CERN COMPUTER NEWSLETTER
Volume 43, issue 2 April–June 2008

Contents

Editors Natalie Pocock, CERN IT Department, 1211 Geneva 23, Switzerland. E-mail cnl.editor@cern.ch. Fax +41 (22) 766 9204.
Web cerncourier.com/articles/cnl.

Produced for CERN by IOP Publishing Dirac House, Temple Back, Bristol BS1 6BE, UK.
Tel +44 (0)117 929 7481.
E-mail jo.nicholas@iop.org.
Fax +44 (0)117 930 0733. Web iop.org.

Published by CERN IT Department

IOP Publishing

Information systems in HEP get INSPIREd

Representatives of the HEP and astrophysics information systems met at DESY in May.

The status of high-energy physics (HEP) information systems has been jointly analyzed by the libraries of CERN, DESY, Fermilab and SLAC. As a result, the four laboratories have started the INSPIRE project – a new platform built by moving the successful SPIRES features and content, curated at DESY, Fermilab and SLAC, into the open-source CDS-Invenio digital library software that was developed by the CERN document server team.

Here I explain how the project was born and how the INSPIRE system is being built.

Current HEP information systems

The different HEP information systems being actively used by physicists can be classified into three main categories:

- the community-based HEP scientific document servers: arXiv (LANL, now Cornell), SPIRES (SLAC, Fermilab, DESY), CDS (CERN), ADS (SAO/NASA, Harvard) and KISS (KEK);
- the publisher-driven scientific document servers, such as PROLA (APS) and ScienceDirect (Elsevier);
- the generalist services, such as Google Web, Google Scholar and other web search engines.

Community-based systems have obvious synergies and they have been collaborating regularly in the past, linking to each other, exchanging general information and metadata. About a year ago, CERN, DESY, Fermilab and SLAC decided to run a user poll jointly with the goal of better understanding the perceptions, behaviours and wishes of the end users of these information systems.

Advertised through online messages and posts to e-mail listboxes between 30 April and 11 June 2007, more than 2100 answers, corresponding to about 10% of the active HEP community, have been collected. Theorists (61%), experimental physicists (22%) and software engineers (6%) from all over the world (22% US, 10% Germany, 8% Italy, 7% UK, 5% CERN) have described their habits and needs in detail. With 83% of responses coming from those using HEP information systems several times per week (and 57% daily), the quality of the feedback has been outstanding.

The poll results [1] show that the community-based services are clearly predominant (91.4%), with SPIRES standing as the most used "first reflex" HEP information system (48.2%). Commercial services are very rarely used in the community (0.1%) and general web search engines (such as Google) are used to some extent, mainly by the younger generation (8% on average, but up to 20% for users with less than two years of experience in HEP). Of course, results returned by these generalist search engines were harvested from the community-based servers and document repositories.

Users have also given their preferences regarding existing functionalities, like access to full text and to citation information. They have given their wishlist of features that they would like to have in the coming years, such as Web 2.0 user-contributed tagging. It is interesting to find that a large proportion of users are willing to invest their time in such a community service.

From this comprehensive picture of the perceptions and needs of the end users of HEP information systems, the SPIRES collaboration, and the CERN Library and CDS, decided to investigate further how a closer collaboration could fully match the community expectations.

SPIRES and CDS-Invenio

The SPIRES-HEP database [2] started in 1974 and was based on SPIRES DBMS, using an IBM mainframe and command line interface. It is run today by SLAC, DESY and Fermilab. In the 1980s an e-mail interface was added and in the early 1990s it became the first US web server exposing deep web content of SPIRES-HEP. It was considered by Tim Berners-Lee as the "killer application" that showed what the Web could bring in the future.

High-quality metadata with human-proofed publication information, links to full text, author affiliations and much more has maintained the attractiveness of the service in the past years. The addition of citation services, such as the cite summary format, has provided physicists with a useful tool to follow up on the impact of documents with their peers. However, SPIRES now suffers from aging technology (SPIRES DBMS), which has resulted in scalability and maintenance issues.

The CERN Library has a history [3] of maintaining preprints since the 1950s. The CERN preprint server appeared on the web in 1993, and later it increased its scope to become the CDS in 2000, used at the same time as the interface to the CERN Library and as the CERN institutional repository (archiving multimedia, notes, conferences, etc). Two sister applications, a conference-management system (Indico) and digital library software (Invenio), were developed in parallel to improve the electronic archiving of the laboratory's assets. The Invenio package was released as open source in 2002 and started to be adopted by services within HEP (e.g. ILC) and outside HEP (e.g. HBZ), receiving valuable contributions from various communities (e.g. EPFL and UAB).

The top-quality metadata curation in SPIRES, combined with the highly performing and scalable software of CDS-Invenio, look like a perfect match to meet the user expectations that were expressed in the poll.

The birth of INSPIRE

In mid-May 2007 SLAC organized the first HEP/PPA Information Resource Summit, where the four laboratories (CERN, DESY, Fermilab and SLAC) decided to conduct a feasibility study of reproducing SPIRES data and features using CDS-Invenio software. The feasibility study ran until autumn 2007 and concluded positively.

Prof. Rolf-Dieter Heuer, CERN director-general designate, gave a keynote speech at the HEP Information Resource Summit, where the INSPIRE system was launched.

A new phase then started where all partners joined forces to replicate SPIRES user-level functionalities in Invenio. Some 760 000 records were converted and loaded. Citation features were empowered and the specific SPIRES syntax was simulated to ensure that users would be able to continue searching as they were used to. The website and search result formats were also configured to resemble the existing SPIRES look and feel.

A year later, DESY organized the second HEP/PPA Information Resource Summit, where the main people from HEP information systems participated. The current research director of DESY and CERN director-general designate Rolf-Dieter Heuer explained his vision of a next-generation HEP information system [4], open access and data publishing in the coming LHC era. The INSPIRE project was announced [5], the new platform explained and a call for further collaborations with all willing parties was initiated to prepare the next phases of the project.

Next steps

After reproducing SPIRES user features, the INSPIRE collaboration will focus in the coming year on cataloguer-level functionalities. Record-editing interfaces, checking and maintenance tools, inputting and harvesting workflows, and record-enrichment tools based on knowledge bases will be treated. Efforts will focus on building strong native tools to enable libraries from the four institutes to share the data-curation workload. It will be challenging to set up an optimized distributed cataloguing environment with the goal of reaching a level of metadata quality as high as the SPIRES standard, and at the same time be able to eliminate the current duplication of work in this area.

In the longer term (2009 onwards), the INSPIRE project will deploy more advanced features, some of them already available in Invenio and others still to be developed. Collaborative tools, such as baskets, alerts and tagging, are ready but they require user authentication throughout the whole HEP community. A coherent solution for all users should be addressed by the INSPIRE partners. Community-shared author/experiment/institute and conference databases will also be considered, in close contact with journal publishers and conference organizers who are likely to have similar needs.

The entire corpus of HEP literature in INSPIRE will be opened to new applications investigating novel text- and data-mining technologies, such as extended citations networks, combined impact metrics, the indexing of plots and tables, open-access dissemination and more.

Conclusion

The SPIRES collaboration and the CDS-Invenio open-source community are joining forces to build INSPIRE, a new HEP information portal, which will integrate present databases and repositories to host the entire body of the HEP literature, aiming to become the reference HEP scientific information platform worldwide. It will empower scientists with new tools to discover and access the results most relevant to their research, enable novel text- and data-mining applications, and deploy new metrics to assess the impact of articles and authors. In addition, it will introduce the Web 2.0 paradigm of user-enriched content in the domain of sciences with community-based approaches to the peer-review process.

INSPIRE represents a natural evolution of scholarly communication built on successful community-based information systems, and it provides a vision for information management in other fields of science. Inspired by the needs of HEP, we hope that the INSPIRE project will be inspiring for other communities.

References

[1] A Gentil-Beccot *et al. Information Resources in High-Energy Physics: Surveying the Present Landscape and Charting the Future Course* ArXiv 08.04.2701.

[2] L Addis *Brief and Biased History of Preprint and Database Activities at the SLAC Library, 1962–1994* www.slac.stanford.edu/spires/papers/history.html.

[3] A Pepe *et al. CERN Document Server Software: the Integrated Digital Library* CERN-OPEN-2005-018.

[4] R D Heuer *et al. Innovation in Scholarly Communication: Vision and Projects from High-Energy Physics* arXiv:0805.2789.

[5] Press release *Interactions* News Wire #37-08 *DESY: High-Energy Physics Labs Join to Build a New Scientific Information System* www.interactions.org/cms/?pid=1026243.

Jean-Yves Le Meur, (IT-UDS) on behalf of the INSPIRE project members

Data Management group faces new challenges

The Data Management (DM) group is ready to face new challenges in LHC computing. We interviewed Alberto Pace, who leads this group, which was created in January.

The IT-DM group meeting was held in May. The group has been created from other teams.

Why a data management group?
IT-DM brings together people in the IT department involved in data-management software development, with the aim of ensuring a coherent process for CERN data-management development.

The group has been created from other teams and includes the Physics Database Service, the development of the LHC Computing Grid (LCG) persistency framework (POOL, CORAL and COOL), the CASTOR team and the developers of the Grid data-management software (DPM, FTS, LFC and associated Grid tools).

What are the activities of the DM group?
The group hosts the Physics Database Services, providing Oracle-based services to the CERN physics community, in particular the LHC experiments and Grid services. The staff members have expertise in providing turnkey database solutions that meet the high reliability and scalability requirements of the experiments. The database services also host research activities on leading-edge technologies that are carried out in collaboration with industry and openlab, in the fields of database replication (between CERN and remote computer centres), performance monitoring, alarms and automated service recovery, disk-based database back-up and fast data-restore techniques. These services are used by the LHC experiments for applications related to detector construction, calibration, bookkeeping, file transfer and event-level metadata.

We aim to provide a persistent and vendor-independent framework to help the experiments in writing physics applications with our software-development projects (POOL, CORAL and COOL). The CERN software libraries and the common support of shared components save dedicated resources, improve the longevity of experiment software and facilitate development. The DM group also provides consultancy for the optimization of database access strategies. Other users of these services are the LCG file catalogue and LCG 3D project (Distributed Deployment of Databases for LCG), which is a joint project between IT, experiments, Grid projects and LCG sites implementing a distributed database service across LCG Tier0 and Tier1 sites.

DM's other main activity is data-management development, including software for data storage (CASTOR) at the Tier0 and some Tier1s and its equivalent for Tier2s; the Grid software to transfer the LHC data between high-energy physics computer centres (file-transfer service, FTS); and the LHC file catalogue (LFC).

This software must be able to handle the 15 million gigabytes that the LHC experiments will send to the CERN Computer Centre. "Handle efficiently" means being able to transfer this volume of data from the experiment, store it on the CERN central disk servers, save a magnetic tape back-up copy and replicate the data to the 11 Tier1s participating in the reconstruction and analysis.

What are the current priorities of the group?
All efforts are focused on LHC start-up final preparation and tests, simultaneously involving the Grid computer centres worldwide via the second phase of the Common Computing Readiness Challenge 2008 (CCRC'08). With experts in physics database operation, the developers of the Tier0 software and those of the Grid data management, we have the ingredients required to evolve towards a coherent strategy aiming to consolidate the existing software. The new structure also bridges the gap between the "physics data" and the "database for physics" activities by creating more awareness of potential cross-technology solutions in both teams.

The group has frequent communications with all major stakeholders (e.g. contact with the IT-DES group database experts and daily meetings with the computer centre deployment and operation team in IT-FIO). Excellent communication channels also exist with both Grid groups (IT-GD and IT-GS), as well as the LHC experiments through the WLCG meetings.

What are the challenges for the service?
The database service for physics is in charge of the operation of the LHC experiment offline databases. Unfortunately the current service does not allow a formal 24/7 on-call support service, but this could be possible if the scope is extended to the experiment's online database. This is currently being discussed with the LHC experiments.

Another aim is to increase synergies with the infrastructure service where additional economies of scale could be achieved with an improved set of common technical solutions and operational procedures.

In addition, the service infrastructure procedures that were developed and tested for the data service set-up and replication are entering a consolidation phase. The LCG 3D project is reaching its end with plans to move its service to final production with only maintenance effort.

How do you see the future of DM software?
The group maintains software coming from three code bases that have been developed partially independently, offering architecture-simplification opportunities leading to reuse of some existing solutions with minimal recoding. For example, the security implementation of DPM can be reused for CASTOR and the two RFIO (remote file input/output) implementations will converge. Similarly, the development experience of database applications is also available for data management.

A working group led by Dirk Düllmann is reviewing the existing architecture with a view to consolidating the software. Tests have started on various high-performance storage solutions to prepare for future challenges required by the LHC experiments. Similarly a review is foreseen on the magnetic tape usage and data access patterns to improve read efficiency by aggregating multiple requests in common operations. Finally, the various database schemas used to store the LFC and the CASTOR name servers will be reviewed. This effort should put the CERN Computer Centre in a position where additional computing capacity could be satisfied with costs proportional to the requested size increase.

Although improvements are possible through architectural changes and consolidation efforts, there is a strong belief and evidence that the LHC data-management software developed at CERN is ready for the experiment commissioning. This is why the entire IT-DM group is ready for the LHC start-up.

CERN receives Duke's Choice Award for its Java development

CERN was honoured by Sun Microsystems at the 2008 JavaOne Conference in May, with the presentation of a gold Duke's Choice Award for the entire collection of Java applications that CERN has developed for the installation and operation of the LHC. James Gosling (inventor of Java and Sun Microsystems vice-president) presented the award to Derek Mathieson (IT department) who was at JavaOne to collect the award on behalf of CERN.

As a Duke's Choice Award winner, CERN was given the opportunity to show some examples of its Java development in the exhibition hall during the week-long conference. Matthias Braeger (TS department) and Andrew Short (IT) staffed the booth with Mathieson, answering questions about the work of CERN from many of the 15 000 conference attendees.

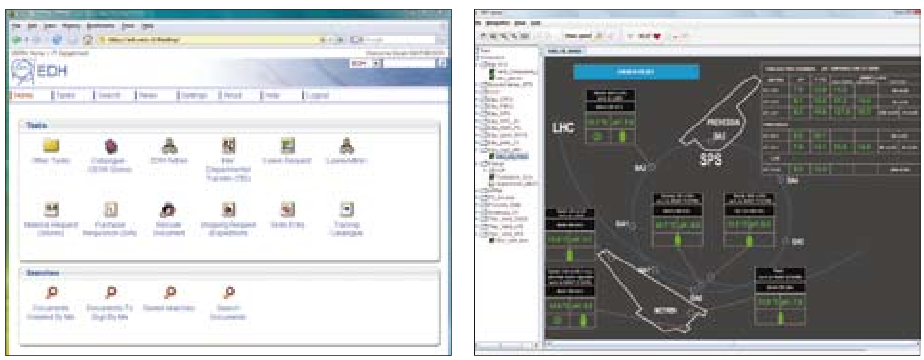
Derek Mathieson also participated in James Gosling's keynote presentation, where he gave a live demonstration of four Java applications to an estimated audience of 10 000 Java developers.

The first application demonstrated was EDH, which is known to most CERN personnel. "CERN's administration, as part of its Advanced Information Systems project, was an early adopter of Java technology with our first production code deployed more than 10 years ago", explained Mathieson. EDH currently consists of 1 million lines of Java, covering 50 procedures with more than 11 000 active users.

The next application was CERN's Technical Infrastructure Monitoring (TIM) application. This tool was developed in the TS department – currently maintained by Braeger – for monitoring CERN's technical infrastructure. More than 34 000 measurement channels are managed by the application, with the results collected and displayed on operators' consoles via a Java Swing application. The tool enables operators to monitor critical systems, such as electricity and water supply, and cooling, 24 hours a day throughout the CERN site.

The third application was the GraXML event viewer made by a Java developer in the ATLAS collaboration – Julius Hrivnac from the PH department. The tool permits a 3D visualization of the Atlas detector geometry with event (collision) data overlaid. With simple movements of the mouse, the detector can be manipulated in 3D space and the reconstructed particle tracks clearly seen as they traverse the various components of the detector.

Derek Mathieson receives the Duke's Choice Award on behalf of CERN from James Gosling.



CERN's EDH application (left) and the Technical Infrastructure Monitoring tool (right).

The GridPP real-time monitor (RTM).

Finally, the GridPP real-time monitoring (RTM) tool was shown. Developed at Imperial College, this tool provides a near real-time view of Grid activity, visualized on a 3D view of the Earth. Data centres are depicted as coloured discs and messages as coloured lines.

Mathieson went on to explain that these four applications are only a small sample of the work done in Java at CERN. Almost all of CERN's administration, as well as key components of the accelerator control software, are written in Java. In the past, Java had a reputation for being slow compared with traditional languages, such as C++ and Fortran. Today, however, with advances in compiler technology, it equals and sometimes outperforms C++ in many benchmarks. This fact, coupled with its first-class security model and threading support, make it an excellent choice for both server-side and desktop applications.

Useful links

EDH: http://ais.cern.ch/apps/edh.
TIM: http://timweb.cern.ch.
GraXML: http://cern.ch/hrivnac/Activities/Packages/GraXML/.
GriddPP RTM: http://Gridportal.hep.ph.ic.ac.uk/rtm.

Derek Mathieson, IT-AIS

IT department shows team spirit on Open Days

Fig. 1. (left) Maria Girone at the Grid Café presenting computing technology to the public. Fig. 2. (right) An image from the new multimedia history of the CERN Computer Centre.

The IT department played a big role in the success of the CERN Open Days held on 5–6 April. CERN opened its doors to 23 000 visitors on the Saturday, which was a day for family of CERN staff to visit. More than 53 000 visitors came on the Sunday when the general public were welcome. The event was billed as the last chance to see the LHC before its scheduled start-up this summer, and the record numbers for such an open-day event show that the public were keen not to miss the chance.

A team of more than 70 volunteers from IT worked hard to show as many people around the computing centre as possible, and to present Grid and computing technology in a Grid Café expo in Restaurant 2. More than 1000 visitors passed through the Computer Centre on the Sunday and about 500 on the Saturday. Several thousand people are estimated to have visited the Grid Café during the event.

In addition, the IT staff helped out with visits at many other visit points, including several of the LHC underground access points and an outdoor children's activity – the human beam dump – which proved very popular. IT staff developed a flow-management system that ensured that visitors had up-to-date information about where queues were the shortest, and an interactive public website that helped visitors to plan their tours in advance. Behind the scenes, many people in the department worked long hours prior to and during the Open Days, to ensure that all of the network and computing infrastructure needed to manage the ambitious event worked reliably.

As part of the extensive preparations for the Open Days, the historical computing material that is on display near the entrance of Building 513 was revamped into a proper Computing Expo for visitors to enjoy. New outreach material was prepared, including a multimedia history of the CERN Computer Centre (figure 2), video tours of the computer centre and related factsheets for training guides, and posters that trace the evolution of the Web and the Grid at CERN.

This Computing Expo and the experience that many IT staff gained from the Open Days will prove valuable in the future because the number of visitors to the Computer Centre is expected to increase substantially now that the LHC experiments are no longer accessible to the public.

The Open Days and the new Computing Expo would not have been possible without substantial sponsorship from a large number of the department's many suppliers and partners. In fact, these partners provided a large proportion of the overall sponsorship received by CERN for the event. Their generosity on this special occasion was particularly appreciated.

Francois Grey and Rosy Mondardini, IT-DI

openlab offers Intel software

Thanks to openlab's collaboration with Intel, a range of their software products for developers is available site-wide at CERN free of charge. One of the most prominent products in Intel's software tools line is the Intel C compiler (with debugger), which, among many features, introduces advanced optimization for the x86 and ia64 architectures. The Math Kernel Library and Integrated Performance Primitives packages are also available. In addition, the Intel Fortran compiler is on offer in the two most recent major versions. Those in need of performance gains will be happy to hear that Intel's optimization products have been published as well. The VTune environment is an advanced performance-tuning suite for Windows and Linux. It comes with Intel Thread Checker and Intel Thread Profiler – two products aimed at developers of threaded software. Both tools support pthreads and OpenMP.

All of the packages are available for Windows and x86 or ia64 Linux. Linux users can use most of the software without installation from the AFS location. Windows users, however, need to copy the installation files from AFS to their local computers and install the software on their own. Instructions for installation are available in the AFS directory.

Depending on the tool, a site-wide simultaneous operation of 10 or 15 instances of a software product is possible. All mentioned programs require a local CERN network connection for operation because the licences are stored and managed in IT. openlab does not provide direct technical support for the published products but is willing to pass issues on to Intel. Also note that the products might be upgraded to newer versions in the future.

Useful links
AFS location: /afs/cern.ch/sw/IntelSoftware.
openlab: http://cern.ch/openlab.

Andrzej Nowak, IT-DI (openlab)

Move signals change at the CERN Printshop

The Printshop has moved to 510-R-007, opposite the former office in a much smaller area. This is in response to the decrease in the number of pages printed here, which went down from 17 million in 2004 to 7 million in 2007. The Printshop personnel have been reduced to one operator provided by an outside company.

The old offset and finishing equipment, which for the most part had been at CERN for more than 30 years, has been sold. This was needed in the early 1980s when the Printshop was producing more than 70 million pages per year.

We have done our best to ensure that this change has minimal impact on the user community. In particular we have negotiated a contract with a printing company to take care of jobs that can no longer be handled in house.

More information is available at http://cern.ch/desktop-publishing/printshop.asp.

Michel Goossens, IT-UDS

openlab expands its teaching activity for CERN programmers

With the LHC nearing start-up, the pressure on HEP software developers is without doubt greater than ever before. Unfortunately, changes in the hardware acquired by the CERN Computing Centre and elsewhere force us to bid farewell to the recent golden age of nearly unlimited resource availability. Frequency scaling has ended and multicore systems have instead become prevalent, but the required memory has become costly in terms of money and power consumed. Such changes prompt programmers to take a close look at parallel programming models and at other ways to optimize their code. To facilitate the transition to more challenging hardware environments, openlab is expanding its training offer.

openlab was actively involved in software training in 2007 and this year will see an expanded portfolio of workshops. The tried-and-tested formula of a multithreading workshop cycle every half year is also being applied in 2008, with one workshop in May and another later this year. Experienced instructors from Intel have provided helpful advice and will continue to aid in this effort. In the second half of 2008 a new type of workshop, related to performance tuning and hardware architecture, will be offered to CERN programmers. The formula has already been tested on a select group of software developers and has even been included in this year's CERN School of Computing. As with the multithreading workshop, the class is being planned as a half-year cycle.

openlab was actively involved in software training in 2007, and this year will see an expanded portfolio of workshops offered to CERN programmers on a half-year cycle.

For more information about openlab, visit the webpages at http://cern.ch/openlab.

Andrzej Nowak, IT-DI (openlab)

PC application maintenance follows a lifecycle

Each computer application that is deployed by the IT-IS group needs to be maintained throughout its lifecycle. This maintenance includes following up new versions released by vendors as well as monitoring all security vulnerabilities discovered in an application. Appropriate actions are taken depending on the severity of the issue. These activities have become more demanding over time, so a range of tools is used at CERN that allows us to deploy Windows applications in an easy, flexible and fast way. These tools also mean that we can constantly monitor the CERN Windows environment in order to know which applications are installed on computers managed by the NICE team.

This article describes the processes involved in the management of an application throughout its lifetime at CERN, followed by three case-studies.

The NICE environment

Managing and supporting applications at CERN is a daily challenge. Currently we have around 6000 Windows-based computers divided into 40 sets. Each set defines its own group of applications with the relevant deployment configuration. Standard computers normally have access to 20 core applications (including big applications like Microsoft Office and Adobe Acrobat Reader, but also small ones such as the CERN Phonebook and 7-zip). In addition to these there are about 100 other applications available for installation on demand. Each one has a support level defined. Given that the application set is large, the variety of support levels is also very large. The support level for any application can be composed of a few elements.

- **Installation** – IT/IS (or another group) prepares the installation package, which is then available via the Computer Management Framework (CMF). As soon as an application appears in the add/remove packages list in CMF, any problem with its installation on a standard NICE computer can be reported to the Helpdesk.
- **Usage** – IT-IS (or another group) provides help for the application. To check whether you can get such help through the Helpdesk, see whether the application provides an e-mail address in the "Contact" column on the CMF add/remove packages page.
- **Forced updates** – IT-IS (or another group) follows up all published application vulnerabilities and, as soon as it is considered critical, the update correcting the problem is deployed on all centrally managed NICE computers. Depending on the severity of the issue, we try to bundle such updates together with monthly Windows security updates.
- **Optional updates** – IT-IS (or another group) follows up published application vulnerabilities and prepares all of the necessary update packages in CMF. Updates are then available through the CMF add/remove packages page and can be freely installed.
- **E-mail notifications** – IT-IS follows up all

published application vulnerabilities and sends e-mails to the users of all affected computers, encouraging them to install an update for an application manually.

An example of an application with the highest support level is the Microsoft Office suite, where the installation is prepared centrally. The Helpdesk can be contacted for problems and all of the updates are implemented by the IT-IS group. On the other hand there are applications, such as Apple QuickTime, where no support is provided (there is no CMF package for it and the Helpdesk will not answer any questions). However, in case of security problems with this application, e-mail notifications are sent to all users who have installed it, asking them to upgrade to the most recent version of QuickTime.

Regardless of the level of application support provided by us, there is always one component of the application management that is involved – monitoring. We always keep an eye on statistics showing which applications are installed on NICE computers and we carefully verify all security announcements published by the competent bodies.

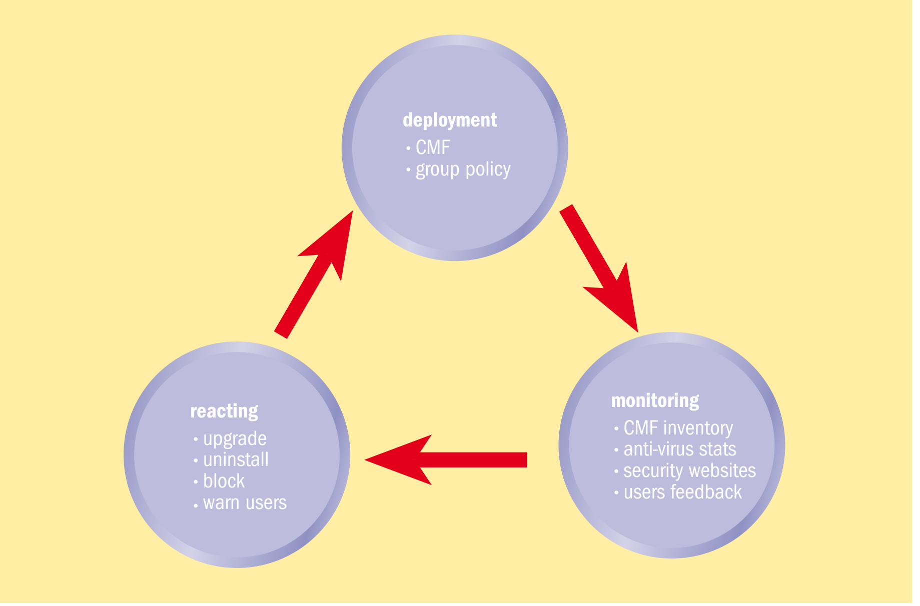


Fig. 1. Throughout the application lifecycle there are three steps repeated in a cycle.

Application lifecycle

Throughout the application lifecycle (from first deployment until phase-out), there are three steps that are repeated in a cycle (figure 1). Everything starts with the deployment of an application and then continues with monitoring (following up vulnerability disclosures as well as collecting the users' feedback through the Helpdesk). Then we have to react according to the severity of issues discovered during the monitoring process. This typically requires the deployment of updates, which closes the application lifecycle and restarts the deployment phase.

Deployment

The deployment of Windows applications at CERN is done using a system called CMF, which is perfectly adapted to the needs of the CERN environment (especially in terms of administrator rights delegation required by several user communities). The main entry point to CMF is the small icon in the Windows system tray (figure 2).

The add/remove packages webpage is accessed by right clicking on the icon. There is a range of applications that can be selected for installation. The support contact details for a given application are also on this page. The CMF User Guide is available at https://cern.ch/cmf/Help/?kbid=001001.

It is important to note that several applications are preinstalled on every NICE computer together with the operating system. These applications are used by the majority of our users on a daily basis and include Microsoft Office, Adobe Acrobat Reader and the CERN phonebook.

Fig. 2. The main entry point of CMF is the small icon in the Windows system tray.

Monitoring

When an application is deployed it becomes part of the pool of monitored applications. From this time onwards we track updates and security vulnerabilities. We consult external websites (e.g. www.sans.org) as well as internal statistics and user feedback. As soon as we are aware of issues, we take appropriate measures.

The main tools used for monitoring are the software inventory and the results of the security scans performed on each computer by CMF. This information is then used to assess the risk of some vulnerability disclosure or simply to verify whether the patch deployment has worked.

Reacting

When we have to react to an event, such as a software update, our goal is to minimize the disturbance to users. We therefore try to group updates together on a monthly basis and deploy them in the smoothest possible way. However, in the case of an emergency, we can react quickly and deploy an update within 30 minutes.

Sometimes we cannot force the installation of an update so we use the e-mail alert system. This retrieves the list of all computers with security issues from the inventory database and sends e-mails to the owners of these computers. These e-mails usually contain simple instructions for how to fix the vulnerabilities. This approach leads to a somewhat slower deployment of updates compared with forced updates (when 100% of computers managed by us receive patches in a short time). Nevertheless, it still proves to be quite effective because most of the users react promptly to these notifications.

Case-studies

The following three cases illustrate aspects of the application lifecycles.

Adobe Acrobat Reader – responding to vulnerability disclosures

Acrobat Reader is widely used to read PDF documents. Consequently it is also frequently targeted by hackers. In early 2007 we were supporting version 7.0.9 of Adobe Acrobat Reader but, following major security breaches, version 8.0 was released. After performing preliminary tests, the progressive deployment of version 8.0 started. During our early phase of deployment, users discovered and reported to the Helpdesk that this new version was causing serious printing problems under certain circumstances. We had to take a decision whether to fix the security issue we were facing and consequently introduce a printing problem, or delay the deployment of the new version until the printing issue was solved.

The decision was taken to stay with the previous version but prepare an update package that would be ready for deployment within 30 minutes on every computer at CERN in case of an emergency. In the next months we constantly monitored security reports looking for announcements of exploits. In addition, we carefully reviewed statistics from our antivirus servers. When it became apparent that the printing issue was still not fixed in the following versions of Acrobat Reader, we decided that we needed to move ahead with the installation of version 8.0. To avoid the observed printing problem, we migrated existing PostScript drivers to use Printer Control Language (PCL) drivers. Once this operation was completed, towards the end of 2007, we were finally able to migrate to Reader version 8.0.

Microsoft Office – product evolution follow-up

In early 2007, Microsoft Office 2003 suite was the default version installed on each new computer, although the previous version (Office XP) was still supported and widely installed at CERN. With the introduction of Office 2007 in November 2006, this brought a lot of new functionalities and we wanted to streamline the situation by keeping only two supported versions at CERN. Therefore we combined the phase-out of Office XP with a smooth introduction of Office 2007 (figure 3).

To do this we initially published Office 2007 as an optional upgrade for more recent computers and we organized new training courses. Then, after some time, Office 2007 became the default Office suite preinstalled on all computers with at least 1 GB of RAM. At the same time the end of support for Office XP was announced and we ensured that all remaining installations were migrated to at least Office 2003.

At the beginning of 2008, Office 2003 (with Service Pack 2) was installed on 80% of NICE computers. The remaining 20% were already using Office 2007. This proportion was a result of the smooth introduction of Office 2007, which we had already planned at the beginning of 2007. In March 2008 the time had come to deploy major updates for both supported Office versions: Service Pack 1 for Office 2007 and Service Pack 3 for Office 2003. To manage this new important deployment successfully we adopted the following procedure:

- service packs are first extensively tested within the IT-IS group;
- they are published via CMF and any user can voluntarily install them;
- after some time we deploy them to all IT computers – the number of old versions installed at CERN drops slightly and the number of new versions increases (marked as period 1 on figure 4);
- normally after 1 month we take the last step and deploy updates to all NICE computers managed centrally by IT/IS; the number of old versions drops dramatically and the number of new versions increases (marked as period 2 on figure 4);
- finally, it takes some time until all CMF local administrators follow and deploy the same critical updates on their computers.

The result of this deployment policy is shown in figure 4.

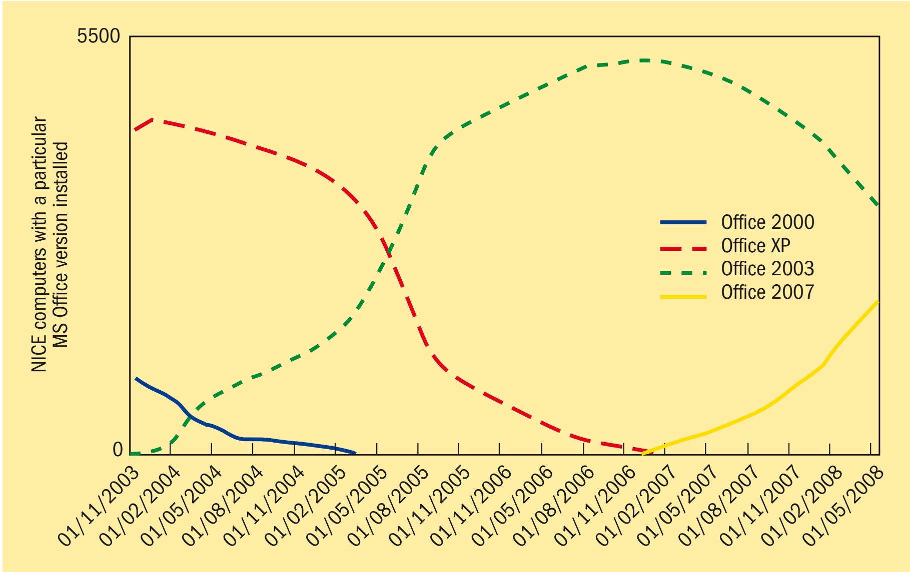


Fig. 3. The evolution of Office suites. Only two versions of Office are supported at a time.

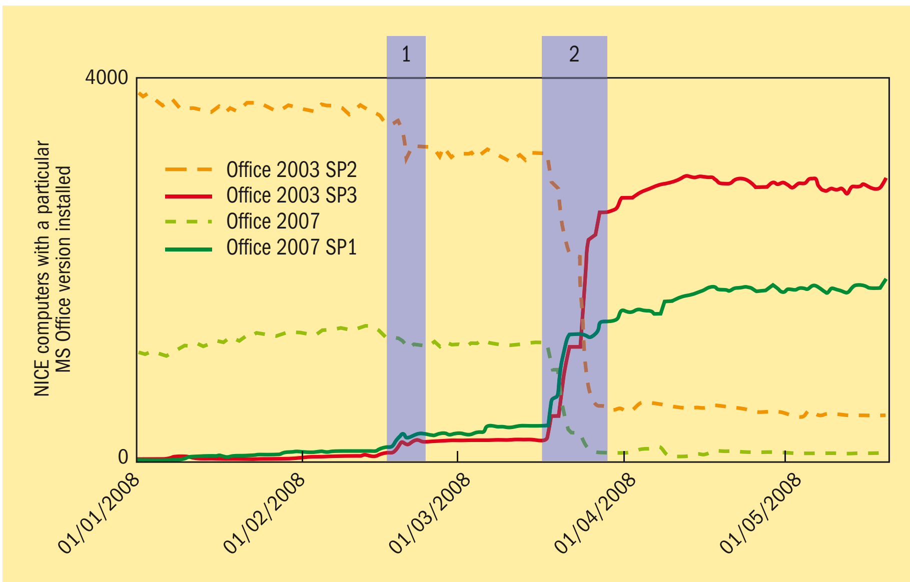


Fig. 4. This shows the progress of the deployment of the Microsoft Office service packs.

Sun Java – managing an unmanaged application

Packaging and installing Java is relatively simple. However, there are many critical Java-based applications used at CERN, such as the accelerator and experiments controls. These applications may require a specific version of Java. Because of this need, Java updates are not centrally managed but there are Java vulnerabilities that are discovered and published almost every month.

Given these constraints we monitor every Java vulnerability disclosure. As soon as updates are released we prepare new packages and publish them via CMF. We then send e-mails to every user who is responsible for the management of an affected computer, giving instructions on how to update manually to the most recent Java version that is needed. This approach is not as efficient as forced deployments but it significantly reduces the risk of a potential attack.

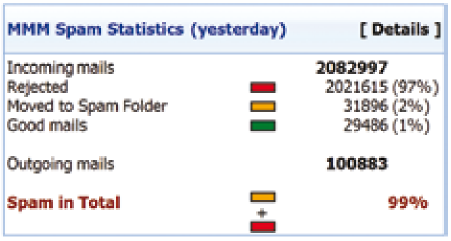
Final thoughts

Software-application management focuses more and more on security rather than application improvement. The number of vulnerability disclosures is increasing and this tendency seems to be here to stay. It is crucial to have efficient and flexible deployment and monitoring tools that allow us to adapt our actions depending on the issue severity. The usability of our computers is as important as security. Therefore the proper balance between fast update deployment and delays that minimize user disturbance is carefully managed on a case-by-case basis.

More details are in the presentation given at the HEPiX Spring 2008 meeting from http://indico.cern.ch/getFile.py/access?contribId=47&sessionId=8&resId=0&materialId=slides&confId=27391.

Sebastien Dellabella, Rafal Otto IT-IS

CERN improves spam filters



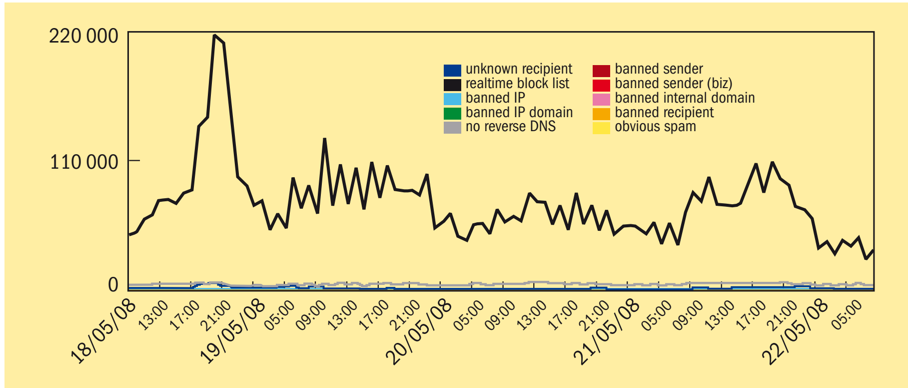


Fig. 1. (above). An example of the daily report on spam at CERN, which is available on the mail services. Fig. 2. (right). Statistics showing the various causes of spam rejection from CERN mailboxes.

In April a visible change in spam filtering was introduced at CERN. This change, due to a new anti-spam system, raised some questions from CERN mail users. As a result we would like to provide more information about spam and how it is filtered at CERN.

Spam is designated as unwanted electronic mail sent in large numbers to as many recipients as possible. It is a nuisance to every e-mail user. The volumes of spam have increased exponentially over the years and have recently stabilized at an enormous level of 100 billion messages every day worldwide (about 90% of all mail messages), generating significant costs for all organizations running a mail system. Spam exists because some people respond to it, generating income for the spammers who send it.

As spammers need lists of e-mail addresses to target, the more publicly visible an address is, and the longer it has existed, the more likely it is to receive spam. For this reason, CERN receives much more spam than average: in 2008 we are receiving between 2 and 3 million mails daily, consisting of 98–99% spam. A report about daily amounts of spam is visible on the CERN mail services page at http://cern.ch/mail (figure 1).

Techniques to contain spam

The process and techniques used to filter spam are basically the same as the previous system and are representative of how most anti-spam systems work. However, we have improved each phase of the process as much as possible to provide perceptibly better filtering.

Detecting spam is difficult. It is a constant race between spammers and anti-spam systems, as spammers constantly try to work around anti-spam measures. The protocols used to exchange e-mail provide little help in identifying spam from legitimate mails. As a result, a mail system has to evaluate the probability of a message being spam in order to decide whether to filter the message. Spam fighting is all about statistics, and consequently it is not an error-proof process – some legitimate mails can be incorrectly considered as spam (false positives) and some spam messages can be undetected (false negatives). The challenge is to keep the false positive and false negative rates as low as possible.

Recent extensions to the electronic mail protocols (Sender Policy Framework, SenderID) provide some mechanisms to prove that a mail is legitimate. CERN mail servers use several of them to prevent outgoing mails from being detected as spam, and to prevent the filtering of incoming mails that contain a proof of their origin. However, too few organizations are implementing these systems so far, resulting in more filtering errors.

The majority of automatic anti-spam systems consist of two levels of filtering: a first level of aggressive action is applied to messages that are identified as spam with a high confidence level; then messages that are likely to be spam but not with as high a confidence are delivered into a dedicated quarantine, or spam folder, so that they can be verified by the user.

First level: reject spam

Before it is delivered, an e-mail message is relayed through various mail servers. At CERN the first level of filtering is essentially implemented by performing various checks on the internet mail server that is trying to send the message to a CERN e-mail address.

The most important check consists of looking up the mail server in several public databases of mail servers known to be spammers to filter them out. This catches 80–90% of incoming spam but is not sufficient because only well known spammers appear in these databases.

Additional checks, like reverse DNS look-ups, are performed to verify the compliance of the sender with the best practices for legitimate mail servers.

Finally, the messages are compared to spam-definition files (similar to the virus-definition files used by anti-virus systems) to see if they contain known spam content or banned words, such as “Viagra” or “porn”, or come from explicitly banned domains and mail addresses. Messages that match well known spam content are blocked, and mail servers sending significant amounts of such messages are automatically completely blocked for a few hours to prevent the less obvious spam sent by the same mail server from being delivered to mailboxes (figure 2).

This first level of filtering now has a very low rate of false positives – in the order of one per million messages. Most important, because it is based on checks against well known public databases of spam content and spammer IP addresses, senders failing these tests are not just unable to send to CERN, but they are typically unable to send to many other mail organizations using the same widely used IP/spam databases. In other words, they are very likely to have difficulties sending this mail altogether.

Given their low error rate, it does not make sense to have these messages delivered to the spam folder: each mailbox would receive hundreds of them each day – too many to make manual checking possible. So a message that fails these checks is never delivered to CERN mailboxes. Instead, CERN mail servers will reject the message with an error message. This approach uses the same error mechanism that allows the reporting of typical mail-related errors, such as a mistyped or invalid mail address. As a result, in case a legitimate message is incorrectly rejected by this first level of filtering, the sender receives an error message (a non-delivery report) with details of why the message was refused. Senders would typically have to contact their local mail administrators to have the problem addressed, such as by removing themselves from public databases of spammer IP addresses.

This first level of filtering is the area in the new anti-spam system that has been improved the most; this is why the amount of spam delivered in the spam folder has been significantly reduced. It was particularly important to improve this step because many CERN users have their e-mail address configured to forward to an external mailbox. Any spam not filtered by the first level would be forwarded to this external mailbox. If CERN forwards more spam than it forwards legitimate messages, then our mail servers would themselves be considered as spammers.

Second level: move spam to folder

Of the initial 2 to 3 million mails, about 100 000 pass the first level of filtering (three times less than with the previous system), but half of them are still spam. The probability of these messages being spam is evaluated by looking at characteristics of the message contents and headers.

Each message is assigned a spam note representative of the probability of the message being spam. A filtering level can be configured by every user in the spam-configuration pages on the mail services website (figure 3). The filtering level defines a threshold so that messages that have a spam probability higher than this are moved to the spam folder. This keeps the amount of spam delivered to the inbox folder to a minimum. However, because of the heuristic approach, the error rate is typically higher than in the first level of filtering. A few percent of these messages can be spam incorrectly delivered to the inbox or legitimate messages moved to the spam folder. Consequently, if the configured filtering level is high, it is important for users to check the spam folder from time to time for false positives. On the other hand, a low filtering level is unlikely to move any legitimate message to the spam folder and more spam have a chance of being delivered to the inbox unfiltered. The medium level provides a good balance.

Fig. 3. The configuration of the anti-spam settings for a typical CERN mailbox.

Finally, a list of trusted e-mail addresses can be set up by each user to prevent messages that are sent by them from being moved to the spam folder. At the moment this “white list” only applies to messages that have passed the first level of filtering, but in the near future it will also be applied during the first level of filtering to reduce the error rate further.

The future of spam fighting

During this summer, further updates will be deployed to improve spam fighting. This includes applying user-defined white lists to the first level of filtering and significantly better integration in webmail and Outlook, with the possibility of easily white- or blacklisting an e-mail address by right-clicking a message, and optionally having all of your contacts and addresses added automatically to the white list. The spam-detection filters will continually be improved and users will not be asked to report spam anymore.

Practical information

When a legitimate message is incorrectly delivered to the spam folder, consider adding the sender’s address to your white list so that further messages from this person will not be moved to the spam folder. If this happens too often, setting the spam filtering to a less aggressive level (low or medium) should solve the issue. Because of the improved efficiency of the spam detection, medium and low settings can now be used safely without delivering significant amounts of spam to the inbox.

If you are aware of a specific message that was sent to you but was never delivered to your inbox, it might have been rejected by the first level of filtering. In this case, the sender receives a detailed error message and should typically deal with their local mail support to solve the problem. The sending organization should be contacted in situations when the mail is not sent by a human (automatic mails). We can also use this information to adjust the spam filter settings if you can provide the sender address (or at least domain) and approximate date and time of the delivery failure to the CERN Computing Helpdesk.

We would like to thank our users for having reported filtering errors in the first days following the introduction of the new anti-spam system. This allowed us to fine-tune the system and improve the settings. Your feedback is valuable for the continuous improvement of the system.

Useful links

Mail support line: mail.support@cern.ch
CERN Mail Services homepage: http://cern.ch/mail (follow the link “spam fight” to configure your spam filtering settings).
Information about spam at CERN: http://cern.ch/mail/antispam.

Alex Lossent, CERN Mail Services

The deadline for submissions to the next issue of CNL is

22 August

Please e-mail your contributions to cnl.editor@cern.ch

If you would like to be informed by e-mail when a new issue of CNL is available, subscribe to the mailing list cern-cnl-info. You can do this from the CERN CNL website at http://cern.ch/cnl.

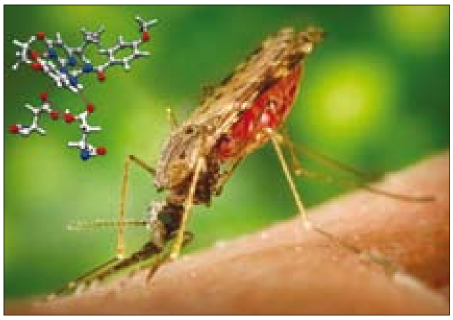
WISDOM unplugged: malaria drug leads graduate to wet lab

Hot on the heels of their initial success, WISDOM collaborators have high hopes as their anti-malaria drug-leads move from computer processors to petri dishes.

"We're excited to be at this stage," said Ana Lucia Da Costa, WISDOM researcher at CNRS in France. "I can't wait until we get the next set of results back."

The WISDOM project, a collaboration of eight core institutions in five countries, began searching for anti-malaria drugs in 2005. Rather than go straight to the lab, the team used a Grid-powered software program to screen for potential drug-leads, searching for small molecules – called ligands – that could bind to and disable the malaria protein plasmepsin. "Disabling this protein essentially starves the parasites to death," said Da Costa.

Plasmepsin, Da Costa explains, is malaria's workhorse. Used by the parasite to attack red blood cells in the human body, it can be disarmed using the perfect weapon: a matching ligand – the lock for malaria's destructive key.



The WISDOM project screened 1 million molecules for their potential to bind to and inhibit plasmepsin, essential for the survival of malaria (left). The parasite can be transferred to a human host via the mosquito. Image courtesy of PD-USGOV. Every year malaria affects more than 500 million of the world's poorest people and kills around 3 million (right). A child dies from malaria every 30 s. An African child has on average between 1.6 and 5.4 episodes of malaria fever each year. Image courtesy of hdptcar.

One in a million

The search for this perfectly matched ligand led the WISDOM team to perform 41 million "dockings", screening a million molecules and then discarding all but the 30 most promising. This select handful then moved off-screen – undergoing *in vitro* evaluation in a wet lab at Chonnam National University in South Korea. In a fantastic result, all 30 computer-selected ligands were able to inhibit plasmepsin in the lab, even at nanomolar concentrations.

"This was a great success," said Da Costa, who helped to analyze the results. "We didn't expect all of the ligands to be active. It shows that our approach worked really well."

Now these 30 ligands have advanced to the next stage – *in vivo* testing in cultures of the malaria parasite at the Commissariat Energie Atomique in France. The testing is ongoing, but initial results show that at least one of the ligands inhibits the parasite lifecycle at micromolar concentrations.

The next step will be to test for toxicity to animal cells. If the ligands are safe, the collaboration can begin to pursue approval for a new drug.

Classical pharmaceutical research often takes 15 years to develop and approve a new drug. Thanks to Grid power, the WISDOM team has fast-forwarded this scenario and is already in the lab after just three years of research. It is using the same approach to expand into the areas of diabetes and avian flu research.

WISDOM uses several Grid infrastructures, including Enabling Grids for E-sciencE (EGEE).

Useful links

WISDOM: http://wisdom.eu-egee.fr/
EGEE: http://www.eu-egee.org/

Danielle Venton, EGEE

● This article was published online in iSGTW on 7 May.

Project gives integrated site security for Grids

Almost 99% of the e-mails that CERN receives each day are spam. While some are nothing but a harmless nuisance, others are malicious and capable of causing a substantial amount of damage. This is just one of the ways in which a site like CERN's can be attacked. To increase awareness and provide security guidance to Grid sites, CERN has led an EC co-funded project entitled Integrated Site Security for Grids (ISSeG), which was completed at the end of March. The final results of this 26-month project are available from the ISSeG website at www.isseg.eu.

This site provides a risk-assessment tool, security recommendations and training material to help sites to improve computer security. The project's focus has been on the security of Grid sites, but the material is applicable to a range of computer centres, particularly those in an academic or technical environment.

Integrated site security

The project's vision has been that Grid security, which focuses on middleware, authentication, authorization and operation across multiple administrative domains, needs to be complemented by comprehensive site security at all of the participating Grid sites. The ISSeG project has created and disseminated practical expertise on the deployment of integrated site security (ISS).

ISS is a practical approach to site security that integrates technical, administrative and educational security solutions, and develops them in a coordinated way. This integration ensures that policies, rules, awareness and training evolve in step with technological and administrative developments.

Creating practical expertise

The project began in February 2006 and has been co-funded by the EU FP6 programme. The consortium comprised the European Organization for Particle Physics (CERN) in Switzerland, Forschungszentrum Karlsruhe (FZK) in Germany and the Science and Technology Facilities Council (STFC) – formerly known as the Council for the Central Laboratory of the Research Councils (CCLRC) – in the UK.

ISSeG created and captured raw expertise through full-scale ISS deployment at CERN and FZK, including flexible and improved security for centrally managed computers, strengthened policies for control networks and increased firewall protection. Experience gained from the two site deployments, as well as site security assessments carried out by a subcontracted company, were used to develop training materials and recommendations as to how to mitigate security risks.

Coordinated by STFC, the results were disseminated via presentations and the ISSeG website to help scientific communities to use this integrated approach to improve their site security measures. In addition, there has been close collaboration with the security groups of the Enabling Grids for E-sciencE (EGEE) project – the Operational Security Coordination Team (OSCT) and the Joint Security Policy Group – which has resulted in their continuous involvement and input to help to shape the further development of ISSeG.

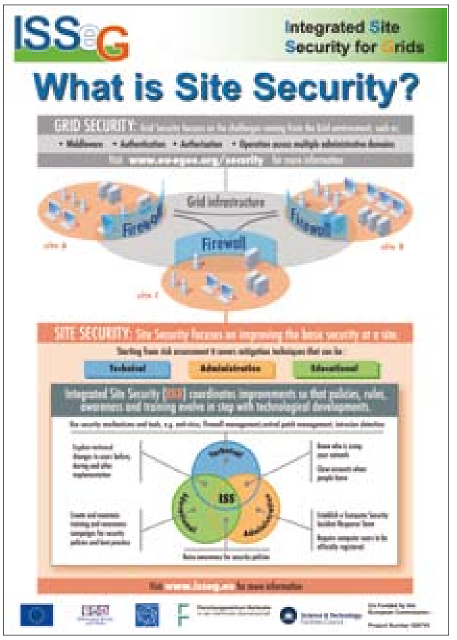
The ISSeG website

Visitors to the website (www.isseg.eu) can download and complete a risk-assessment questionnaire and hence, via a prioritized list of threats specific to their site, receive tailored site-security recommendations. A generic set of the top threats and recommendations for Grid sites can also be viewed directly by following the links.

Training materials for general users, system administrators, software developers and managers are also available from the website, including introductory material, training presentations, security checklists and downloadable printable materials and posters in PDF format.

Acknowledging the usefulness of the content, the OSCT will take over the maintenance of the ISSeG website to ensure its continued availability beyond the lifetime of the ISSeG project. This agreement was formalized in a memorandum of understanding between ISSeG and phase three of the EGEE project, which began on 1 May.

Kate Bradshaw, ISSeG project



This "What is Site Security?" poster was shown by ISSeG at EGEE'07. It is available to download from the ISSeG website.

BalticGrid-II increases the impact, adoption and reach of Grid technologies in Baltic states

The launch of the second phase of the BalticGrid project – BalticGrid-II – was held in Vilnius on 13–15 May. The BalticGrid first phase extended the reach of Enabling Grids for E-sciencE (EGEE; www.eu-egee.org) technologies by creating a gLite-based e-infrastructure, which included the Baltic states. This has brought the knowledge and use of Grids in the region to a level comparable to other EU member states.

BalticGrid-II is designed to increase the impact, adoption and reach, and to improve further the support of services and users of the recently created e-infrastructure. This will be achieved by extending the gLite-based e-infrastructure to Belarus; its interoperation with UNICORE (www.unicore.eu/) and ARC-based (www.norduGrid.org/middleware/) Grid resources in the region; and identifying and addressing the needs of new scientific communities.

Currently the e-infrastructure has 29 sites comprising 1994 CPUs and 217 TB of storage, and this is envisaged to grow in capacity and capability in the next two years. The project consortium is composed of 13 institutions in seven countries: 7 from the Baltic states (Estonia 2, Latvia 2 and Lithuania 3), 2 from Belarus, 2 from Poland, 1 from Sweden and CERN. This two-year project, which started on 1 May, is co-funded by the EC (€2.9 million) under the 7th Framework Programme (http://cordis.europa.eu/fp7/home_en.html).

BalticGrid-II is organized into eight activities: management, education and outreach, application identification, policy and standards, Grid operations, network provisioning, application support, and development of enhanced application services. In total the project will deliver 450 person-months, or about 20 full-time two-year posts.

The map shows the BalticGrid-II partners: CERN, Poznan, Krakow, Minsk, Vilnius, Riga, Tartu, Tallinn and Stockholm.

KTH, Sweden, as coordinator, will manage and run the project. As in the first phase, CERN will play a strategic role in providing support to the expansion and the delivery of a production-level e-infrastructure with production-quality management and operation. CERN will also contribute essential know-how in gLite deployment, operations and support. Specifically, CERN will support the expansion of the e-infrastructure to Belarus, with the goal of creating at least one gLite-based resource centre in the country to be administered by local partners. In addition, CERN will continue with its liaison role to promote further the collaboration between BalticGrid-II and EGEE, and with other similar projects.

BalticGrid applications span several scientific domains: HEP (CMS, LHCb), life sciences and material sciences. In BalticGrid-II this will be expanded to include new areas, including thermonuclear fusion, engineering modelling tools and linguistics. HEP remains one of the main applications, with project partner KBFI hosting the Estonian Tier2.

In BalticGrid, as with similar regional e-infrastructure projects, the CERN personnel hired for the project are from the region and CERN is looking forward to welcoming a project associate from Lithuania in June.

For more information about the BalticGrid-II project, visit the website at www.balticGrid.eu/.

Frédéric Hemmer, Maria Barroso Lopez and Florida Estrella, IT department

Accelerator databases migrate to architecture based on Oracle RAC and NAS storage systems

During March, three of the most critical accelerator databases were migrated from old Sun servers running Solaris with local attached storage area networks (SANs) to the new database infrastructure designed by the IT-DES database group. This article explains what this new database infrastructure is and it presents some results and experience from the migration.

Brief introduction to the accelerator databases
There are three main databases:

- **ACCLOG (logging service)** The aim of this database is to have a clear understanding of what is going on in and around the LHC, by logging heterogeneous time-series of data to achieve the following objectives: manage information required to improve the performance of the machine and individual systems; comply with traceability safety regulations to record beam history; make long-term operational statistics available for management; avoid duplicate logging developments.

The service was originally intended for the LHC but the scope has widened to the whole accelerator complex. The current (April) phase of hardware commissioning is extremely demanding and important because this is how equipment experts determine the behaviour of their systems.

This database is accessed using different interfaces: Java API, PL/SQL API and a web-based graphical interface that includes time-scaling functionality.

The ACCLOG database should run for the entire lifetime of the LHC with an expected growth of 10 TB per year.

- **ACCMEAS (measurements service)** This has similar functionality to the ACCLOG database but is faster and has a shorter lifetime. It acts as a buffer from which the data can be viewed, automatically filtered and the useful long-term data transferred to the logging database.
- **ACCCON (settings and controls configuration)** This is the most critical database holding all data to pilot the accelerators. "Controls configuration" makes the addressing of remote equipment possible and "settings configuration" gives instructions to the particle beams. The settings configuration part is particularly crucial because the settings are not static and constantly need to be tuned (like an aeroplane but without an automatic pilot). A potential absence of these settings (i.e. a database missing) could lead to the LHC beam being dumped by the operator.

New infrastructure
For the last 12 months the IT-DES database group has been deploying a new infrastructure to host databases. This offers high availability by means of technologies at both the hardware level via network attached storage (NAS) and the logical level via Oracle's Real Availability Cluster (RAC).

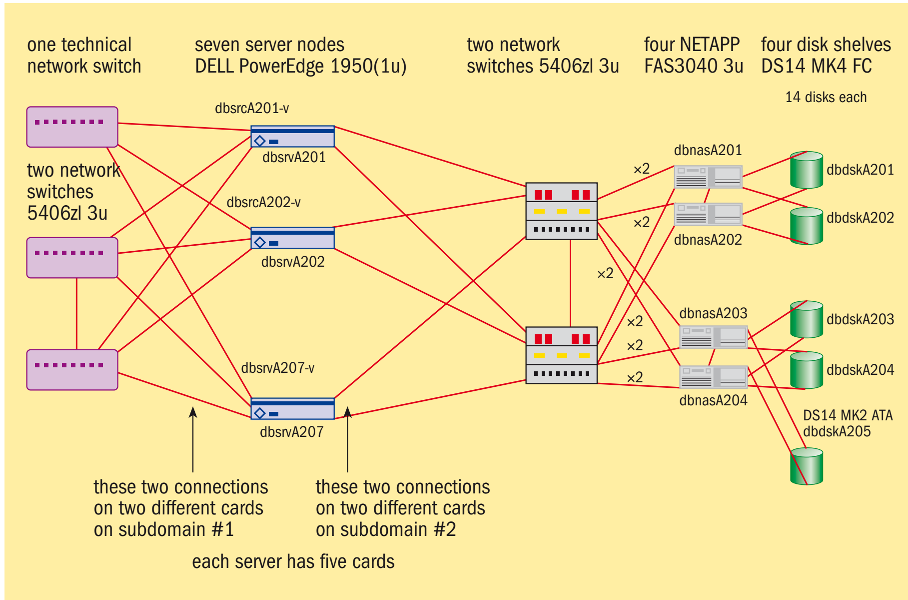


Fig. 1. Accelerator database design. The infrastructure was designed by the IT-DES group.

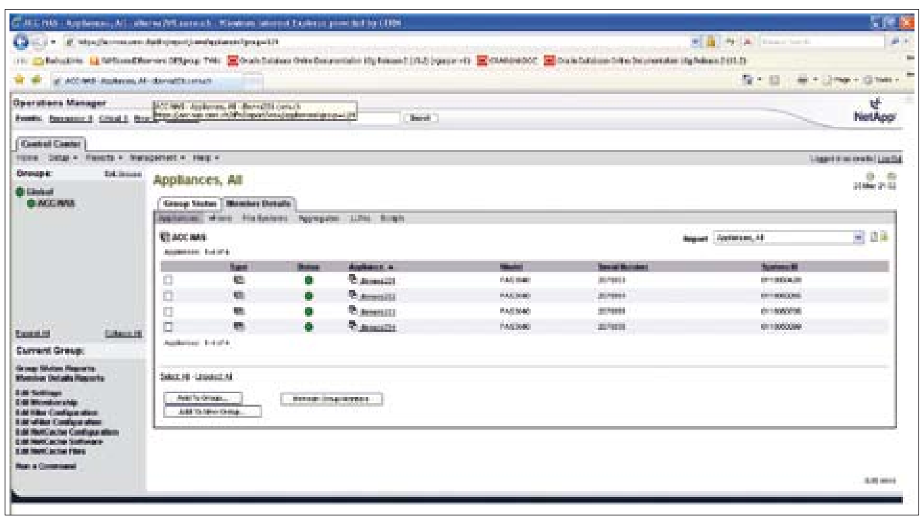


Fig. 2. Monitoring and administration can be carried out via the operation manager.

NAS storage differs from SAN storage in two main ways: the protocols used and the way in which they handle information. Any system connected via an Ethernet network can access a NAS storage, typically using TCP/IP-based protocols like network file system (NFS) for Unix-like clients or common internet file system (CIFS) for Windows-like clients. Both NFS and CIFS are protocols for accessing the data like a local file system. Information is represented as in a file system, so operations like granting privileges or doing operations on files are done using the operating system rules. On the other hand, a SAN system uses fibre channels and the information is accessed at the block level.

Today NAS and SAN technologies are largely unified and the key differences for the IT-DES systems are simplicity, ease of management and built-in monitoring.

Oracle RAC allows multiple computers (typically two or four nodes with multicore CPUs each on DES clusters) to run the Oracle Relational Database Management System (RDBMS) software simultaneously while accessing a single database, thus providing a clustered database. Oracle RAC uses a technology called cache fusion to keep coherency among data. A RAC system improves high availability as it makes the database resistant to server failures. A range of interventions can be carried out without any downtime. Scalability is improved because new resources can be added to the system.

Figure 1 shows that all systems have redundancy except for the technical network switch. From the server point of view there is a bonding link on both sides via two Ethernet links working in an active/passive configuration. On the left the bonding link serves as an Oracle RAC interconnect and on the right it establishes connection with the NAS switches. On the NAS side the Data Ontap operating system provides trunking under the name of a virtual interface (vif). On the left of the NAS filers, a first level is made up of two vifs with two Ethernet links each that work in a load-balancing fashion. They are joined by a second-level vif that works on an active/passive mode between both first-level vifs. On the right of the NAS filers the connection to the disk shelves is via fibre channel. Reaching any of the disks can be accomplished via any of the NAS filers in the cluster, thanks to multipathing.

More about the NAS hardware

The storage is done by Netapp appliances. As we see in figure 1, a NAS cluster consists of two filers working in an active/active mode. This means that, in case of failure, the surviving filer will undertake the role of its partner and start to serve the partner's data. The operating system running on the filers is Data Ontap version 7.2.4.

The NAS filer comes with a lot of out-of-the-box functionalities to deal with data integrity and resilience to failure. Four of its main features are as follows:

- **Continuous media scrubbing** The purpose of this is to detect and scrub media errors. This minimizes the storage-system disruption due to media errors while a system is in a degraded or reconstruction mode. Performance impact is negligible.
- **Predictive disk failure and rapid redundant arrays of inexpensive disks (RAID) recovery** Data Ontap monitors disk performance so that when certain conditions occur (i.e. more than 100 media errors in a period of one week) it can remove a disk from its RAID group for testing and, if necessary, fail the disk. Initially, Data Ontap will mark the culprit disk as pre-failed and start to copy the contents onto a spare disk. Once this is finished it will mark the disk as failed. Doing this avoids the rebuild time, performance degradation and possible data loss due to an additional disk failure during RAID reconstruction.
- **Double parity RAID** This is special high-performance implementation of RAID6 that allows up to two failing disks.
- **Direct connection to Netapp support, auto-support feature** This speeds up the supply chain because a disk is shipped as soon as it has been detected as failed.

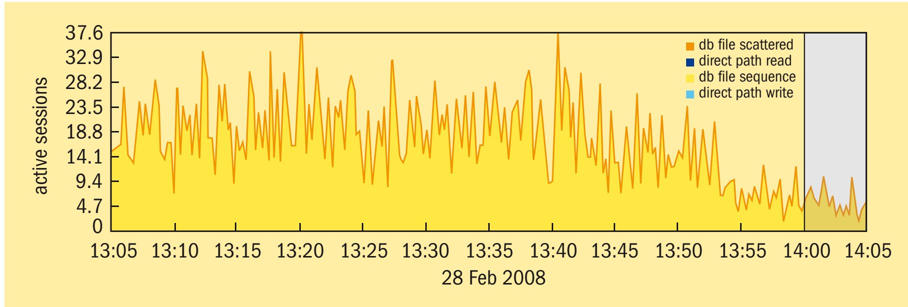


Fig. 3. The active sessions that are waiting to use the I/O subsystem on SUNLHCLOG.

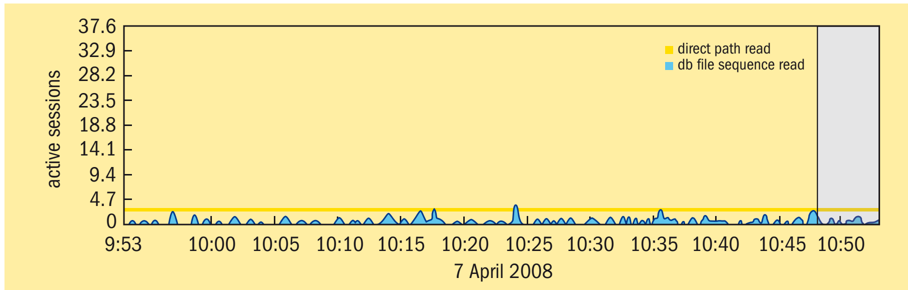


Fig. 4. The active sessions that are waiting to use the I/O subsystem on the new hardware.

Moreover, built-in snapshot technology allows versioning of data if done in coordination with the Oracle database, thereby ensuring that the copies are consistent. For databases, data consistency is only achieved if the database is in back-up mode while a snapshot is taken.

Data are grouped in volumes by means of Netapp FlexVolume technology that allows flexibility in the way that the storage system assigns resources to the volume. Extra features like autosize or reducing the size of the volume on the fly make this technology very interesting. Volumes get their storage from one or more RAID groups, gathered together by an aggregate. The aggregate is the biggest unit of storage. It holds metadata about the volumes it contains. In our case the RAID groups consist of 13 disks plus one spare. We work with double-parity RAIDs, so only 11 disks will be used for data and two for parity information.

Monitoring and administration can be done via the console or via a web interface engine called the operation manager (figure 2). This provides reporting, programming API against the filers, configuration files to deploy settings, role-based access to the interface and the possibility of creating rules to filter events to be sent to the managers.

Access to data is provided via NFS with Oracle-optimized mount options. This allows a lot of flexibility because the only requirement to access the data is to have a machine with filer access allowed. This also gives the possibility of mounting volumes belonging to different filer clusters.

Oracle RAC

The migration also implied an upgrade of the Oracle version used from Oracle 9.2.0.8 32 bits and Sun cluster ware to Oracle RAC 10g version 10.2.0.3 64 bits.

A RAC system of two instances is provided for each database. Installation was done using Red Hat package manager developed by the DES group, as well as Quattor for the configuration of the Oracle cluster-ready services (CRS). Access to the databases is done by services that the CRS system guarantees to balance among the instances. For example (from tnsnames.ora):

```
dbabco=(DESCRIPTION=
(ADDRESS=(PROTOCOL=TCP)
(HOST=dbsrva202-v) (PORT=1574) )
(ADDRESS=(PROTOCOL=TCP)
(HOST=dbsrva203-v) (PORT=1574) )
(LOAD_BALANCE=off)
(CONNECT_DATA=(SERVER=DEDICATED)
(SERVICE_NAME=acccon_s.cern.ch)
(FAILOVER_MODE=(TYPE=SELECT)
(METHOD=BASIC))))
```

In this case we indicate to the listener that we do not want load balancing. This means that new requests will go to the first node and, if there is an error, to the second. Clients aware of fast application notification (FAN) will also benefit from the virtual IP service because it knows that a reconnection is needed without waiting for TCP/IP timeouts. Failover method "SELECT" means that a client will still continue to retrieve the results of a select statement if the first connection fails.

Notes about migration and performance results

The benefits for each database after several weeks of running on the new system have been analyzed by the AB department database administrators, Ronny Billen and

Chris Roderick.

- **ACCLOG (logging service)** The old hardware had difficulties keeping the input/output (I/O) load, which was a cause of contention (figure 3). As a result of the time spent waiting for the I/O subsystem on the old server, the CPU usage was at 100%. The new server is currently running with around 25% CPU usage, but this can be further optimized (figure 4).

The downtime of the logging service during the switchover between the old and new servers caused a significant backlog of logging data – cached by the PVSS systems. Once the logging service was reopened, we saw that the application server used to write logging data to the database was overloaded and had become the new bottleneck in the system.

A second application server was configured and used to ease the situation. Once the logging service returned to a stable situation, the audit data showed that data can be written to the new database at similar rates (per client) to the best performance seen on the old database server a few years ago (when the total volume of data was at least an order of magnitude less; figure 5).

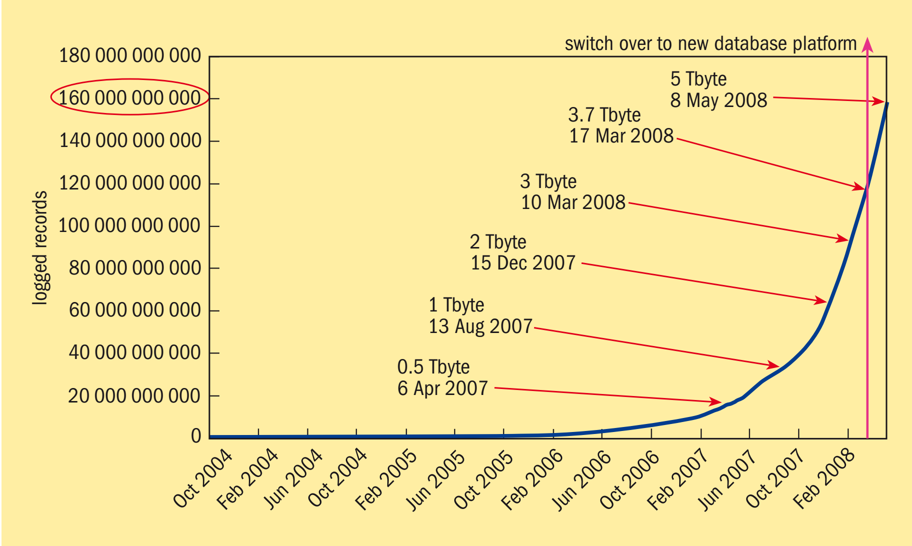


Fig. 5. Database size evolution. The data volume was significantly less a few years ago.

- **ACCMEAS (measurements service)** The data throughput to the new dedicated server for measurements has significantly increased. On the old server we had observed throughputs per client of between 30 000 and 20 000 records per second. On the new server we have already witnessed throughput at close to 70 000 records per second, but surely with the ability to maintain this rate across a greater number of clients.

A critical job that runs periodically, analyzes measured data and sends filtered data of interest to the logging database, which now runs significantly faster for similar data volumes: old server (optimal conditions) = 3 min; new server = 22 s. The new system runs eight times as fast.

The migration to the new service went smoothly. The new service was reopened after approximately three hours.

- **ACCCON (settings and controls configuration)** The performance of the applications that interact with this database suffered badly from the demands placed on the old LSA database server (SUNLHCLOG) by the logging and measurement services. Since moving to a dedicated database server, the users are extremely happy with the performance and the rapid I/O times. One example of the improvements observed is a complex query used to return accelerator equipment settings, which was taking around 1 min and is now taking less than 5 s to return some 200 000 records – the new system is 12 times as fast.

All controls-related databases were also migrated to the new ACCCON RAC server. The migration went relatively smoothly, with the exception of one very large settings-related table, which suffered from an Oracle-related bug affecting index organized tables and exp/imp utility. The new service was fully reopened after approximately 16 hours.

Further tuning is being done by the IT-DES group with the co-operation of AB-CO database developer's team.

Conclusion

The migration of the accelerators databases to a new infrastructure was successfully accomplished. The new hardware has more than a year of proven performance enhancements on other services, such as ITCORE and CASTOR databases, and it is based on high availability at the hardware and the logical level. Hardware high availability is achieved by the duplication of switches, NAS appliances, network bonding and logical high-availability through Oracle RAC, which makes the system easy to scale. It is foreseen to add an extra layer of protection for the very critical databases with Oracle data guard.

After several weeks of running on the new system, the improvement in performance is quite noticeable and has had a positive impact on the user community. Further work is now being carried out to tune, manage and administer these databases.

Useful links

http://indico.cern.ch/contributionDisplay.py?contribId=110&sessionId=28&confId=3580.
www.oracle.com/technology/products/database/clustering/index.html.
http://quattor.web.cern.ch/quattor.
http://storagemojo.com/2007/09/19/cerns-data-corruption-research/.
www.linux-foundation.org/en/Net:Bonding.
www.netapp.com.

Ruben Domingo Gaspar Aparicio, IT-DES

If you want to be informed by e-mail when a new CNL is available, subscribe to the mailing list cern-cnl-info. You can do this from the CERN CNL website at http://cern.ch/cnl.

3D viewer offers another dimension for PVSS

If you have ever watched a science-fiction film where the cockpit of a spacecraft is shown, you will have spotted an animated 3D display showing an overview of the complete apparatus, with flashing red highlighting the damaged parts.

And what about the LHC and the experiments at CERN? The operators in the control rooms sit in front of screens displaying mysterious, yet standard-looking, operation panels (many of these run on top of the common application platform, PVSS). With the new 3D viewer developed by IT-CO, such 3D displays will become available.

Let's have a look at an example of the 3D viewer in action. In this series of images (figure 1), the state of a fictitious subdetector is shown. The colours of the working elements signify the temperature: green being OK, yellow signalling that attention is needed and red signalling danger of damage.

The idea of a 3D overview display showing the state of the detector first emerged in the CMS experiment. It was implemented by Robert Gomez-Reino and immediately attracted the attention of many users. However, it had an important limitation – being a complex blend of Java and ActiveX technologies, it could only be used on the Windows platform. This limitation turned out to be a show-stopper preventing the adaptation of the CMS display for use in other experiments.

Luckily, in recent versions of PVSS the Qt library was harnessed to power the user interface. The Qt application framework, known for its excellent cross-platform portability, is a workhorse for projects such as the KDE desktop environment (a leading, open-source desktop environment for Unix) Lucasfilm, Skype and Google Earth. Its open architecture allows us to extend the set of "widget'' types (elements used to create user-interface applications), using the same code-base on all platforms. PVSS makes full use of this feature, allowing the PVSS platform to be extended with custom "external widget object" extensions, written with Qt. This has been done for the 3D viewer presented in this article.

The rendering and animation of the 3D scene in the viewer is powered by the Coin3D library – an open-source implementation of the OpenInventor high-level programming interface for 3D graphics. The mechanism to embed the Coin3D-based graphics scene in a Qt application is also readily available from the Coin3D webpages, which made the initial integration of these software components straightforward. However, the set of shape types, offered out-of-the-box, quickly appeared insufficient – shapes representing a segment of a cone, polyhedra or trapezoids had to be added. Again, we were able to identify and reuse the code of an existing open-source project – HEPVis – that fitted our needs exactly.

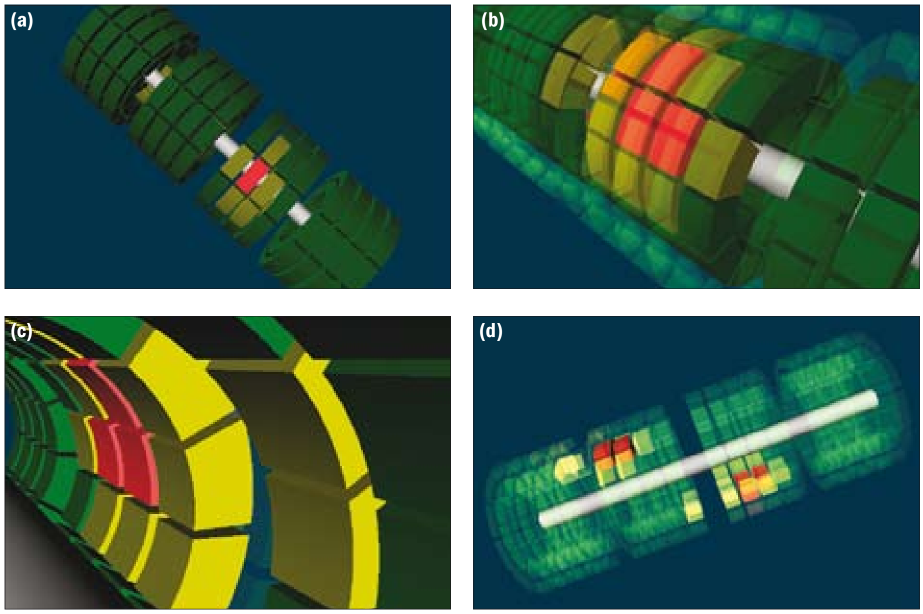


Fig. 1. (a) The operator spotted an anomaly in one of the cylinders. "Inner" elements are not visible. (b) Dimming the external layer reveals a problem in the inner cylinders. (c) "Flying" through the detectors allows inspection of the problem details. (d) Dimming all of the parts except the ones indicating problems allows for spatial correlation.

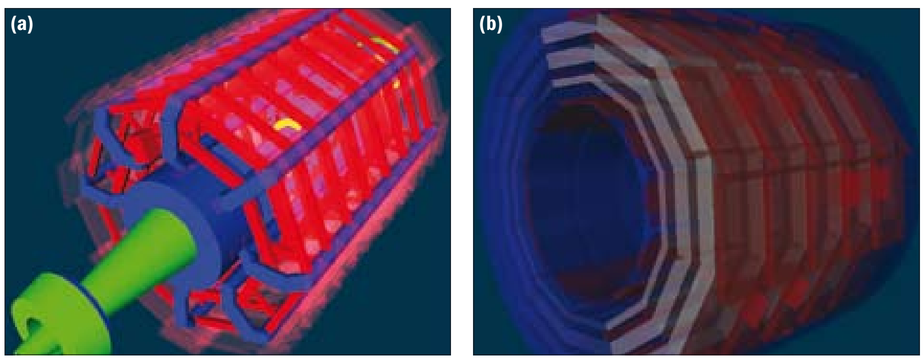


Fig. 2. Visualizations that make use of the data from the detectors' geometry databases. (a) Image from the ATLAS detector. (b) Image from the CMS muon subdetector.

The set of technologies being integrated does not, however, end here. To display the views of detectors, such as the ones in figure 2, the data stored in pre-existing geometry databases can be used. Unexpectedly, the data used mainly by the offline data-analysis programs find an exciting application in the detector online systems. All that is left to do is to establish the connection between the geometrical representation of the detector parts and actual parts of the existing control systems. Many subdetectors will soon have a 3D view of their hardware ready to be embedded into their operational panels.

The 3D viewer is not limited to visualizations of the detector. It is a generic, fully programmable UI widget, the power of which could be unleashed using PVSS scripting features, such as 3D histograms, charts and trend plots.

Useful links

3D viewer homepage (JCOP framework downloads page): http://itcobe.web.cern.ch/itcobe/Projects/Framework/Download/Components/3DViewer/welcome.html.
PVSS (process control and visualization): www.pvss.com.
Qt cross-platform application framework: http://trolltech.com/products/qt.
Coin3D high-level 3D graphics toolkit: www.coin3d.org/.
HEPVis class library extension to the OpenInventor toolkit; maintained in the OpenScientist code base: http://openscientist.lal.in2p3.fr/.

Piotr Golonka and Alvar Cuevas i Fajardo, IT-CO

New techniques are developed to manage virtual machine images

Virtual machines (VMs) can add agility to computing environments by satisfying different execution environment requirements on single machines, providing flexible testing environments and allowing more robust services with Live Migration. In these scenarios the infrastructure needs to manage the VM images being deployed efficiently, which can vary in number and range in size from several hundred megabytes to several gigabytes. Also, up-to-date VM images need to be generated automatically because doing this manually is time consuming and error prone.

Virtual appliances extend the concept of VMs and offer a preconfigured integrated application deployment inside a VM, meaning that it can be configured more easily and has a smaller impact on the hosts of the applications and services. It opens the door to a novel and expanding ecosystem for deploying services and computing, such as cloud computing.

Fig. 1. The OS Farm user interface is a service that provides virtual machine images.

OS Farm

OS Farm is a service similar to rPath, which provides images and virtual appliances for use in VMs. OS Farm takes advantage of a tool and library developed in the IT-GD group, called Libfsimage, that generates VM images with many different flavours of Linux, including Scientific Linux CERN (SLC), Debian and Ubuntu. By using Libfsimage in the backend, OS Farm provides a service to create and store VM images and virtual appliances. This is similar to rPath, but the focus differs in that images are created on demand and are more faithful to the originating OS. OS Farm provides a web interface through which users can configure their images. Users can choose between a number of software repositories, offering a rich set of software packages that can be selected and added to the image configuration. On creation, images are stored in a browseable repository from which they can be retrieved later. Users can also submit a configuration specified in XML or interact through a Simple Object Access Protocol (SOAP) interface.

Automatically generating an image from scratch is a long process, often requiring downloading and installing hundreds of packages, usually taking several minutes. However, parts of images are identical and can be shared between different images. Image generation is optimized by layering the different parts of an image. For example, the "core" of two different images can be a shared layer, while additional configuration can be appended on top. Thanks to the copy-on-write, or "snapshot", capability of Linux's Logical Volume Manager, the application of an existing layer is instantaneous. Using this technique, the generation time can be shortened to around 1 min.

To optimize the generation time further, each image in the repository is tagged with a checksum of its configuration. If an image is requested that already exists in the repository, it can be returned instantly. The repository thus also serves as a cache of recently generated images.

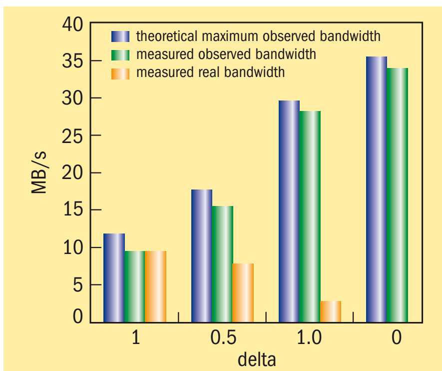


Fig. 2. Comparison of CBT bandwidths in different degrees of commonality.

Content-based transfer

Across different VM images there can be a substantial amount of commonality. The way in which the data is laid out in the images allows the commonality to be identified. For example, comparing two images made from different batch systems in the computer centre showed that 84% of their blocks were identical. Comparing two asymmetrical SLC3 and SLC4 images showed that 48% of the blocks contained in SLC3 were found in SLC4, and 22% of SLC4's blocks were found in SLC3.

Content-based transfer (CBT) is a technique that efficiently transfers VM image data from a source host to a target host. It takes advantage of commonality between images to speed up the transfer and reduce the load on the network. We can calculate the hypothetical maximal observed bandwidth from the degree of commonality in the image being transferred, the physical network bandwidth and the disk-read bandwidth. "Observed bandwidth" here means the bandwidth observed from transferring the image, although only the uncommon blocks and identification data have actually been transferred. Moreover, for the common blocks, which do not need to be transmitted, the bandwidth is bound by the disk, and for the differential blocks the bandwidth is bound by the network.

Experimental analysis shows that our implementation of CBT achieves an observed bandwidth close to the theoretical maximum and reduces the load on the network. For example, on a 100 mbit network, given 90% commonality, CBT achieves an observed bandwidth of 28 MB/s – a speed-up of 2.2 times as much as the maximum. At the same time, the impact is reduced to less than 3 MB/s. Both OS Farm and CBT are proofs of concept, but they are freely available for download and can be used under open-source licences.

Useful links

OS Farm project: http://sourceforge.net/projects/osfarm/.
OS Farm demo: http://cern.ch/osfarm.
Content-based transfer library: http://hbjerke.web.cern.ch/hbjerke/cba/cba.xml.

Håvard Bjerke, IT-DI (openlab)

WLCG workshop examines the readiness for LHC data taking

A WLCG collaboration workshop was held on 21–25 April at CERN and had more than 230 registered participants. It was the third event of its kind. The main theme of the workshop was readiness for LHC data taking and processing, which naturally included discussions about the February and May runs of the Common Computing Readiness Challenge (CCRC'08; described in the January–March *CNL*), as well as the experiments' full dress rehearsals and other similar activities.

Parallel sessions were organized on database services (Maria Girone), monitoring (James Casey), as well as on the experiments. All presentations can be accessed through the agenda page at http://indico.cern.ch/conferenceTimetable.py?confId=6552, together with notes taken by kind volunteers. Live webcasting was available for remote (passive) participants. Thanks are due not only to all speakers, conveners and note-takers, but also to all for attending and in particular those who participated in the discussions.

The workshop started with a review of the WLCG critical service standards – a simple set of proven techniques for designing, implementing, deploying and operating robust and resilient services. These are targeted at the priorities set by the experiments and are one of the key metrics by which CCRC'08 is being evaluated. Unfortunately, not all middleware components are yet able to take advantage of these techniques, which (as highlighted in a review by Maarten Litmaath) are of interest to many other communities beyond WLCG. Birger Koblitz explained the work continuing within ATLAS to apply these methodologies to ATLAS-specific services, again driven by the critical services lists.

Discussions then turned to storage optimizations, as well as ongoing issues with SRM v2.2 production services (the common mass-storage interface). These resulted in a new set of metrics for evaluation in the May run of CCRC'08 – to be revisited in a future article.

The next two full days were devoted to an analysis of the results of the February run, as well as preparation for May and beyond. Very detailed presentations were made by the experiments, together with discussions on service issues and those seen by the sites. Explaining exactly what is required of different sites and/or tiers remains a problem that is being addressed urgently. Specifically, the middleware and storage-ware versions required for May are well defined and are available through the CCRC'08 wiki.

The final topic discussed in the plenary sessions was future WLCG operations, which included an introduction to the EGI design study (Dieter Kranzlmüller), WLCG requirements (Ian Bird) and the EGEE-III operations workplan (Maite Barroso Lopez). The importance of a timely and non-disruptive transition to the post-EGEE environment was stressed by many – 2010 LHC data-taking and processing cannot be placed at risk by changes in this area.

In addition to the CCRC'08 post mortem foreseen for 12–13 June at CERN (http://indico.cern.ch/conferenceDisplay.py?confId=23563), future events are planned during the EGEE'08 conference in Istanbul, and another WLCG collaboration workshop in conjunction with the CHEP'09 conference in Prague (http://indico.cern.ch/conferenceDisplay.py?confId=16861).

Jamie Shiers, IT-GS

EGEE project in transition to an expanded level

More than 90 participants – some from as far away as Taiwan – came to the two-day transition meeting to EGEE-III, held on 6 May at CERN. Another dozen people "attended" virtually, via video.

The goal of the meeting, said EGEE project leader Bob Jones, was how best to make the transition from the existing, project-based Grid computing infrastructure – built to support the LHC project and applications from other scientific disciplines – and adapt and expand it to a new level (EGEE-III), which is scheduled to last from 2008 to 2010. EGEE-III has expanded and now represents 250 sites from 48 countries, working on applications from areas of science as diverse as geophysics and multimedia. The long-term plan is then to shift to a sustainable, long-term, European-wide permanent structure to be called the European Grid Initiative (EGI).

"By 2010 we want EGI to be fully in place as something that can be returned to again and again by a variety of projects from other disciplines of a similar size to LHC, which each last differing time

EGEE-III project leader Bob Jones addressed the transition meeting at CERN in May.

periods. In the next four years we see a tremendous number of big, data-intensive projects coming online, such as the Square Kilometre Array or the Very Large Telescope," he said. "The role of this meeting, and of EGEE-III as a whole, is how to seamlessly make the transition, and coordinate the work of the 120 different institutions and thousands of people involved. Coordinating the integration and interaction among the different National Grid initiatives is key."

Other speakers emphasized the mechanics of this, calling for more automated tools, simpler and more robust middleware, and an easier to use interface.

"The Grid is based on the idea of give and take, with users 'paying' in kind with resources and computer time for the data and resources that they have used. You help to process and in return you get access," Jones explained. "In order to make people willing to participate, you have to make the procedures as simple as possible."

Attendees pointed out that it was symbolic that many of the tools used for running the conference – such as the PPT tool for time- and record-keeping, the SIMBA e-mail and InDiCo (which eliminated the need for any memory sticks to carry slides), and the Engineering Data Management Service (EDMS) – were originally developed at CERN for the LHC.

"We make full use of these tools," said Jones. "Their seamless integration shows how a lot of CERN IT can be used to support other programs."

Dan Drollette, iSGTW editor

CERN welcomes HEPiX back to Geneva in May

For the first time since 1992 (and only the second time in its history), HEPiX met at CERN for the Spring 2008 meeting held on 5–9 May. Some 95 people registered and there were a record 65 submitted talks, completely filling the available time slots. The format was slightly different, with plenary talks on some aspect of the LHC or the computing for it, followed by the usual days split into blocks by topic. All slides are on the web at http://indico.cern.ch/conferenceDisplay.py?confId=27391.


HEPiX met at CERN for the Spring 2008 meeting. Image courtesy of Maximilien Brice.

The plenaries

The meeting was opened by CERN's chief scientific officer, Jos Engelen, with a status report on the LHC and its experiments. He explained how the final cooldown had started and was expected to be finished by July. In parallel, hardware commissioning of the magnets was ongoing. Full cooldown from room temperature takes about six weeks plus about three weeks for stabilization. When the whole ring is cold, there will be about 130 tonnes of helium circulating. Other plenaries included a review of the WLCG project by the project leader, Ian Bird; an explanation of the goals of the Combined Computing Readiness Challenge by Jamie Shiers; a description of the activities of the CERN openlab project by its chief technical officer, Sverre Jarp; and a review of networking in the LHC era by David Foster. Bob Jones, project director of EGEE-III, took the opportunity to brief the audience about this new phase of EGEE.

New data centres

Once again there were many talks on new data centres (RAL, BNL, St Louis) and upgraded ones (NIKHEF and SLAC, if you include the addition of a second SUN Black Box – a converted shipping container that is delivered full of computers in racks with cooling and ventilation built in). During these data centre talks, no fewer than three incidents were reported from three different sites of electrical problems leading to fires and burned equipment. Of course, all sites reported increasing purchases as we build up to LHC data taking. To aid with these acquisitions, the HEPiX Benchmarking Working Group is investigating which benchmarks are the most appropriate for 2008 and beyond, and attention is focusing more and more on the integer part of SPECint2006. Among the new data centres discussed, BNL described a plan to build and make operational a new 600 m^2 data centre in record time (18 months instead of the more usual 30 months minimum) and for a limited budget of less than $5 million. It will be interesting to monitor the progress of this.

File systems

In previous meetings, Lustre had appeared as an interesting file system. This week, judging by the number of times it was talked about, it is now the file system of choice across much of the HEP world. This was confirmed by the presentation of the results of carefully prepared tests performed by the HEPiX Storage Working Group for whom Lustre is now the recommendation for a cluster file system. In the area of LHC data storage, members of the corresponding IT group explained their short- and long-term plans for improving the reliability and efficiency of storing the data and making it accessible at all levels of the WLCG. At CERN this is based on CASTOR and it was interesting to note that, at the same time as RAL (UK) is switching from dCache to CASTOR, CNAF (Bologna) is switching away from CASTOR to a combination of GPFS and TSM–IBM products.

Operating systems

Scientific Linux (SL) futures were discussed in detail in the presence of the principle authors, and a number of decisions were taken regarding support lifetimes and future directions. For example, it was agreed that FNAL will attempt to maintain support for versions 3 and 4 of SL until the third quarter of 2010. This may permit some sites to remain on version 4, skip the current version 5 and move directly to version 6, which is expected to be released by then. Talking of new versions of operating systems, there was an update of the CERN talk from the last meeting on the recommendations for upgrading Windows to Vista (soon to become the default on newly acquired desktops) and several labs have now also adopted or are seriously considering similar policies.

Invited talk

The meeting featured an invited talk from Google, for which at least 50 people came in from CERN's IT and PH departments. The speaker gave a review of data storage at Google. The problems and issues were familiar but the numbers were on a different scale to most HEP sites. Their largest individual clusters typically have 5000+ servers and will be serving 10 000+ clients with 5+ PB of files. A single cell of their online database can manage up to 6000 TB of data spread across 3000+ nodes, supporting 500 000+ operations per second, sometimes with higher peaks.

Magnetic tapes

Another highlight was a presentation on making better use of magnetic tapes. The speaker, Charles Curran of CERN/IT, gave a personal view on how the useful cycle of read and writes could be improved by developing codes that take more account of the features of modern devices that actually access the data on the tape.

Other sciences

One interesting point is that a growing number of HEP labs are adopting or adding new sciences. It may or may not be surprising that many are meeting common problems and it appears that once again HEPiX is being used as a suitable forum for sharing these experiences. This has been the way, and still is, for the pure HEP labs since the founding of the group in 1991. This expansion is also proved by the continuing representation of two of the world's leading Genome Sequencing Centres (for the second consecutive time).

Security

As is now traditional at HEPiX, the meeting closed with sessions on security, where experts from CERN scared the audience with tales of known or suspected security exposures in mail, operating systems and, especially, web browsing and web applications.

During the final session tribute was paid to Wojciech Wojcik of IN2P3 in Lyon, for whom this will be the last meeting before his retirement. Not only has he served for the past few years as European coordinator of HEPiX, but he also claims the unique record of having attended every meeting of the group since it started in 1991.

The next HEPiX meetings include ASGC in Taiwan on 20–24 October, one in spring 2009, probably in the Nordic area, and possibly NERSC (Berkeley, California) in October 2009.

Alan Silverman, IT-DI

Information corner

Bookshop offers more titles

Since Roger Woolnough established the Bookshop in 1994 it has been a big success. It was a great idea because CERN staff are constantly in need of information. The Bookshop complements the Library for material that our scientists and engineers like to have at hand. In spite of the internet and the fact that e-books are gaining popularity, printed books are still very sought after.

The Bookshop has around 750 titles: 300 in computing science, 300 in physics and 150 in other subjects.

The books are purchased from about 15 publishing houses. However, the offer is not limited to these publishers – the staff are ready to order any title that exists in print. In addition, we welcome suggestions from the user community for new titles that they want to be added to the stock.

The title list is available from http://cdsweb.cern.ch/collection/CERN%20Bookshop and http://cdsweb/tools/itbook.py.

The Bookshop has opened new subject groups, including IT security, radiation protection, detectors and experimental techniques, and project management.

Its most recent acquisitions include the following titles.

Computing
Andrew S Tanenbaum *Modern Operating Systems* 3rd edn.
Steven Feuerstein *Oracle PL/SQL Best Practices*.
Penelope Coventry *MS Office SharePoint Designer 2007 Step by Step*.
Ken Coar and Rich Bowen *Apache Cookbook* 2nd edn.
Mark G Sobell *A Practical Guide to Ubuntu Linux*.

Physics
Olivier Morsch *Quantum Bits and Quantum Secrets: How Quantum Physics is Revolutionizing Codes and Computers*.
Steven Weinberg *Cosmology*.
Zyun Francis Ezawa *Quantum Hall Effects: Field Theoretical Approach and Related Topics*.
Thomas F DeLaney and Hanne M Kooy *Proton and Charged Particle Radiotherapy*.
Helmut Wiedemann *Particle Accelerator Physics* 3rd edn.
Andrew Sessler and Ted Wilson *Engines of Discovery*.

Other subjects
Jean-Yves Le Meur *Faux pas*.
Andrew Whittaker *Speak the Culture France: Be Fluent in French Life and Culture*.
Harald Fritzsch *Escape from Leipzig*.

The Bookshop is located in the Central Library, Building 52/1-052 and is open on weekdays from 8.30 a.m. to 5.30 p.m. It can also be contacted by e-mail (bookshop@cern.ch). CERN users can buy books and CDs at discounted prices.

Jutta Megies and Roger have now retired and the service has been passed on to the Library. Joanne Yeomans has taken over general responsibility from Roger, and Eva Papp has taken over Jutta's tasks.

Eva Papp, CERN Bookshop

Calendar

June

30 June – 1 July **EGI Workshop**
Geneva, Switzerland
http://web.eu-egi.eu/events/workshops/geneva-2008

30 June – 4 July **High Performance Computing and Grids Workshop**
Cosenza, Italy
www.hpcc.unical.it/hpc2008/

July

5–8 **International Conference on Software and Data Technologies**
Oporto, Portugal
www.icsoft.org/

6–18 **International Summer School on Grid Computing**
Balatonfüred, Hungary
www.iceage-eu.org/issgc08/

21–24 **4th International Conference on Computer Science and Information Systems**
Athens, Greece
www.atiner.gr/docs/Computer.htm

August

18–29 **Introduction to High Performance Computing**
Stockholm, Sweden
www.pdc.kth.se/systems_support/training/2008/summerschool/

25 August – 5 September **CERN School of Computing**
Gjøvik, Norway
www.cern.ch/CSC

26–29 **14th International European Conference on Parallel and Distributed Computing (Euro-Par 2008)**
Las Palmas de Gran Canaria, Spain
www.caos.uab.es/europar2008/

Berners-Lee offers advice on social sites

When Tim Berners-Lee handed his CERN supervisor his proposal for a HyperText project, did he have any idea that years later the web would become so ingrained in people's lives? Today sites such as *Facebook* boast around 70 million users and *Myspace* a mighty 200 million. When interviewed by the BBC recently, Tim was asked what advice he would give to users uploading personal information to sites such as these. Here is his response:

"Imagine that everything that you are typing is being read by the person you are applying to for your first job. Imagine that it's all going to be seen by your parents and your grandparents and your grandchildren.

"The danger is when you put something into a public space to share it with a few friends you forget that it's actually a public space or that the list of friends is huge or that some of them can't be trusted not to put it somewhere else."

For the full interview, see http://news.bbc.co.uk/2/hi/technology/7300434.stm.

As CERN computer users, we have to be aware not only of our personal reputations if using social networking sites, but also of the organization's reputation.

For more information on the pitfalls of social networking, readers may find this *ComputerWorld* article useful: www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=311395.

Kate Bradshaw, IT-DI

Fake mail targets CERN

CERN's computing rules (Operational Circular No. 5) require that passwords remain confidential and must never be given to anyone, not even Helpdesk or other support personnel. If you think that your password could have been exposed then you must change it immediately. For further information, consult http://computingrules.web.cern.ch/ComputingRules/.

Cybercriminals are making growing use of fake e-mails and websites to steal account and identity information from users. This technique is called phishing. The e-mails often look real and try to trick you into giving your password or personal data.

CERN has recently been targeted by a fake e-mail that asked people to send their username, password and date of birth. This case was blocked by the CERN mail service, but future cases may still occur. Users need to be vigilant about any requests for passwords or personal data. Fake e-mails should be deleted immediately.

Contact helpdesk@cern.ch if you have any questions on this topic.

The CERN Computer Security Team