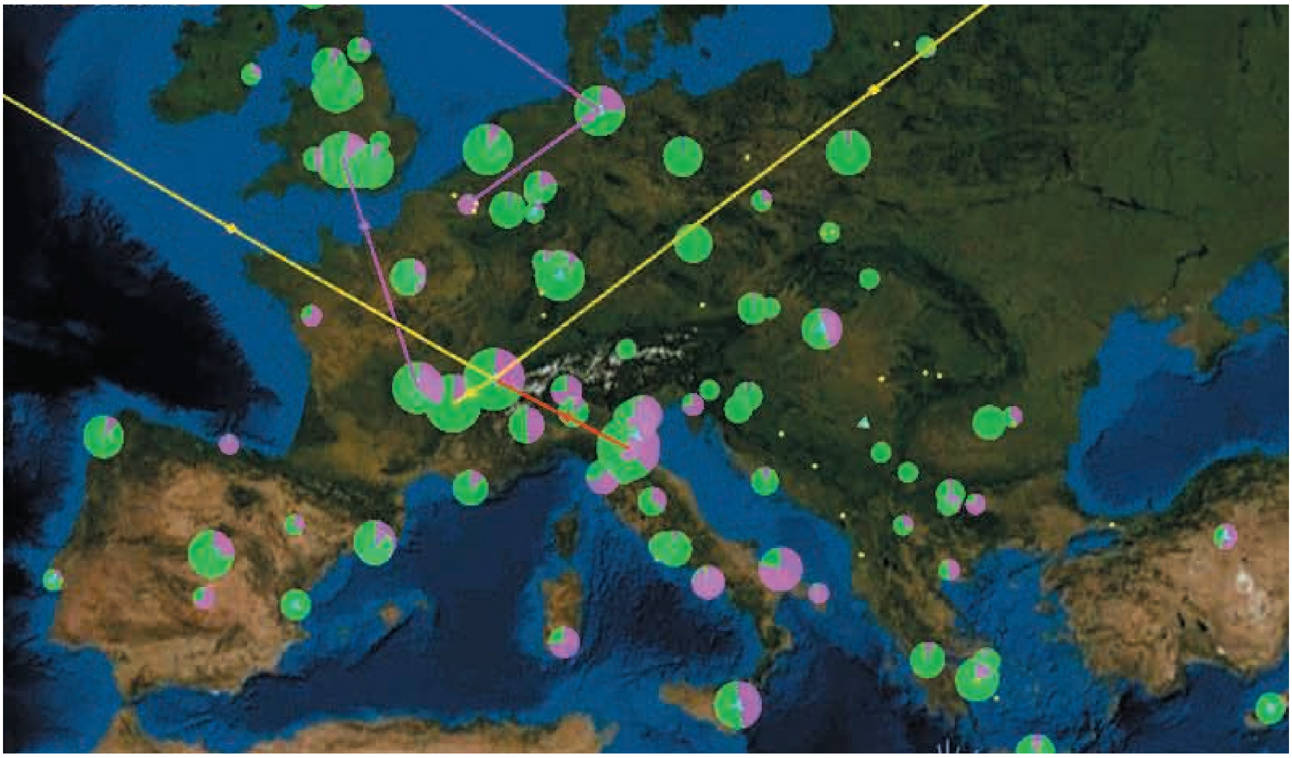
In 2007, the 10Gbit/s connectivity between CERN and each of the Tier-1 sites was completed. This will support data rates from CERN to the Tier-1s in excess of 1GB/s, which is what is expected once the LHC reaches design performance. Over a year, and including raw data, analysis results and simulations, CERN will store some 15 Petabytes of data for the LHC, with the Tier-1s committing to jointly store a backup of the raw data. The Tier-2 sites have continued to grow in power and reliability, and now supply over 50% of the total computing capacity needed for LHC data analysis, which is estimated at about 100,000 processors.

One example amongst many of how CERN openlab activities have been contributing in a tangible way to the LCG agenda is the scaling up of the distribution of databases from CERN to the Tier-1 centres. This is necessary so that the experiments will have up-to-date detector calibration constants when reconstructing data on the Grid. During the past eighteen months, major improvements have been made in the rates at which databases could be distributed, through the CERN openlab partnership with Oracle.

The Enabling Grids for E-sciencE project (EGEE), which manages the world's largest multi-science Grid infrastructure, is also the largest Grid infrastructure contributing to LHC computing. CERN's IT Department leads this project, which has 91 institutional partners. EGEE has started a third two-year phase of its existence in May 2008, with co-funding from the European Commission. This support is essential for CERN and the LHC. For example, EGEE supports a third of the 60-person grid deployment group at CERN.

In 2007, the throughput of the EGEE Grid reached more than 100,000 jobs per day, about 70% of which was related to the LHC. More than 5000 people have registered to use the EGEE infrastructure, and an estimated 12,000 scientists benefit from its existence. EGEE continues to promote Grid technology in other scientific disciplines and in industry, where a close collaboration with CERN openlab has led to several joint activities. One example in 2007 was the joint EGEE-openlab event on Computing for Finance, held at CERN, and featuring speakers from leading banks and firms providing Grid-related financial services.

Another important contribution of EGEE is to improve interoperation with other major Grid projects, notably the US-based Open Science Grid, which are also supporting LHC computing. This sort of interoperation is an important practical step towards open Grid standards, which many view as essential in order to increase industrial commitment for Grids.



Real time monitor showing some of the European sites the LHC Computing Grid in operation. Green indicates jobs being processed, mauve queued. Lines indicate jobs being transferred for processing. Monitor courtesy of Imperial College, GridPP and EGEE.

THE STATUS

Catalyzing collaboration

CERN openlab, a framework for partnership with industry launched in 2003, is now five years old. During these five years, CERN openlab has come to play an increasingly pivotal role in the strategy of CERN's IT Department, enabling test and validation of new commercial technologies.

Such testing in CERN's demanding IT environment provides the partners with valuable feedback on their products. This also allows CERN to assess the merits of new technologies for possible future use, in particular for the major international Grid projects that CERN leads. These are the LHC Computing Grid (LCG) project, which provides the core Grid services for the LHC experiments, and the Enabling Grids for E-sciencE (EGEE) project, co-funded by the EU, which runs a Grid infrastructure for a wide range of scientific and industrial applications.

Beyond the technical results achieved, which are amply described in this report, CERN openlab also provides a place for the different industrial partners to get to know each other and work together. This benefit has become particularly visible during the course of the last year.

As a case in point, let me mention the flourishing collaboration between Oracle and Intel, which was not anticipated at the outset when these two companies joined CERN openlab. This collaboration evolved out of discussions that have occurred during various CERN openlab meetings, and has led to initiatives such as an Oracle Intel Architect meeting, held at CERN in February 2008. The collaboration between the two firms, in the context of CERN openlab, was even highlighted by the CEO of Intel, Paul Otellini, during a keynote speech he made at Oracle OpenWorld in November 2007. CERN and Oracle also celebrated 25 years of collaboration marked by a joint event on innovation.

Another positive development during the past year has been the collaboration initiated with ProCurve Networking by HP. ProCurve was already a significant supplier of networking equipment to CERN, but thanks to the existing HP participation in CERN openlab, the relationship has expanded now to cover long-term, exploratory research in the area of network monitoring and security. This research is fully integrated in the CERN openlab framework. ProCurve is sponsoring two CERN fellows in this area, a significant commitment within the context of HP's overall contribution to CERN openlab.

EDS joined as a contributor last year and has developed a new Grid monitoring tool, GridMap, which found a very good resonance in the Grid user community.

In conclusion, I would like to thank all the openlab partners and contributors for their support of this ambitious partnership. As the eyes of the world turn to CERN this year, in anticipation of the launch of the LHC, I look forward to a period of exciting technical developments, as our joint efforts continue to support the major Grid projects that will analyze the LHC data, and help physicists open new frontiers in our understanding of the Universe.

Wolfgang von Rüden
Head of CERN openlab
Head of CERN's IT Department

The openlab team: a broad partnership

The work of CERN openlab is carried out by people from several groups in CERN's IT Department, in particular CS (Communication Systems), DES (Databases and Engineering Systems), IS (Internet Services) and PSS (Physics Services Support) now DM (Data Management), as well as the Department's security section and communications team. Many of these are sponsored by the openlab partners, the EU or national programmes. There is, in addition, close collaboration with computing experts in the LHC experiments. A list of the IT Department staff most closely involved in the CERN openlab activities is given below. It should be emphasized that the openlab partners contribute significant amounts of manpower to these activities through the time and effort invested by many of their engineers and scientists from around the world. Principal liaisons with partners and contributors, both at the technical and partnership management levels, are also listed below. In addition, substantial contributions are made by students participating in the CERN openlab student programme, both directly to openlab activities (7 students during summer 2007) and more widely to LCG, EGEE and other Grid-related activities in the IT Department (25 students).

CERN openlab Board of Sponsors

Robert Aymar	CERN (head of Board)
Wolfgang von Rüden	Head of CERN openlab
Michel Bénard	HP
Richard Dracott	Intel
Sergio Giacoletto	Oracle

CERN openlab Management Unit

François Fluckiger	CERN openlab Manager
Sverre Jarp	Chief Technology Officer
Séverine Pizzera	Administrative Assistant

CERN openlab Staff and Fellows (sponsor indicated)

Gyorgy Balazs	Technical Student (CERN)
Håvard Bjerke	Fellow (Intel)
Daniel Filipe Rodrigues	Fellow (EDS)
Xavier Gréhant	Fellow (HP)
Milosz Marian Hulboj	Fellow (HP ProCurve)
Ryszard Jurga	Staff (HP ProCurve)
Andreas Hirtsius	Staff (Intel)
Andrzej Nowak	Fellow (EU Marie Curie)
Eva Dafonte Perez	Staff (Oracle)
José M. Dana Pérez	Fellow (HP)
Anton Topurov	Fellow (Oracle)
Dawid Wójcik	Fellow (Oracle)

CERN liaison

Ian Bird, Liaison with EDS
Sverre Jarp, Liaison with HP, Intel
Jean-Michel Jouanigot, Liaison with HP-Procurve
Mats Möller, Liaison with Oracle

Other IT Department staff contributing to CERN openlab

Dirk Düllmann	DM Deputy Group Leader
David Foster	CS Group Leader
Maria Girone	DM group
Eric Grancher	DES Deputy Group Leader
François Grey	Communications team
Denise Heagerty	Security section
Christopher Lambert	DES group
Alberto Pace	DM Group Leader
Markus Schulz	GD Group Leader
Jamie Shiers	GS Group Leader

Industry Partner Liaisons with CERN (Technical/Management)

Peter Toft	HP (T)
Arnaud Pierson	HP (M)
Dan Ford	HP ProCurve (T)
Herbert Cornelius	Intel (T)
Stephan Gillich	Intel (M)
Russel R. Beutler	Intel (M) until February 2008
Claudio Bellini	Intel (M) from March 2008
Bjørn Engsig	Oracle (T)
Graeme Kerr	Oracle (T)
Monica Marinucci Lopez	Oracle (M)

Industry Contributor Liaison with CERN

Pär Andler	F-Secure
Mika Rautila	Stonesoft
Rolf Kubli	EDS
Max Böhm	EDS

From left to right, front row kneeling: Andrzej Nowak, François Fluckiger, Ryszard Jurga, Séverine Pizzera, Daniel Da Rocha Cunha Rodrigues, Sverre Jarp. From left to right, back row standing: Jose Dana Perez, Xavier Gréhant, Gyorgy Balazs, Dawid Wójcik, Milosz Hulboj, Håvard Bjerke.

THE RESULTS

Platform Competence Centre

Growing workloads and rising energy costs have made data centre energy efficiency a critical issue for many organizations, especially for those with older facilities that lack the power and thermal capacity needed to support growing needs. This is a challenge that is highly relevant to CERN. In 2007, a particular emphasis of the Platform Competence Centre – one of the four broad activity areas of CERN openlab – was on evaluating ways to improve the performance of the CERN Computer Centre.

The computing model for the LHC data analysis and simulation is based on a high throughput implementation where independent jobs can efficiently be executed in parallel in a computer with multiple processors, provided there is sufficient memory. Because of this model, it is relatively easy to expand total capacity by adding more computers. Unfortunately, thermal issues in today's computers limit the capacity of CERN's current data centre. This represents a real challenge to the full deployment of the required resources.

For the CERN Computer Centre, global measurements have shown that one Watt saved on server power consumption has the added benefit of saving almost an additional Watt on power requirements for the cooling system. In other words, one Watt saved on power consumption translates into a total saving of almost two Watts. For many organizations, achieving a similar reduction would translate into direct savings. However, for CERN the benefits are indirect, since the computing needs of the physicists are basically without limits. Any power-efficiency gains will simply be used to install additional computers.

In this context, the multi-core evolution, which enables a significant increase in computing power within a constant processor power envelope, is particularly beneficial for CERN. For years, the heart of CERN's computing resources consisted of hundreds of Servers based on two earlier generation Intel® Xeon® Processors. Although frequency increases in successive processor generations delivered the required performance gains, the corresponding rise in the amount of dissipated power was a serious problem. Fortunately, the move to multi-core processors enabled ongoing improvements in overall performance, without a corresponding increase in processor power consumption. Although the amount of memory had to be kept constant per core, the power savings compared to a non-multicore scenario have been impressive.

Like other computer centres, CERN uses the SPECint benchmark developed by the Standard Performance Evaluation Corporation (SPEC). Dividing a computing system's SPECint result by the amount of power it consumes provides a SPECint/Watt ratio, an important parameter for assessing the performance of the CERN Computer Centre. Around 2004, CERN had been using mostly enterprise class Intel processors, namely Xeon® processors based on Intel Netburst® microarchitecture. The systems using these processors, which made up the backbone of CERN's computing centre at the time, had a SPECint/Watt ratio of about 6.

By early 2008, there were about 2400 compute nodes used for physics applications in CERN's computing centre: about 400 single-core systems, about 700 dual-core systems, and about 1300 quad-core systems. With the introduction of the new Intel® Xeon® dual core system based on Intel CoreTM Microarchitecture, the SPECint/Watt ratio rose to values between 20 and 30. Today, Intel's quad-core systems are beginning to dominate in the Computer Centre, since they provide a slightly better SPECint/Watt ratio to dual-core, while occupying less physical space. It is expected that Intel's next generation quad core Xeon® processor will improve the SPECint per Watt ratio further. Much of the testing of these new systems has been carried out in the context of the Platform Competence Centre.

CPU and disk array aisles in the CERN Computer Centre

Left : View of the CERN Computer Centre, April 2008

A survey of strategies for limiting power consumption in the CERN data Centre was carried out for the Platform Competence Centre and has led to list of recommendations and forecasts, based on different technological choices (see Page 9). One important conclusion is that without the move to Intel® Core™ microarchitecture, there would have been no way to maintain the data centre power consumption within its bound of 2.5MW.

As well as the focus on data centre performance issues, the Platform Competence Centre continues to pursue a range of specific technical challenges, ranging from compiler optimization to competing in the Top500 Supercomputer rankings. The Top500 runs are also used as "burn-in" tests for our new servers, since the tests are rather demanding. One focus area is performance monitoring, a joint project with HP Labs which aims to ensure that the perfmon2 interface, originally developed for Itanium processors, is integrated into the Linux kernel for all hardware platforms. CERN openlab is contributing to the thorough and intense testing on Intel Xeon® and Itanium® processors, helping to increase sophistication in pfmon, the user tool, by ensuring comprehensive symbol resolution, and developing a sophisticated and intuitive graphical user interface, gpfmon.

Another area where CERN openlab continues to contribute is in compiler optimization, where the aim is to improve performance of a wide range of different jobs by influencing the back-end code generator. The corresponding test suites used are based on millions of lines of C++ source code used for high energy physics. The current work in this area is targeting further improvements in execution time and expanding to more complex benchmarks.

Finally, in 2007 a major push was made by CERN openlab in the area of multi-threading, which exploits the new multicore processors in a more fine-grained manner, and is an important complementary approach to the traditional multiprocessing for improving the performance of high throughput computing. Two heavily over-subscribed workshops on multi-threading were successfully run in 2007 for CERN staff and participants in the LHC experiments. The workshops consisted of one day of lectures and one day of practical exercises. Five lecturers, two from Intel and three from CERN, taught classes of 45 persons.

The plan is to continue these workshops in the future and the next one is planned for end-May 2008. Tutorials on computer architecture and performance monitoring are also being established in the Platform Competence Centre.

As well as the educational initiative, multi-threading has been used to make advances in the parallelization of the Geant4 toolkit for the simulation of the passage of particles through matter. Geant4 is widely used in high energy physics, but also in medical and space science. To prepare for the future as LHC gets underway, openlab started several initiatives with Intel to look at future processors and possible future languages to support the processors. This work is expected to be one of the foundations for the next openlab period.

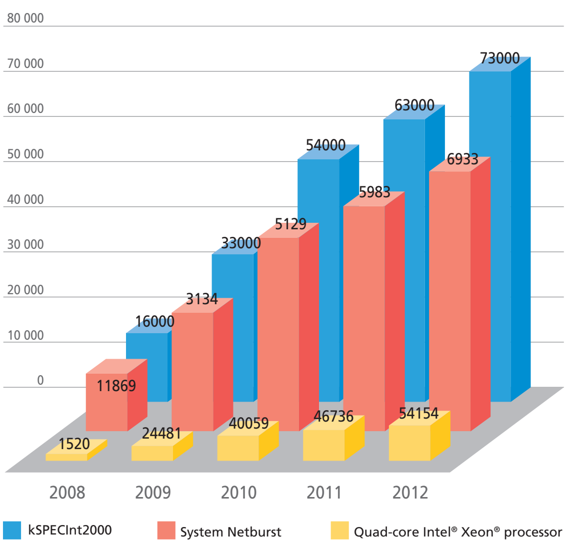
Herbert Cornelius teaching at the multi-threading workshop

Energy saving strategies

The conclusions of a study carried out as part of the Platform Competence Centre programme of work are listed below. The following strategies are expected to prove valuable in the CERN computing environment, and may be useful to other organizations seeking to increase computing capacity, while reducing power consumption and total cost. A more detailed description of these strategies is given in the CERN openlab report "Strategies for increasing data centre power efficiency" (February 2008). Two key forecasts for power consumption, taken from this report, are also shown below.

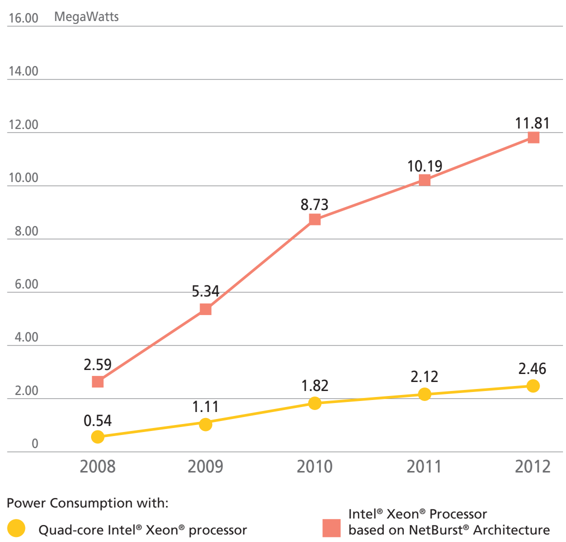
1. Make power efficiency part of the server tendering process, by including the impact of power consumption in estimating total lifecycle costs.
2. Carefully select new servers that have optimal power consumption for the workload at hand.
3. Optimize your data centre for power efficiency.
4. Evaluate software technologies for improving the overall energy efficiency of your systems and data centre.

Projections of CPU requirements - CERN



The number of servers required to meet CERN's projected computing requirements through 2012. The back row shows CERN's projected computing requirements. The middle row shows the number of Netburst based servers required to meet those requirements. The front row shows the number of quad-core servers (based on the Intel Core microarchitecture) required to deliver the same level of performance.

Projected Power Consumption for CPU nodes



Projected power consumption for the main CERN data centre through 2012. If CERN were using only quad-core processors based on the Intel Core microarchitecture rather than single-core processors based on the older Netburst microarchitecture, it would reduce power consumption by nearly a factor of five, as shown here.

Grid Interoperability Centre

The Grid Interoperability Centre focuses on testing and validation of partner software solutions in connection with the EGEE project and its middleware stack, gLite. One of the most significant advances for Grid technology in the last couple of years has been the progress of virtualization techniques, which provide a practical way to achieve a central part of the 'Grid vision' of on-demand computing, namely Grid resources that are independent of the software environment needed to run applications.

Over the last year, CERN openlab has helped raise awareness of virtualization techniques in the high-energy physics community, and demonstrating the specific advantages of the Xen virtualization open source platform. HP is a long-time contributor to Xen and the first major company to work with the platform. For CERN openlab, the focus has been on system testing and the use of Xen for server consolidation, improving flexibility and potentially also security on the Grid. Other work in this area includes the development of an interface called OSFarm for generating operating system images for use with Xen Virtual Machines, a methodology for transferring virtual images between systems (Content-Based transfers), as well as benchmarking of virtual machines.

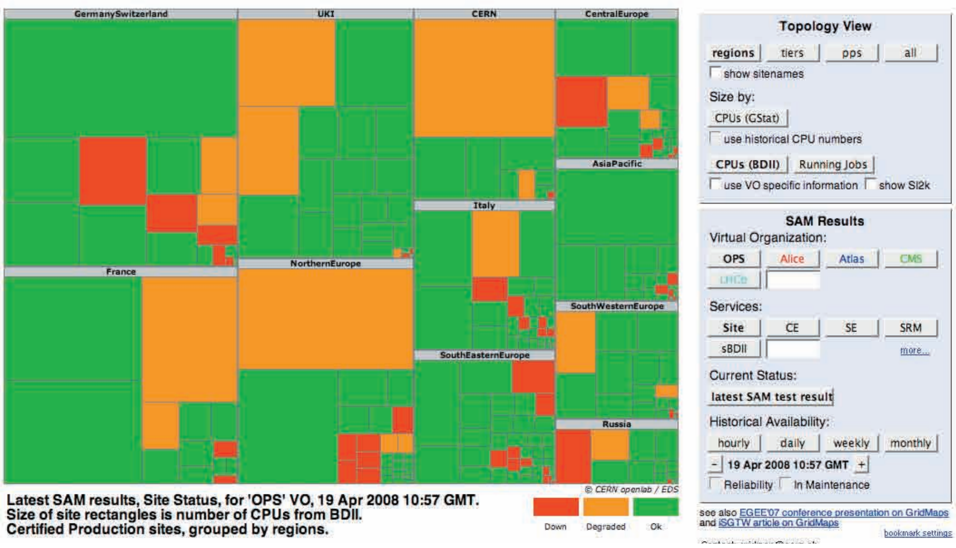
The integration of Tycoon, an open source market-based system for allocation of Grid resources developed by HP, with gLite was successfully carried out in 2007. This resulted in an interface, "Tycoon-gLite", allowing easy deployments of Grid elements in a Tycoon environment. Work is ongoing to promote the use of this environment by Grid administrators.

A study of Grid resource scheduling techniques is being carried out by CERN openlab. This work is being done in collaboration with several Grid experts at CERN, and covers different Grid infrastructures and various approaches to scheduling. A resource model and a simulator, Levellab, have been developed, to integrate novel low-level performance prediction models and evaluate the efficiency of grid scheduling algorithms.

The CERN openlab contributor, EDS, has developed an intuitive overview of Grid resources, called GridMap, that allows operators to rapidly assess the status of the Grid. This top-level visualization has been deployed in prototype form since October 2007, and is already widely used both at CERN and at many other Grid nodes. It allows for views of availability, reliability and maintenance status of resources. GridMap is also being used to monitor critical services during the final tests of the LHC Computing Grid prior to start-up of the LHC.

In order to support the testing that goes on in various CERN openlab activities, specifications have been developed for a new HP/Intel blade system, with 128 Xeon processors and capacity to expand with future processor technologies over the next years. The new blade system will be used for a variety of purposes, including virtualization tests, and testing of Grid-specific software such as HP's Tycoon, as well as for compiler tests, performance monitoring of HEP applications, and hands-on teaching in future workshops that CERN openlab will run.

In the first phase of CERN openlab, a similar workhorse was established, called the CERN opencluster, based primarily on Itanium technology. The opencluster continues to be actively used and is very reliable. Current usage focuses on support of urgent computational fluid dynamics calculations being run to ascertain the heat flow around the huge LHC experiments in their underground caverns. This is to ensure, prior to the LHC startup, that there are no potentially dangerous hotspots.



GridMap Prototype Visualizing the "State" of the Grid

Opposite page: CPU aisle in CERN Computer Centre

CERN openlab

Relational Database Activity

The CERN openlab Relational Database Activity focuses on testing and practical implementation of new database techniques for the LHC Grid environment, as well as for CERN's general IT services. During the last year, CERN openlab has extensively used Oracle Data Guard to migrate existing servers and storage to new hardware. As a result, a complete technology upgrade based on dual-CPU quad-core servers has been obtained with minimal downtime. CERN database administrators create a physical standby database on the new technology platform, then execute an Oracle Data Guard switchover to move production to the new platform. The new platform is easily tested prior to assuming the production role. If necessary, production can be quickly switched back to the original system without data loss.

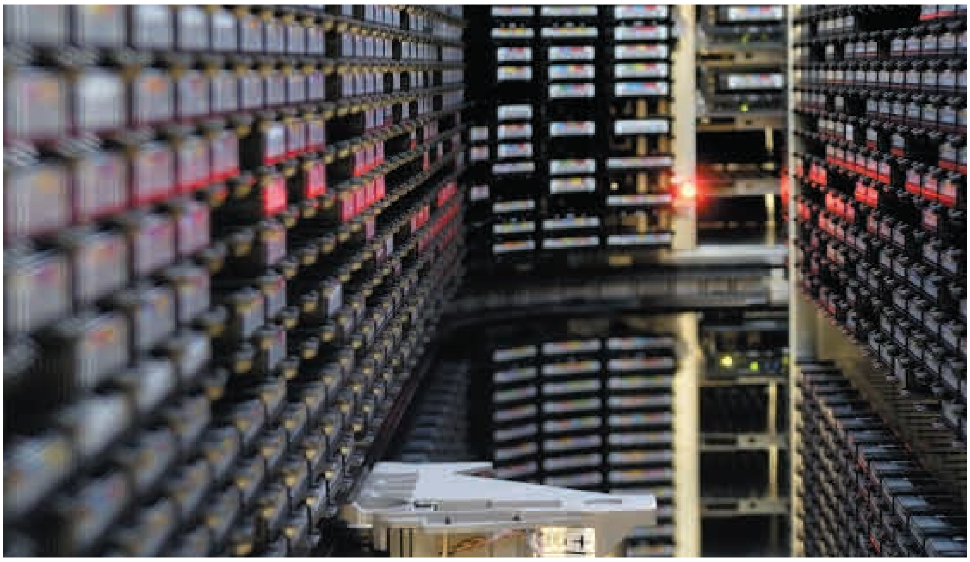
As many other sites use similar hardware to CERN, this technology is proving of interest for other data centres using Oracle Real Application Clusters (RAC). In the multi-Tier Grid computing model foreseen for the LHC experiments, the tasks of the Tier-0, located at CERN, are the raw data recording, first pass processing and distribution to external Tier-1 sites. The 11 Tier-1 sites perform all subsequent reprocessing, as well as additional tasks that depend on the precise computing models of the various experiments. Oracle Database RAC clusters of typically 32 cores will be connected via Oracle streams replication from the experiment sites to the CERN Computer Centre, as well as to the Tier-1 sites, creating a world-wide distributed database network.

CERN uses a downstream database capture configuration in order to decouple Tier-0 production databases from Tier-1 replicas, and avoid network problems. As part of the CERN openlab activity, fail-over procedures have been implemented to isolate Tier-1 interventions and downtime. In collaboration with the Tier-1 sites, CERN has also evaluated many recovery scenarios at Tier-0 and Tier-1 sites and implemented and tested the appropriate procedures needed to synchronize the streams databases after a recovery situation.

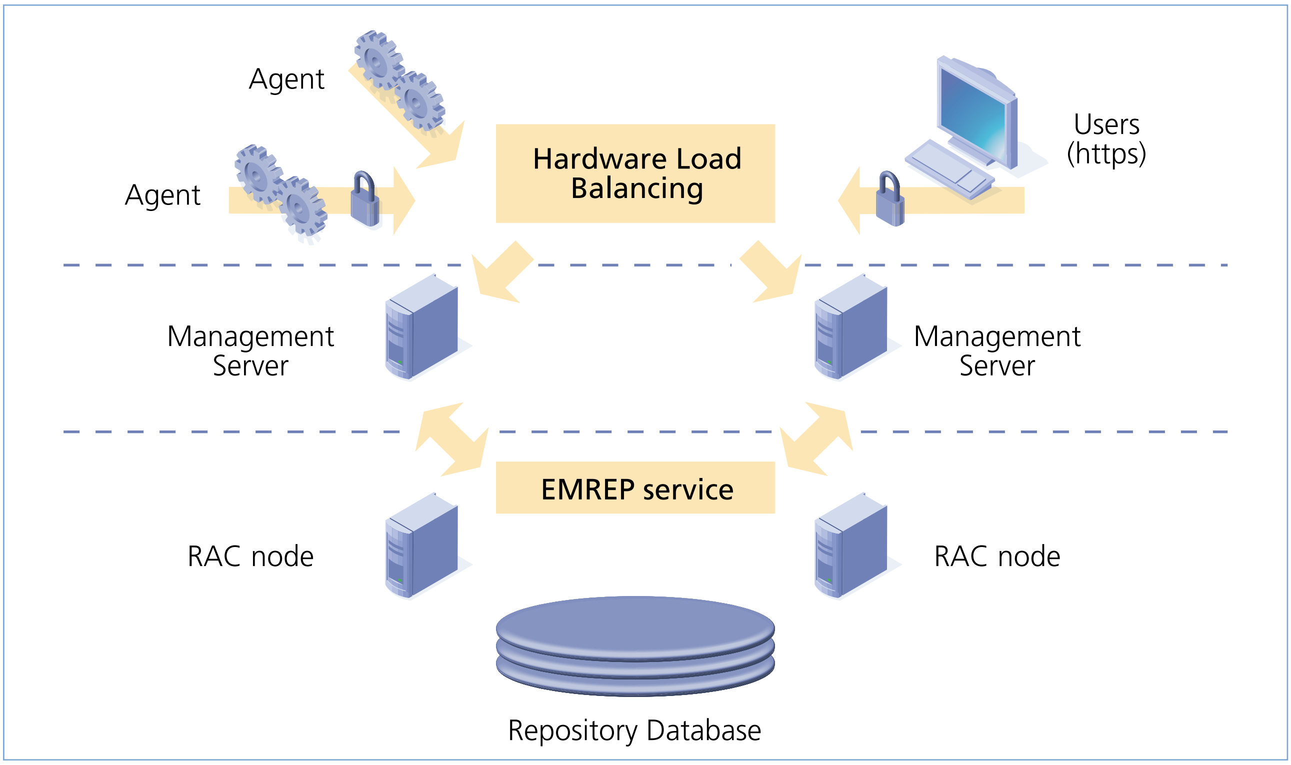
In close collaboration with the Oracle Streams developers important optimizations have been implemented to achieve significant throughput improvements. As a result, the replication rate over the Tier-1 sites has been increased by almost a factor 10. For the ATLAS Streams setup, filtering rules were also re-written increasing more than threefold the replication throughput. Regular monitoring of the databases and streams performance is available to the LHC experiments and Tier-1 sites, based on Oracle Enterprise Manager and an Oracle Streams monitoring tool developed at CERN

CERN openlab has also contributed in greatly improving our monitoring capabilities through the use of Oracle Enterprise Manager 10g (EM10g). We successfully migrated to the most recent Linux platform, while at the same time implementing a technical architecture capable of handling the ever increasing data volume generated by CERN's expanding adoption of Oracle EM. These measures help CERN's IT Department to maximize the benefits of EM10g and to improve their service delivery and service quality through problem avoidance and real time problem analysis. The latest architecture for EM10g is illustrated in the figure on the next page.

Part of our implementation has been focused on extending the out-of-the-box monitoring functionality. One example is that we now actively monitor for any unwanted changes that could compromise the SYS schema, which stores all views and base tables of the Oracle data dictionary. Another example, which was highly appreciated at the Oracle OpenWorld joint



Database systems for the LHC Computing Grid must keep track of files stored both on disk arrays (opposite page) and in tape storage systems (this page) at CERN and in the 11 major Tier-1 centres.



The new Oracle EM10g architecture used at CERN. There are 825 targets, 28 cluster databases, 140 database instances and 146 hosts.

presentation with the Configuration Management team, is the monitoring of backup activity and subsequent alerting of failures to meet pre-defined expectations. Overall, such improvements are giving administrators greater confidence in the stability of their services, and their ability to meet agreed service levels. This in turn allows them to concentrate more on adding value to those services rather than firefighting.

Virtualization of Oracle RAC using the Xen platform was another project undertaken in 2007 by a CERN openlab student. Oracle RAC was successfully installed in the virtualized environment and performance tests were done to compare virtualized setup versus classical single instance setup. Virtualization gave a better usage of hardware, less power consumption and a better isolation of the user applications. The only drawback was the lack of support for Oracle database running in this type of environment. But since the study was performed by CERN openlab, Oracle has announced its own Xen-based virtualisation software - Oracle VM.

Another area of work concerns the use of Oracle Relational Database Management System (RDBMS) to make comparisons of different Intel processor technologies. This sort of test and validation project is an example of the collaborative environment that CERN openlab provides its industrial partners. The performance results with dual-socket systems showed more than 70% improvement of the quad-core processors over the previous dual-core generation.

An area of activity has been on the optimization of PVSS, a supervisory control and data acquisition system used widely in the LHC and associated experiments. Crucial LHC hardware is dependent on PVSS. The out-of-the-box performance for running PVSS with Oracle database allowed archiving of only 100 changes per second, but through optimization of the application code and usage of advanced Oracle features like RAC and partitioning, CERN openlab was able to achieve a practical target of 150,000 changes per second at a stable throughput while queries are being performed on the system.

Throughout the year, under the joint software testing initiative, we have tested early beta releases and various new functionalities of the Oracle database. Particular focus has been on Real Application Testing, where the Workload Capture and Replay feature is used to provide an additional level of convenience and confidence in the databases upgrade process. This has been proven with both PVSS and CASTOR name server workloads.

Networking and Security Activity

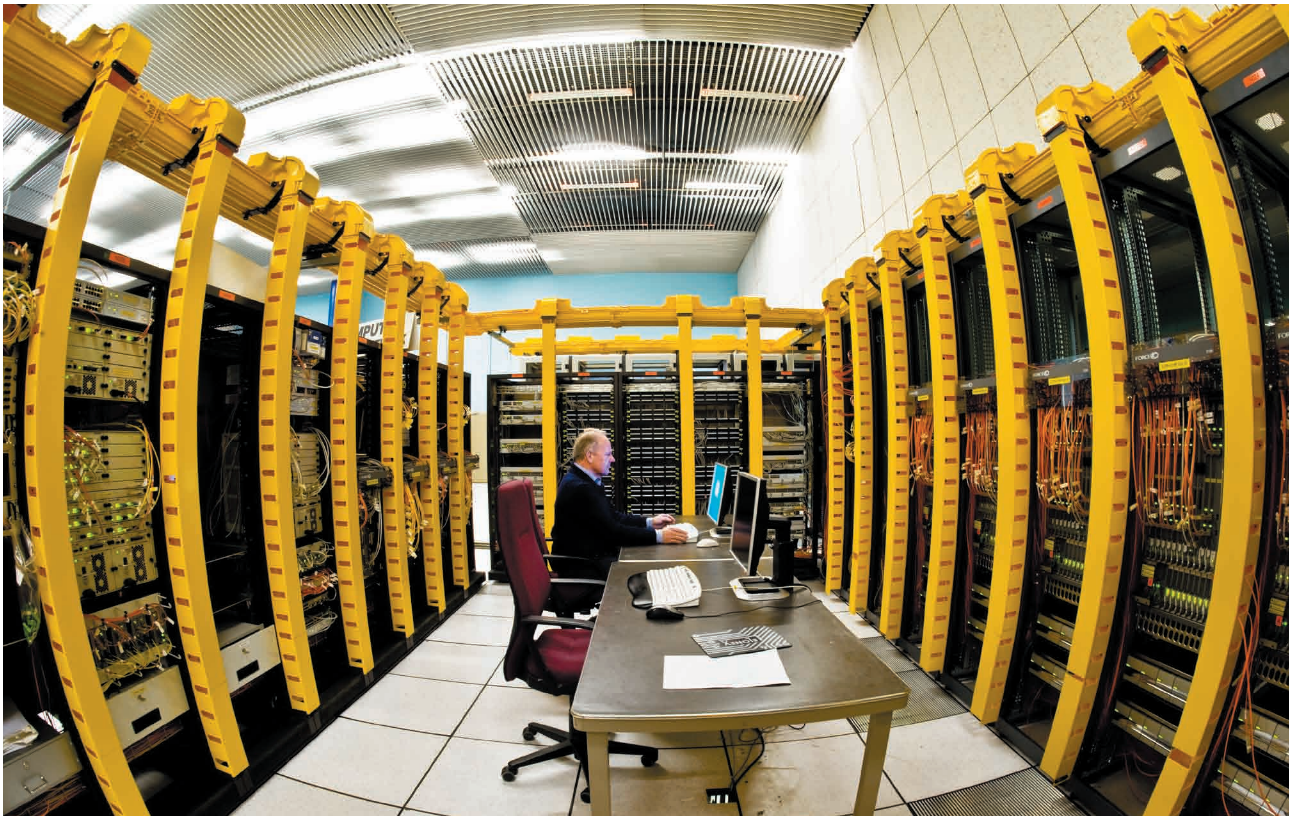
A major CERN openlab project was launched in 2007 in collaboration with ProCurve Networking by HP. Called CINBAD (CERN Investigation of Network Behaviour and Anomaly Detection), the project mission is to understand the behaviour of large computer networks (10,000 or more nodes) in the context of high performance computing and large campus installations such as CERN. The goals of the project are to be able to detect traffic anomalies in such systems, perform trend analysis, automatically take counter measures and provide post-mortem analysis facilities.

This is a challenging research activity as it must address large scale issues, requiring collection and storage of large quantities of data. The starting point in 2007 has been to define requirements and ensure a common understanding of precise definitions, for example of what constitutes an anomaly or a trend. The project is tentatively divided into three phases: data collection and network management; data analysis and algorithm development; performance and scalability analysis.

Work has begun in order to implement a high performance collector and associated storage, as well as a configuration mechanism to setup agents in order to collect traffic samples from all CERN devices for several days, and identify potential performance issues. Analysing the stored data will require defining the characteristics of "zero-day" anomalies, understanding the possible correlation of the data with, for example, antivirus and intrusion detection systems, network incidents, layer 3 topology and route changes. Work is underway to identify the best technologies for database storage and further analysis of the historical data.

A separate networking project has been the continued evaluation of Intel 10Gb/s networking cards. The first generation of these cards had been successfully tested with high-throughput disk servers, but cost as well as throughput issues remained. Meanwhile, production disk servers with 1Gb/s network interface cards have now reached their limit in terms of throughput and capacity. CERN openlab tests have shown that second generation 10Gb/s cards have much better characteristics for CERN's needs, and in particular native speed (9.49 Gbp/s) was reached, driver support for Linux kernels is available and reasonable cost levels have been attained.

View of networking hub in CERN computer centre, with campus backbone (left) and LCG backbone (right).



CERN
openlab

Publications and Presentations

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

PRESENTATIONS:

- X. Grehant and I. Demeure, SmartDomains, 2nd authoring workshop, EuroSys 2007, Lisbon, Portugal, March 2007
- J.M. Dana, CERN Snippets Dissected, Gelato ICE '07, San Jose, California, USA, April 15-18 2007
- Sverre Jarp, Perfmon2: A leap forward in performance Monitoring, PH/SFT Many-Core Workshop, CERN, April 16th
- J.M. Dana, Tycoon market-based system: integration with EGEE, 2nd EGEE User Forum, Manchester, UK, May 9-11 2007
- C. Lambert, EM10g implementation at CERN, Oracle Enterprise Manager User Group, Munich, Germany, June 2007
- S. Jarp, How good is the match between LHC software and current/future processors, CHEP07 Plenary session, 5 Sep 2007
- X. Grehant, SmartDomains, Invited talk for Kite Special Interest Group in Virtualization, London, September 2007
- M. Böhm and R. Kubli, Visualizing the State of the Grid with GridMaps, EGEE'07 Conference, Budapest, Hungary. Oct. 2007
- J.M. Dana, An update on HP's Tycoon at CERN: integration with EGEE, EGEE '07 Conference, Budapest, Hungary, 1-5 October 2007
- L.M. López, C. Lambert OOW, Securing CERN Oracle Databases using Grid Control, Oracle OpenWorld, San Francisco, USA, November 2007
- E. Grancher, A. Topurov, Real life experiences of RAC scalability with a 6 node cluster, UK Oracle User Group Conference, Birmingham, UK, December 2007
- Sverre Jarp, Parallel Processing Model and Research at CERN, Gelato Many-Core Workshop (Shanghai China / Fudan University, March 27-28)
- Sverre Jarp, Issues facing CERN and HEP in the manycore computing era PH/SFT Many-Core Workshop, CERN, April 14th

PUBLICATIONS:

- X. Grehant, O. Pernet, S. Jarp, I. Demeure and P. Toft, Xen Management with SmartFrog, VHPC'07, August 2007
- A. Nowak, openlab/Intel workshop offers programmers multicore training, CERN Computer Newsletter, Sept. 2007
- S.Jarp, R.Jurga, A.Nowak, Perfmon2 : A leap forward in Performance Monitoring, Proceedings of CHEP07, Vancouver, Canada, September 2007
- M. Girone, J. Wojcieszuk, D. Wojcik, G. Kerr, Cross-Platform Database Migration with Oracle Data Guard: a CERN Case Study, Oracle Technology Network (October 2007).
- M. Böhm and R. Kubli, Link of the week - Visualizing the state of your grid with GridMaps, iSGTW, Oct. 2007; CNL, Nov. 2007
- H. K. F. Bjerke, D. Shiyachki, Andreas Unterkircher, Irfan Habib, Tools and Techniques for Managing Virtual Machine Images, submitted to 3rd Workshop on Virtualization in High-Performance Cluster and Grid Computing (VHPC '08)
- I. Habib, H. Bjerke, A. Anjum, R. McClatchey and P. Bloodsworth, Benchmarking Virtualization Hypervisors for e-Science, submitted to 3rd Workshop on Virtualization in High-Performance Cluster and Grid Computing (VHPC '08)
- A. Dechert, IT-Service Management at CERN and How It Can Be Improved by the Usage of Oracle Enterprise Manager, Diploma Thesis, Karlsruhe University of Applied Sciences 2008

POSTERS:

- X. Grehant, J.M. Dana, and H. Bjerke, Grid and Virtualization Experience in CERN openlab, EuroSys 2007, Lisbon, March 2007
- E. Grancher, A. Topurov Oracle RAC (Real Application Cluster) application scalability, CHEP07, Vancouver, Canada, Sept. 2007
- M. Böhm and R. Kubli, Top-level Grid Services Monitoring Visualization, EGEE'07 Conference, Budapest, Hungary. Oct. 07
- J.M. Dana, Tycoon @ CERN, Distributed Computing Workshop, London (UK) May 21, 2008

CERN OPENLAB REPORTS:

- X. Grehant and J.M. Dana, Providing Virtual Execution Environments: a Twofold Illustration, May 2007
- R. Erazm Jurga and M. M. Hulboj, Packet Sampling for Network Monitoring, December 2007
- A. Hirstius, S. Jarp, A. Nowak, Strategies for increasing data centre power efficiency - An overview of CERN's approach to energy efficient computing, February 2008

TEACHING:

- CERN openlab student programme Teaching series, July-August 2007. Nine lectures: **Server hardware**, A. Hirstius/ CERN; **Linux kernel**, J. Iven/CERN; **Oracle RDBMS**, B. Engsig/Oracle; **Computer Security**, S. Lopienski/CERN; **Compilers**, L. Pollock/ U. Delaware; **Benchmarking**, S.Jarp/ CERN; **Networking**, M. Swany/ U. Delaware; **Virtualization**, H.Bjerke/CERN; **LCG**, L. Poncet/CERN; **gLite**, M. Schulz/CERN.
- X. Grehant. Grid Computing. ENST Paris, in Distributed Systems. April 2008

Events and Outreach

As well as the many excellent technical results that CERN openlab provides, the partnership 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 to a wider audience, including partner clients, the press and the general public.

CERN is regularly visited by top delegations from governments and industry, as well as customer and press visits organised by openlab partners. These groups are briefed about CERN openlab in a dedicated VIP meeting room known as the CERN openlab openspace.

ProCurve Networking for HP Press Event, CERN, 11 September 2007
Introduction to CERN openlab as part of product launch event

Oracle 11g launch, 12 September 2007, CERN
Oracle Database 11g launch 'Innovate Faster' held at CERN

Oracle OpenWorld, San Francisco, October 2007
CERN participation in Intel CEO keynote speech, CERN presentations

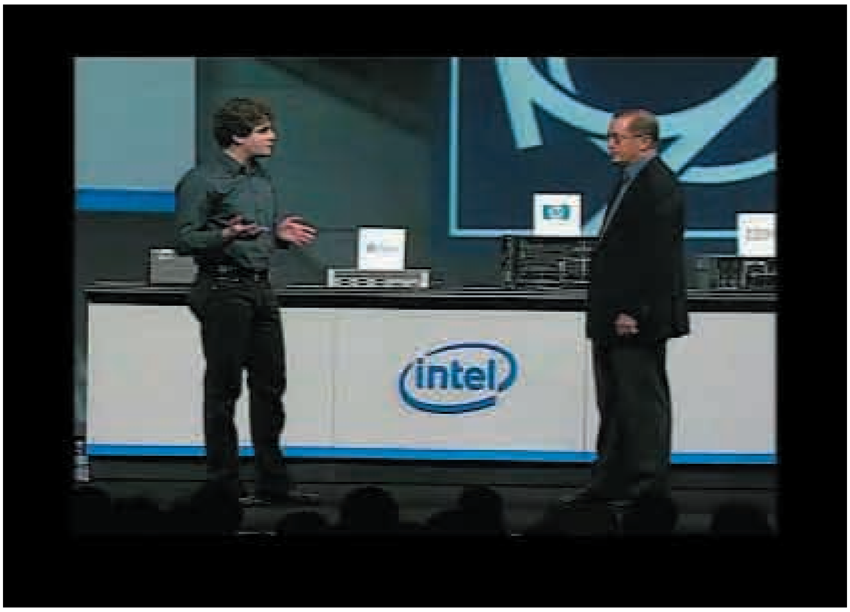
Intel Customer Visit, CERN, 14 November 2007
Tour of CERN and review of CERN openlab results for Intel corporate customers

First event on Computing for Finance, 21 November 2007, CERN
Joint public event by CERN openlab and EGEE, featuring speakers from banking and financial services sectors.

CERN-Oracle Event, 18 December 2007, CERN
Joint Oracle-CERN meeting and celebratory event to mark 25 years of collaboration.

From R&D to Innovation: Future Business Opportunities, 6 February 2008
Oracle VIP conference and roundtable discussion held at CERN

Procurve Networking for HP client events, 20 February and 10 April 2008
VIP and Swiss client events held at CERN, presentation of CERN openlab



video capture of Dirk Düllmann of the CERN IT Department on stage with Paul Otellini, CEO of Intel, at Oracle OpenWorld (October 2007).

Education

The CERN openlab student programme was launched in 2002, to enable undergraduate, Masters and Ph.D. students to get hands-on experience with Grid technology. In 2007, 25 computer science and physics students took part in the programme at CERN, coming from Europe, Canada and Brazil. The students work on Grid-related projects supervised by openlab staff as well as staff from the LCG and EGEE Grid projects, other Grid-related groups in the Department, and Grid-related activities in the rest of CERN.

Several of these students were co-funded by openlab partners. The students are listed below, with home institute and project topic. The programme included study tours to the HP-Intel Business Solutions Laboratory in Grenoble, as well as to the EPFL (Ecole Polytechnique Federale de Lausanne). A dedicated lecture series for the students was given by Grid experts in CERN's IT Department.

CERN openlab students during an excursion to the EPFL, Ecole Polytechnique Fédérale de Lausanne, August 2007

- Viacheslav Burenkov, Imperial College, UK (CERN School of Computing support)
- Lars Mikael Bärlund, HIP, Finland, (Digital pen guestbook)
- Serena Cameirano, University of Dublin, Trinity Coll., Ireland, (Provisioning computing resources with high availability)
- Maria Christoforaki, University of Patras, Greece (Content delivery and streaming in P2P overlay networks)
- Wojciech Czech, University of Science and Technology in Kraków, Poland (Grid Monitoring Messaging system)
- Alexandre Dehlay, ENST-Bretagne, France (Indico internationalization)
- Jorge Canizales Diaz, University of Madrid, Spain (Performance monitoring)
- Antonio Dominguez, University of Cadiz, Spain (Grid-based financial software)
- Manuel Entrena Casas, University of Granada, Spain (Monitoring interface system)
- Rodolfo Gabriel Esteves Jaramillo, University of Waterloo, Canada (Compiler optimization)
- Paulo Ricardo Motta Gomes, University Campina Grande, Brazil (Virtualisation (Xen) to set up gLite training infrastructures)
- Irfan Habib, University of the West of England, Bristol, UK (Platform virtualization in a Grid environment)
- Piotr Jurga, University of Bydgoszcz, Poland (Power measurements)
- Thomas Koeckerbauer, University of Linz, Austria (Xen for VGrid portal)
- Anais Labarre, University Lyon 1, France (Mobile Grid access for Satellite imagery)
- Katia Leal Algara, Rey Juan Carlos University, Spain (Digital Libraries Services on gLite)
- Maria Leitner, University in Vienna, Austria (Oracle Real Application Clusters Virtualisation)
- Jim Levy, University Lyon 1, France (Web Stastistics for J2EE Public Service users)
- Christopher Nater, University in Edinburgh, UK (Reviews and updates to GRID user documentation for EGEE and LCG)
- Eduard Pauna, University of Bucharest, Romania (Develop Service Availability Monitoring Security sensors)
- Robert Preissl, University of Linz, Austria (Porting physics software to BOINC)
- David Soriano, University of Madrid, Spain (Lightweight GUI for Ganga data location)
- Andrea Sottoriva, University of Amsterdam, The Netherland (Scalabitity analysis of Tycoon)
- Danica Stojiljkovic, University of Belgrade, Serbia (Automatic deployment of gLite User Interface with ETICS)
- David Weir, Imperial College London, UK (BOINC and ATLAS computing using Virtualization)
- Krzysztof Wyszynski, University of Silesia, Poland (Database Access Management framework extension)

The Board of Sponsors

This is the second annual report of the second phase of CERN openlab, which was launched in 2006. It was presented to the Board of Sponsors at the Annual Sponsors meeting, 24-25 April 2008. Present at the meeting were: (front row) François Flückiger (CERN), Arnaud Pierson (HP), Renaud Barillère (CERN), Sverre Jarp (CERN); (second row) Alberto Pace (CERN), Séverine Pizzera (CERN), Wolfgang von Rüden (CERN), Stephan Gillich (Intel), Graeme Kerr (Oracle), Mats Möller (CERN), Max Böhm (EDS); (back row) Jean-Michel Jouanigot (CERN), Andrew Bulloch (Oracle), Martin Antony Walker (HP), Michel Benard (HP), Bjørn Engsig (Oracle), Dan Ford (HP ProCurve), Pierre Klatt (EDS), Rolf Kubli (EDS).

CERN openlab partners

HP is a leading global provider of products, technologies, solutions and services to consumers and businesses. The company's offerings span IT infrastructure, personal computing and access devices, global services and imaging and printing. HP completed its merger transaction involving Compaq Computer Corp. on May 3, 2002. 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.

CERN openlab contributors

EDS is a leading global technology services company delivering business solutions to its clients. EDS founded the information technology outsourcing industry more than 45 years ago. Today, EDS delivers a broad portfolio of information technology and business process outsourcing services to clients in the manufacturing, financial services, healthcare, communications, energy, transportation, and consumer and retail industries and to governments around

For More Information:

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

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

http://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.

CERN
openlab

www.cern.ch/openlab

PARTNERS

hp
invent

intel

ORACLE

CONTRIBUTORS

EDS