CERN COMPUTER NEWSLETTER

Volume 42, issue 3 June–August 2007

Featured article
Learn more about Vista on p4

Contents

Editors Nicole Crémel and Hannelore Hämmerle, CERN IT Department, 1211 Geneva 23, Switzerland. E-mail cnl.editor@cern.ch. Fax +41 (22) 766 8500. Web cerncourier.com/articles/cnl.

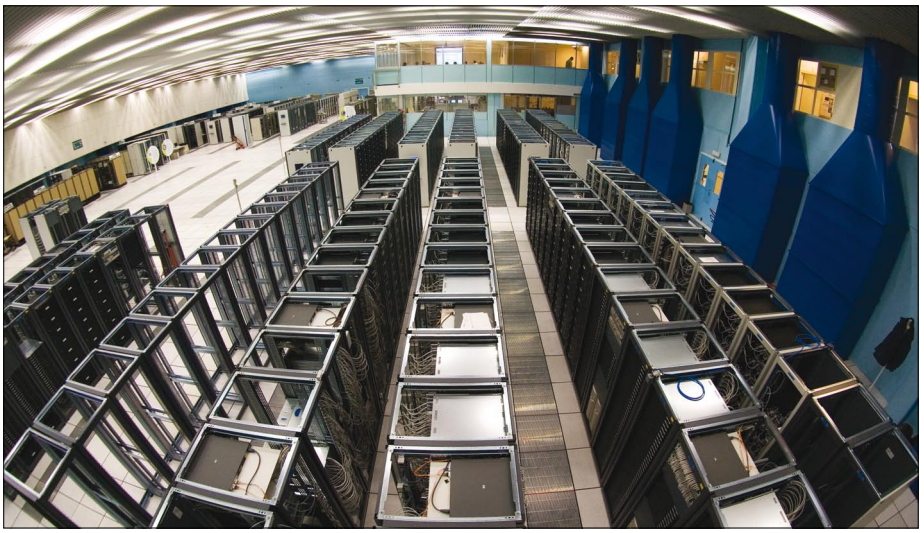
Advisory board Wolfgang von Rüden (head of IT Department), François Grey (IT Communication team leader), Christine Sutton (*CERN Courier* editor), Tim Smith (group leader, User and Document Services).

Produced for CERN by IOP Publishing Dirac House, Temple Back, Bristol BS1 6BE, UK. Tel +44 (0)117 929 7481. E-mail joseph.tennant@iop.org. Fax +44 (0)117 920 0733. Web iop.org.

Published by CERN IT Department

IOP Publishing

FIO prepares Tier-0 centre for LHC data



The Computer Centre at CERN is the Tier-0 site of the Worldwide LHC Computing Grid.

Since the beginning of the year *CNL* has been visiting the groups in CERN's IT Department that provide support to the LHC Computing Grid (LCG). This issue we interview Tony Cass, the leader of the Fabric Infrastructure and Operations group (FIO). With nearly 60 members, the size of the group reflects the complexity of the tasks needed to run the CERN Computer Centre, which is the Tier-0 site of the Worldwide LCG collaboration and also offers Tier-1 functionality.

How is FIO structured to cope with the major changes that are occurring in the Computer Centre, in preparation for the LHC?

There are five sections in the group. In alphabetical order, the first is the Fabric Development section, which looks after the ELFms (the Extremely Large Fabric management system); this includes the LEMON monitoring system, Quattor for system configuration, and the CASTOR storage manager. These are grouped to ensure that all the software we support has a common design and development methodology, and to enable people to move between different software projects as priorities change. This has worked over the last year with people from ELFms, where development efforts are ramping down, moving to support CASTOR. Additional effort is needed for CASTOR at present, as increasingly realistic tests in preparation for LHC operations have highlighted some challenges. However, CASTOR2 is proving to be generally more robust than CASTOR1, and has sustained data throughput at more than 1 GB/s as well as meeting targets for data transfer to the Tier-1 sites.

The second section is Fabric Service, which is at the front line of delivering services to physicists. The team looks after disk and batch services, configures them and makes sure they are running to meet the experiments' needs. The team also works closely with the Grid Deployment group to make sure that what we have in the Computer Centre is available at the Grid level, and that we are reporting the status of our machines correctly.

Then there is the Linux and AFS section. Although aging, AFS is a reliable file system that, together with the Kerberos authentication system, underpins much of how we manage our clusters. On the Linux side, the team has a challenging task given our technologically aware and demanding

Editorial

user community. In addition to general support, the team focuses on optimizing the Linux kernel to deliver the highest performance, especially for the file systems, for CASTOR and for networking.

The Technology and Storage Infrastructure section looks after the tape robots. We now have about 12 PB of tape capacity and must ramp up significantly for LHC operations next year. We're also switching over from the obsolete 9940 drives and Powderhorn silos to new IBM and StorageTek robots. So there's a lot of effort behind the scenes to copy 5 PB of existing data to new media. This section also looks after procurement. We installed 1200 PC boxes last year and we expect to buy at least as many this year. Finally, this team also handles hardware monitoring. We have several thousand disks now and it is important to understand how they and the RAIDs (Redundant Arrays of Independent Disks) are functioning.

Last but not least we have the System Administration and Operations section. This groups the system administration team – which provides a *piquet* (on-call) service and first-level cover for most of our systems – and long-term planning for the machine room. To install 2500 PCs you have to sort out racks, cabling and a lot of logistics well in advance. And of course, a growing preoccupation is to provide adequate cooling for all this equipment.

Power and cooling seems to be an increasing challenge for data centres. What is your view?
Looking into the future there are some eye-opening statistics. At the moment we use less than 1 MW to power the equipment in the Computer Centre. If the cooling stops, the temperature rises by half a degree per second. By 2009, when the rest of the equipment will be installed, temperatures will increase by one degree per second if cooling fails. This leaves little time to protect the equipment.

My biggest concern, though, is that computing power is projected to grow for years to come. Although chips are becoming more efficient this is not occurring fast enough to offset the overall growth in processing needs from the LHC experiments. A conservative estimate is that the Computer Centre will consume 20 MW by 2020. If we extrapolate from the growth during the Large Electron Positron collider era, however, 100 MW may be a more accurate estimate. Based on preliminary investigations it will not be possible to upgrade the Computer Centre for this, so we will need another data centre at CERN or elsewhere. Nothing is excluded and various hosting solutions are being investigated.

How is the status of Tier-0 services visible to users and to the Grid?
The LEMON status displays are open for all to look at but are designed for a specialist audience. A recent success for the group is SLS, the Service Level Status display. This is a user-oriented view of the service status that helps users understand how the Tier-0 and Tier-1 services at CERN are behaving, and allows them to drill down to individual components if necessary. This year we aim to deliver an XML interface to our monitoring database. This will enable experiment production managers to take decisions on what to do: for example, throttling back on production if there are problems with CASTOR.

Tony Cass, the group leader of IT/FIO.

In a similar vein, we want to avoid parallel monitoring systems querying sites on the Grid, so our data is being fed into systems like NAGIOS and GANGLIA even if we don't use those monitoring tools ourselves.

Reliability remains an issue for the Grid. How is the Tier-0 faring in this respect?
We're processing 5000–6000 jobs round the clock, with peaks of 40 000 jobs in the queue. We easily run 200 000 jobs each week and I'm pleased that ours is the most reliable site on the Grid. We've met the target every month since it has been measured and we're easily 95–97% reliable.

Considerable ingenuity is required to maintain this sort of reliability, especially given the evolving nature of the Grid middleware. For example, the middleware queries the batch system once per minute per job to check if it is still running. At a Tier-2 site with a few hundred jobs, that is manageable. But for 6000 jobs we're getting hundreds of queries per second. So we had to develop a caching system to avoid overloading when just reporting that everything is OK.

I would emphasize, though, that the bottom line to providing such high reliability is the level of commitment of everyone in the group. I've lost count of the number of times people have come in at weekends and at night to fix even minor problems. This commitment of the people in the group is ultimately what makes a 24/7 service a reality.

Computing articles featured in this month's CERN Courier

The articles listed below appear in the June issue of *CERN Courier*. Full-text articles and the rest of the issue's contents are available at www.cerncourier.com.

Computing News

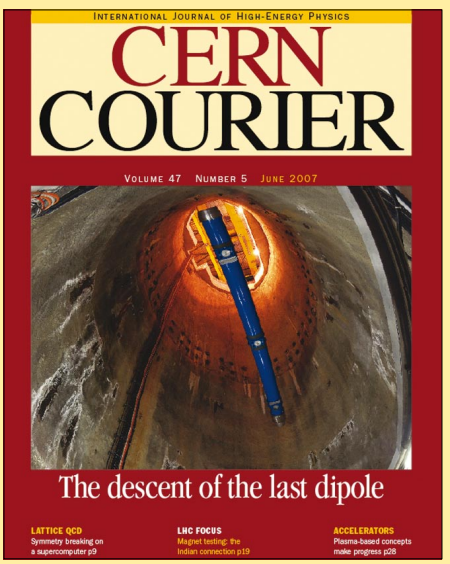
- **French LCG steadily ramps up**
Members of the French section of the LCG project meet to discuss progress.

- **CERN wins prize for storage systems implementation**
Computerworld recognises ALICE data-acquisition system with award.

- **PACE makes way for European collaboration**
A new initiative will ensure that important research has access to supercomputers.

- **CERN openlab gains new contributor**
The information-technology service provider EDS joins CERN openlab.

- **Thin clients use less energy at work**
Studies show that thin clients consume up to 50% less electricity.



- **GridPP extends to deal with LHC data**
PPARC enables GridPP to run until 2011.

Calendar of events

IT strengthens firewall security

On 29 May CERN increased again the security of its firewall. All servers that need to be directly accessible from the internet (e.g. mail and web) must now have prior authorization and be configured explicitly in CERN's main firewall. This strengthening of the firewall also affects the LXPLUS Linux cluster by restricting its access from off-site to the SSH server.

In the past users could expose some applications to the internet without prior agreement, but this has led to security incidents that could have placed the whole site at risk and/or impacted the reputation of the laboratory. We have had to take action to close this loophole.

CERN began to strengthen its firewall some years ago by protecting services that were known to be targeted by attackers. Now any service offered to the internet is considered to be a target for attack. The situation has become too dangerous to permit external exposure without justification and without security checks being carried out in advance.

To make it easier to manage firewall access for the many CERN services that must be made available to the internet, IT department's Communication Systems group has integrated firewall management into their network database tools. This enables registered system administrators to view and request firewall modifications for the devices they manage. This new system has been described in a previous article, "CERN upgrades firewall to meet requirements of LHC" (see http://cerncourier.com/articles/cnl/4/3/4/1).

Denise Heagerty, IT security team

CERN intercepts more than 90% of spam mail

Spam is a difficult issue, not to say a nightmare, for all mail administrators. We have published articles about this in the past, and provide regular updates to assure users that the mail team is continually fighting spam.

CERN e-mail addresses receive up to 2 million spams each day. Spam falls into two categories:

- the "obvious" spam (more than 90%), which is immediately suppressed and is never delivered to mailboxes or forwarded;
- the remaining 10%, which is not as obviously detected and might occasionally contain some legitimate messages. For this reason the "not-so-obvious" spam mails are sent to the destination mailbox. When the destination mailbox is at CERN these suspicious messages are delivered to the Spam Folder, according to the user filtering level that is configured with the Spam Fight tool on the MMM website.

When a "mail forward" is activated, the 10% of the not-so-obvious spam are also forwarded (they might include some legitimate messages). It is up to the internet service provider or the home institute hosting the destination mailbox to provide its own special folder for suspicious messages. This same rule applies to users who have set their physical e-mail address (PEM) to their local institute or at home.

We would remind users that the "MMM forward" facility should only be used to forward incoming e-mails to another address temporarily. To forward mail on a permanent basis, users should set their PEM to their external e-mail address and avoid having two different mailboxes. (For more details, see the article "Users may need to request deletion of mailbox" at http://cerncourier.com/articles/cnl/4/1/4/1.)

The mail team

Same authentication for AIS and NICE/MAIL

Over recent years the IT department has been streamlining CERN users' access to all of the central computing services. The long-term goal is to converge on a unique "CERN authentication" mechanism, which will increase computer security and make it easier to maintain accounts.

The next step of this process took place on 4 June. Now all users at CERN require the same username and password to authenticate on the AIS applications (such as EDH, HRT, CET, APT, ERT, CRA and Foundation), and on NICE (Windows) and MAIL. We would like to remind you that this "CERN authentication" can also be used on EDMS, Indico, CDS and SIMBA.

So, CERN users must now authenticate on the AIS applications using this username and password for the CERN authentication.

The AIS and NICE teams

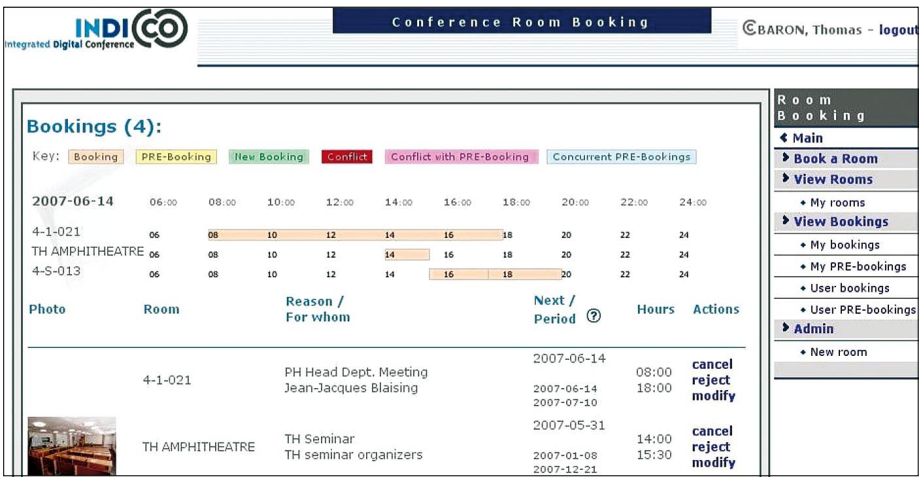
Booking rooms with Indico is child's play

Good news for all meeting enthusiasts – booking a room at CERN has never been so easy. On 26 May the aging Conference Rooms Booking System (CRBS) application was replaced with an integrated Indico module that greatly eases the process of reserving a meeting room.

Indico, the Integrated Digital Conference tool, will guide you through the booking process by proposing enhanced room searching interfaces and improved options. Managing your bookings is child's play with this new application: simply log on and you'll be presented with the list of your bookings, with options to cancel or modify them. A stand-alone interface also exists for those who do not yet manage their meetings through Indico.

There is also good news for room managers, since the new application enables them to choose between several new methods for handling incoming requests: for example, it is now possible to accept all bookings automatically. Room data is managed through the AIS/Gesloc application to avoid database duplication.

The Indico team



The Indico application enables users to cancel or modify their conference room bookings.

IT/IS evaluates Windows Vista

On 30 January Microsoft released Vista, a new version of the Windows desktop operating system. Since then we have been evaluating the operating system to ensure that the CERN computing infrastructure can support Windows Vista natively.

The last issue of *CNL* (April–May 2007) gave a humorous announcement of the arrival of Windows Vista, with a cartoon by Patrick Chappatte. We would like to introduce Vista in a more structured way. The following article describes the benefits of Vista compared to XP, includes a selection of new features and the hardware required to run Vista, and discusses issues of application compatibility.

Why Windows Vista at CERN?

In October 2001 Microsoft released Windows XP. At that time the IT department's Internet Services group (IT/IS) started to evaluate it and prepared a pilot deployment in 2002. After one year Windows XP became the default Windows operating system at CERN and started to replace Windows 2000 computers. The operation ended in 2006 when the IT/IS group followed Microsoft and stopped supporting Windows 2000. The move from Windows 2000 to XP was smooth because it was begun so early. Learning from that experience we know that this is the right time to start Windows Vista at CERN.

Another motivation for deploying Vista is that it is more secure than Windows XP. The design and development of this new operating system incorporates Security Development Lifecycle, which addresses concerns about security in the core components of the Windows kernel. As a result Vista contains many security-related features; the most visible one, User Account Control, is described later.

Today all desktop computers bought at CERN support 64 bits but are installed with the 32 bit version of Windows XP. This limits the address space and the total memory size to 4 GB. What is more, the Windows operating system reserves half of this address space to the kernel, limiting the total address space to 2 GB. Several applications used at CERN in the field of engineering, such as CAD, structural calculations, monitoring and databases, are limited by the 32 bit address space and may soon require a 64 bit desktop operating system. At CERN we need to ensure that the central computing infrastructure is compatible and can support this new generation of operating system. Although there is a 64 bit version of Windows XP we will jump directly to 64 bit Vista. It doesn't make sense to introduce 64 bit XP when Microsoft will stop supporting it in a few years' time.

What is new?

There are many changes in the new operating system, and we will highlight those that are the most important and visible for the end-user.

Graphical user interface

The most visible change is the graphical user interface. The sidebar on the right side of the screen (figure 1, note 1) contains small applications ("gadgets") that can keep information easily accessible. The sidebar can be customized to show, for instance, a calendar, the local weather forecast, the time or the content of an RSS feed that you have subscribed to.

Depending on the capabilities of the graphics card, the look and feel of your system can differ slightly. With the most advanced cards Vista activates the Windows Aero user interface. In this case the windows and start menu are transparent, which is called the "glass effect" (figure 1, notes 2 and 3). By moving the mouse cursor over the minimized items in the task bar on the bottom of the screen you can see thumbnails of the application windows, which can help to quickly identify the window you want to maximize (figure 2, note 4).

Another nice graphical effect, called "flip 3D", occurs when you change the active window using Win+Tab keys (figure 3). The Alt+Tab keys will give the same effect but in two dimensions.

File system

The major improvements to the file system are in the search and index capabilities. The Windows desktop search, which was optional on XP, is now a native component of the operating system and its performance has much improved. With Vista you can run a more advanced and almost instant search of the content of your disk.

Users can also specify additional tags, save them in the metadata information that is kept with every file and later search for documents that have a defined tag. Figure 4 shows how a Vista tag can be added to a

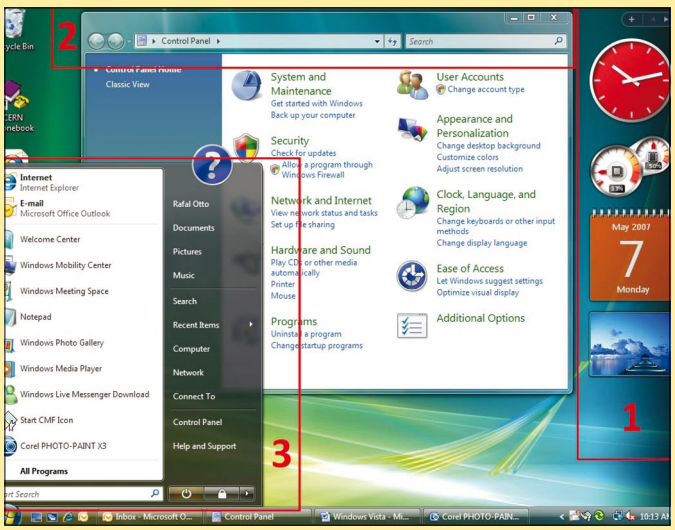


Fig. 1. (1) The sidebar on the right can be customized to contain frequently accessed information. (2,3) Advanced graphics cards activate the Aero user interface and provide a transparent effect.

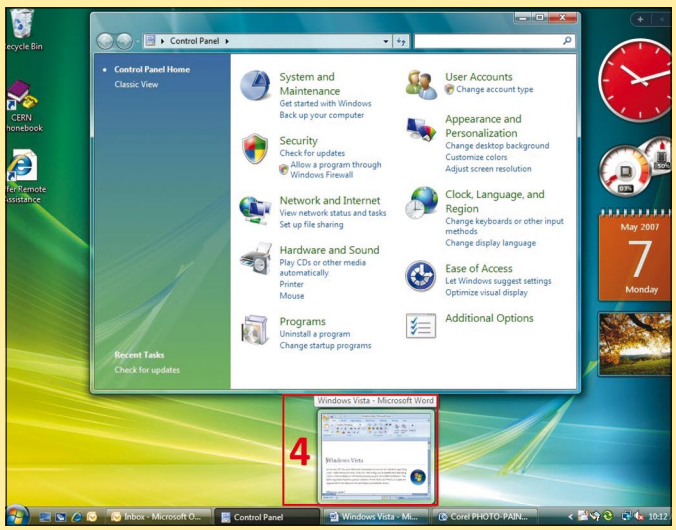
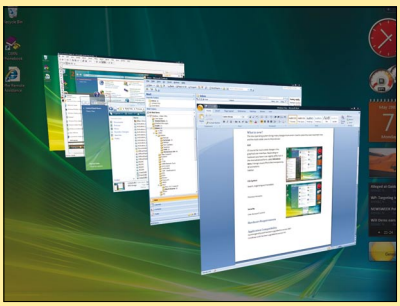
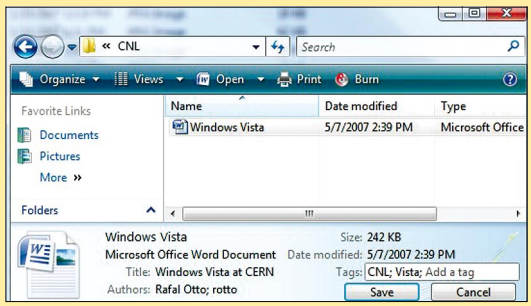
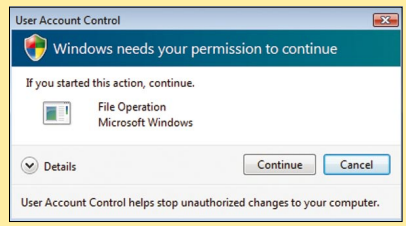
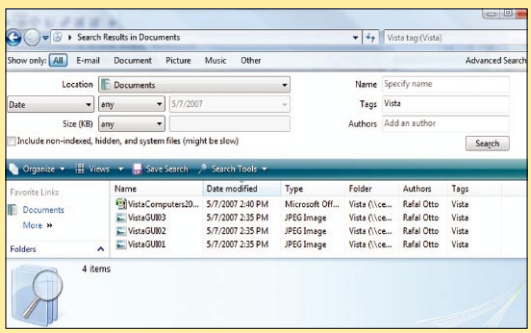


Fig. 2. (4) Moving the cursor over minimized items at the bottom of the screen enables users to see thumbnails of the application windows. This can help to identify the window to maximize.









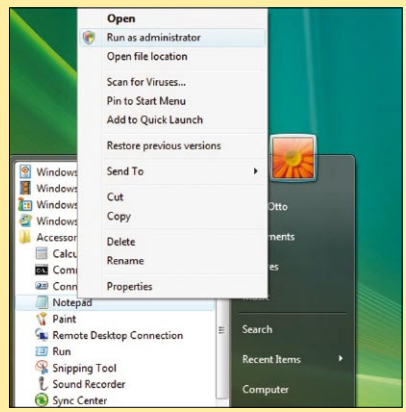


Fig. 3 (above). The "flip 3D" effect is achieved with the Win+Tab keys.
Fig. 4 (middle top). A Vista tag can be added to a Word document.
Fig. 5 (middle bottom). It is possible to search for files that have a Vista tag.
Fig. 6 (right top). The User Account Control feature prompts local administrators to authorize changes that are made to the computer.
Fig. 7 (right bottom): Users can force the system to run certain applications by selecting "Run as administrator".

Microsoft Word document, and figure 5 shows how you can search for files that have a defined Vista tag.

Another improvement is the Previous Versions feature, known as Shadow Copy on earlier Windows servers. Now you can restore previous versions of a document stored on your local disk whenever you need it, as well as recover deleted items. On Windows XP you can do that only with files stored on central servers.

Security

A list of all the security-related changes introduced in Vista would be long and probably uninteresting for most end-users, but one of them is worth mentioning because it will be visible to the user. This feature, called User Account Control, replaces the NICE Non-Admin tool. In Vista, each user who is a local administrator by default runs applications without administrator rights. Each time an application requires administrative action the user is prompted to agree for privileges to be elevated (figure 6) and becomes the local administrator just for the duration of that action. A user can also force the system to run certain applications with elevated privileges by right-clicking the application icon and selecting "Run as administrator" (figure 7).

Hardware requirements

Before installing Windows Vista on your computer you should be aware of its hardware requirements (see table). Microsoft distinguishes two levels of hardware compliance for Vista. A computer that is "Vista capable" can run Vista, although some features, especially graphical effects, will be disabled. A "Vista Premium ready" will run all features and graphical effects.

Table 1: hardware requirements for Windows Vista

	Vista capable	Vista Premium ready	CERN recommendation
Processor	800 MHz	1.0 GHz	1.0 GHz
Memory	512 MB RAM	1 GB RAM	1.5 GB RAM
Graphics card	DirectX 9 capable	DirectX 9 capable GPU with Hardware Pixel Shader 2.0 and WDDM driver support	DirectX 9 capable GPU with Hardware Pixel Shader 2.0 and WDDM driver support
Graphics memory	N/A	128 MB RAM supports up to 2 304 000 total pixels (e.g. 1920 × 1200) or 512 MB+ for greater resolutions such as 2560 × 1600	128 MB RAM supports up to 2 304 000 total pixels (e.g. 1920 × 1200) or 512 MB+ for greater resolutions such as 2560 × 1600
HDD capacity	20 GB	40 GB	60 GB
HDD free space	15 GB	15 GB	30 GB

Beyond the minimal hardware configurations, our tests show that 1.5 GB RAM will boost performance significantly, especially if the graphics card does not have the dedicated 256 or 512 MB memory. Similarly it is recommended to have at least 50% of the hard disk free, so a 60 GB hard disk is advised. This will avoid the decrease in performance that occurs when the operating system runs on a fragmented hard disk.

Application compatibility

Although Vista can run most Windows XP applications natively, there are no benefits in running the legacy desktop application on Vista compared to XP. That is why it is important to ensure that the desktop applications are compatible and support the latest Vista features. For this reason, Hummingbird Exceed has been upgraded to version 2007 and the CorelDraw suite has been upgraded to version X3. Similarly we plan to only support the 2007 release of Microsoft Office on Vista.

The status of Vista at CERN

Today a limited number of computers are running Vista at CERN. In August we plan to start a pilot project to enable end-users to install Vista on their computers. We are investigating certain issues with Microsoft and hope that these will be solved by then.

During the first phase of the pilot project we will propose the 32 bit version of Vista, and the 64 bit version will follow a few months later.

If the project is successful Vista may become the default installed operating system on new standard Windows desktop PCs before the end of 2007.

Rafal Otto and Alberto Pace, IT/IS

Dashboard monitors experiments

"This project started," says Julia Andreeva, "from some users saying that the Grid does not work." Andreeva, who works in CERN's IT department, coordinates a project that aims to show from the user perspective exactly what is working in a Grid and what is not.

The prototype for this project (Experiment Dashboard) began running in October 2005, monitoring jobs for the Compact Muon Solenoid physics experiment at CERN's Large Hadron Collider.

The Experiment Dashboard now provides a monitoring service for all four main LHC experiments. With it, researchers can access information about job processing, data management, test transfers and site efficiency.

At the beginning of this year the project reached significant milestones: putting a data management monitoring system for the ATLAS experiment into production and setting up job monitoring for the LHCb and ALICE experiments.

A special quality of this monitoring system is its ability to run on several Grid "flavours". The LHC experiments are using three Grid infrastructures: Enabling Grids for E-sciencE, Open Science Grid and NorduGrid.

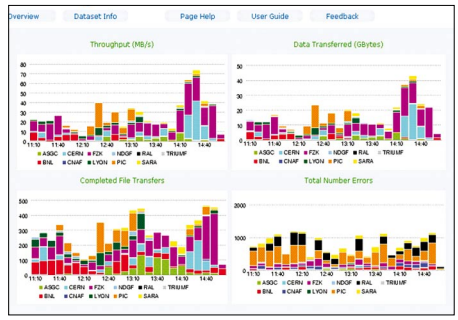
"Most monitoring tools are developed for a specific infrastructure," says Andreeva. "The advantage of the Dashboard is that it provides transparent monitoring for all infrastructures used by the LHC community."

JULIA ANDREEVA

The team of developers behind Experiment Dashboard. From left, Ricardo Rocha, Benjamin Gaidioz, Gerhild Maier, Julia Andreeva, Irina Sidorova and Pablo Saiz.

This quality makes Experiment Dashboard attractive to other virtual organizations, outside the LHC community, says Andreeva. The project has collaborators from Academia Sinica Grid Computing Centre in Taiwan, Moscow University and the Joint Institute for Nuclear Research in Russia, and Laboratoire de l'accélérateur linéaire in France.

The web-based framework of Dashboard also is able to give data in a form to be read by humans (standard HTML web programming language) and in a form that can be read by software applications (such as Extensible Markup Language or Comma Separated Values).



Dashboard provides a distributed data management monitoring system for ATLAS.

Several areas of Experiment Dashboard are targeted for improvement in the coming months: "We want to make it more user friendly and simple to understand," explains Andreeva. "And we want to add new sources of information to fill in areas where we are missing data. This will make it more accurate and comprehensive."

In the long term, Andreeva says, the principal task of the Experiment Dashboard will be to adapt to the needs of the LHC experiments as the experiments themselves evolve.

Danielle Venton, iSGTW editor

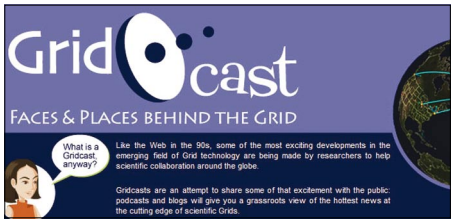
This article was published in *iSGTW* on 25 April.

Gridcast.org provides access to Grid forum

If you were unable to attend the joint 20th Open Grid Forum and the second Enabling Grids for E-sciencE user forum in Manchester, England, then the Gridcast at www.gridcast.org will give you a virtual tour of the event.

Gridcasts are a combination of blogs, podcasts and videos aimed at sharing "the hottest news at the cutting edge of scientific Grids" with the public.

In Manchester eight people with different science backgrounds participated in the blogging, sharing ideas, experiences and impressions – anything that might help web viewers feel as though they were there.



LCG-France ramps up for LHC

In March more than 70 members of LCG-France gathered to discuss the progress of French involvement in the global Grid infrastructure to meet the computing needs of the Large Hadron Collider, CERN's new particle accelerator, which is set to go online late this year. This meeting highlighted the challenging aspects of the Worldwide LHC Computing Grid (LCG) and the progress made during the last several months.

Many new members have recently joined LCG-France as part of the preparation for the LHC start-up. "The required data-processing capacity, the complexity of the installation and the management of the computing, as well as the effort needed to provide and operate a reliable platform for the experiments, have led to an important increase in the staff and in the number of French sites," said Fairouz Malek, scientific manager for LCG-France.

To cope with the large amount of data the LHC will produce, the LCG is organized in a "tier" system. Several Tier-3 sites, which will carry out analysis but not data storage, have joined the ranks of LCG-France. Currently the project is made up of a Tier-1 site operated by the IN2P3 Computing Centre in Lyon (which will receive data directly from the Tier-0 site at CERN), three Tier-2 sites and four Tier-3 sites. Several sites in other countries work with the French Tier-1 site; among these are collaborators from China, Belgium, Japan and Romania.

The French sites support all four LHC physics experiments. They also provide computing resources for non-LHC experiments and projects, mainly through Enabling Grids for E-sciencE.

The French infrastructure of the LCG was successfully tested in 2006 during data transfer exercises. All the French sites contributed to the exercises and the goals of the exercises were all reached.

The French Tier-1 site, whose budget has risen noticeably, has to sort out problems with electric installation and air conditioning so that it can cope with the planned increase of its computing

equipment. In spite of the difficulties faced, all measures have been taken to keep the Tier-1 operational while those improvements are made.

Finally, this meeting was also a chance for site representatives and members of the ALICE, ATLAS, CMS and LHCb physics experiments to exchange views and information, which helped strengthen links between the groups. The needs of the experiments, not only in terms of computing and data-storage capacity but also in terms of quality of service and bandwidth requirements, were clearly presented.

"The sites had the opportunity to hear about the particular needs of each experiment and to expose, in turn, their own constraints concerning the operations of the computing resources," said Fabio Hernandez, technical manager of LCG-France and deputy director of the IN2P3 Computing Centre. "These exchanges will certainly allow us to further improve the service for the benefit of the experiments."

Learn more by visiting the LCG-France website (http://lcg.in2p3.fr/wiki/index.php/Accueil).

Gaëlle Shifrin, IN2P3 Computing Centre (CNRS)

This article was published in *iSGTW* on 9 May.

UK's GridPP awaits collider data

Particle physicists in the United Kingdom have built the country's largest Grid as part of the GridPP project. Developed to analyse data from the upcoming Large Hadron Collider (LHC) at CERN, the Grid currently spans 17 UK sites and includes almost 10 000 CPUs and more than 500 TB of data storage, with available storage reserves on disk and tape of 1 PB. It is also an integral part of the Enabling Grids for E-sciencE (EGEE) project, contributing most of the computing power for the UK and Ireland.

GridPP is a collaboration of all the UK universities with particle physics departments. It also includes Rutherford Appleton Laboratory (RAL) in Oxfordshire, which acts as the UK Tier-1 centre for LHC computing. Steve Lloyd, GridPP's chairman from Queen Mary, University of London, explains the structure: "Apart from the Tier-1, the GridPP sites are joined into four regional Tier-2s: London, SouthGrid, NorthGrid and ScotGrid. These are used to share expertise and support. This works well, particularly for smaller sites where it would be difficult to provide dedicated help on the whole range of Grid issues."

Set up in 2001 initially for three years, GridPP was extended in 2004 to take it up to the switch-on of the LHC in 2007. The project is now funded by the recently created Science and Technology Facilities Council, and a further £30 million was announced last month to extend its remit until 2011.

"The first part of the project aimed to create a prototype Grid – which we did very well," said Dave Britton from Imperial College London, who will be project leader for GridPP3. "We're now focused on running a large-scale, stable, easy-to-use Grid integrated with EGEE and the WLCG [Worldwide Large Hadron Collider Computing Grid]. From later this year we'll see how successful this has been, when we start dealing with the petabytes of data from the LHC."

In addition to providing hardware, GridPP plays a role in applications and middleware. Among other achievements GridPP members have helped develop Grid accounting for the WLCG and EGEE projects, and a user interface for the ATLAS and LHCb physics experiments at LHC. The Real Time Monitor, another GridPP-developed tool, has held centre stage at numerous talks and conferences worldwide. RAL also hosts a Grid Operations Centre that monitors the overall EGEE grid, which has about 200 sites, tens of thousands of CPUs and petabytes of storage.

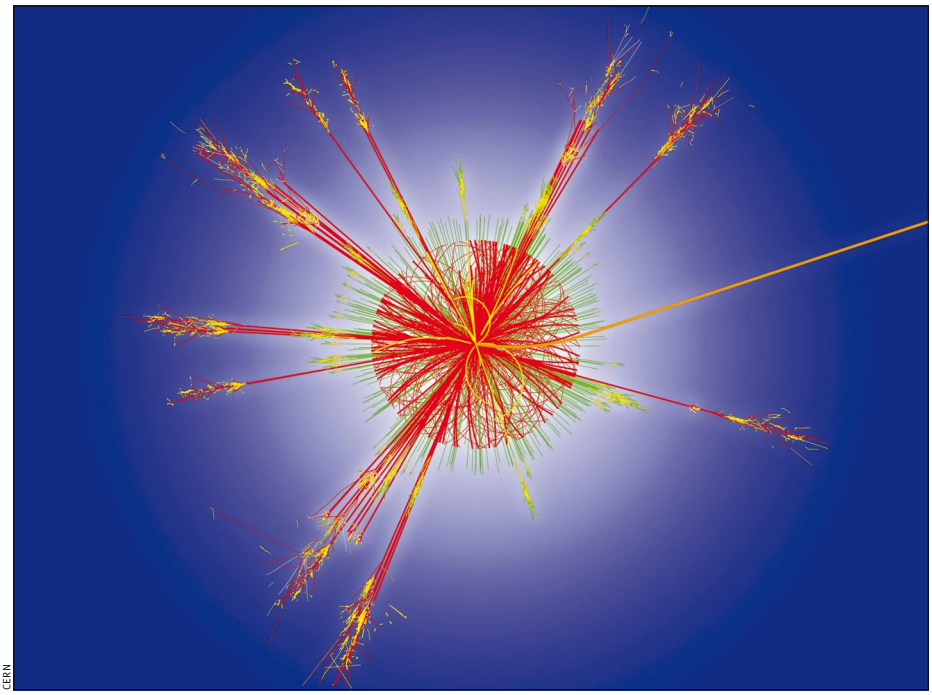
GridPP sites have contributed to what is now the EGEE Grid since its very first days, when RAL was one of only five sites in the first European DataGrid test-bed. Although the EGEE Grid now consists of hundreds of sites, the UK still provides a substantial proportion of its infrastructure – in the recent biomedical data challenge searching for drugs against malaria, known as WISDOM, GridPP provided nearly half the computing hours used.

Working with non-particle physicists is an important part of GridPP's remit. "Although our Grid was built to analyse particle physics data, when we have spare capacity we're able to share it with other scientists worldwide," said Tony Doyle, GridPP's current project leader. "We're happy to have contributed millions of hours of computer time to help find drugs against malaria and avian flu."

More information is available on the project's website (www.gridpp.ac.uk).

Sarah Pearce, GridPP, iSGTW contributing editor

This article was published in *iSGTW* on 11 April.



CERN

Simulated event of the collision of two protons in the ATLAS particle physics experiment.

GRIDPP

The GridPP stand at the EGEE'06 conference.

Collaboration workspaces ease development of web-sharing sites

Since the web was created at CERN in 1990 it has continued to evolve. You may have heard of Web 2.0, a buzzword that reflects the trend towards more collaboration and sharing between users. The web is now flourishing with collaborative applications, such as blogs, wiki sites (like the collaborative encyclopedia Wikipedia), social networking sites (e.g. Myspace) and online photo galleries.

The introduction of "collaboration workspaces" at CERN means that users can now easily set up their own collaborative web applications (figure 1). A collaboration workspace is a new type of website that is available from the central web management interface at http://cern.ch/web. The websites are based on SharePoint technology. The aim is to make it quicker and easier to share information and documents between CERN co-workers as well with people outside CERN.

Save time editing

Collaboration workspaces do not require a website editor tool; you can do most things from your web browser, including creating new pages and editing existing pages. In addition, all workspaces use a default template so that your site is ready to use immediately and there is no need to worry about formatting. See figure 2 for the types of content that can be created in collaboration workspaces.

Use wiki pages for quick web publishing

A wiki is a type of website whose content can be modified by several authors from the site itself. You can use wiki page libraries to store web content that changes often and involves more than one author, such as knowledge bases or operational procedures. Changes are tracked automatically so that it is possible to view older versions of a document and review changes.

Work with document libraries

Document libraries enable you to store and share any type of document in your collaboration workspace. Most importantly they provide an efficient way for several people to work on the same document: document libraries can be configured to automatically keep track of document versions as a document is edited.

The check-in/check-out feature enables a user to lock a shared document for some time, preventing concurrent editing. Approval workflow enables modified documents to be reviewed as necessary.

Create web forms and questionnaires

An important feature of collaboration workspaces is the ability to easily create web forms to collect data (figure 3). For example, it only takes a few minutes to create an online questionnaire: all that has to be done is to define the questions and the type of answers (e.g. yes/no, choice between possible values, or free text).

The collected data is stored in lists (figure 4). A list is similar to a spreadsheet or a database table that is managed online. To create a web form simply define the columns of a custom list, then a corresponding web form is generated automatically to feed the list.

Similarly, users can create and publish different views of the collected data. The list data can also be exported to a spreadsheet for analysis.

Use lists to share information

Several predefined lists exist in collaboration workspaces, such as contact lists, task lists and shared calendars. These provide a convenient way to share this information with others.

The lists can be connected to Outlook as additional contact, task or calendar folders, so that you can manage them side by side with your private contact, task and calendar folders (figure 5). The connection to Outlook also enables the online information to be synchronized with mobile devices such as phones and PDAs, as well as providing an offline copy of the information.

Exchange points of view on web forums

The discussion board feature enables web discussion forums to be set up quickly and to be opened to contributors outside CERN.

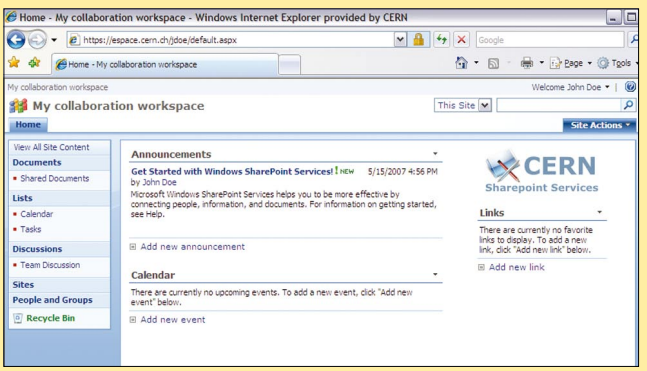


Fig. 1. It is easy to set up a new collaboration workspace.

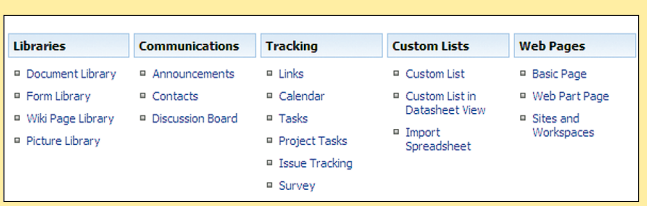


Fig. 2. Various types of content can be included in the website.

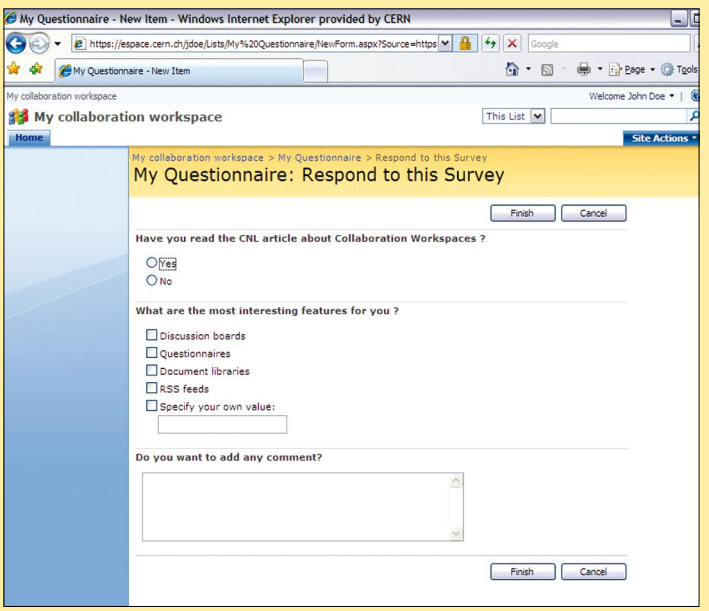


Fig. 3. Workspaces can be used to create online questionnaires.

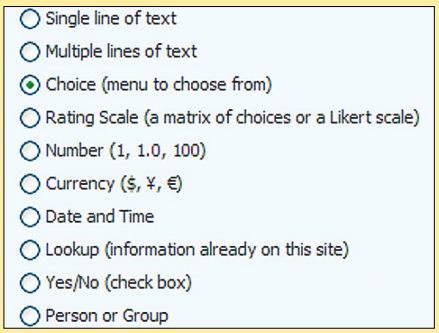


Fig. 4 (left). Different types of data can be stored in lists.

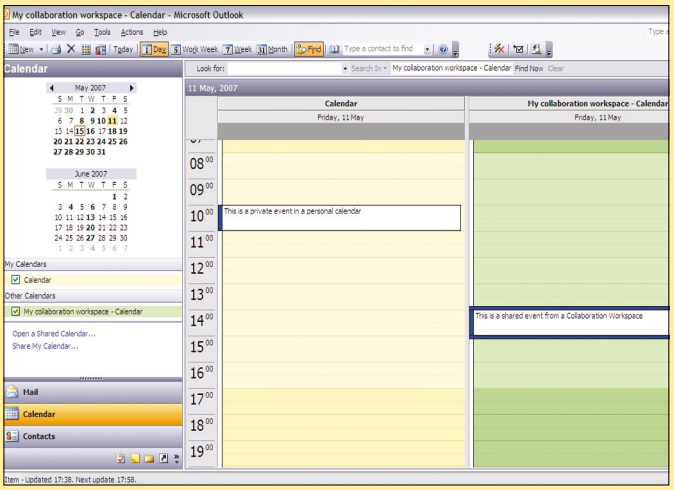


Fig. 5. Private and shared calendars side-by-side in Outlook.

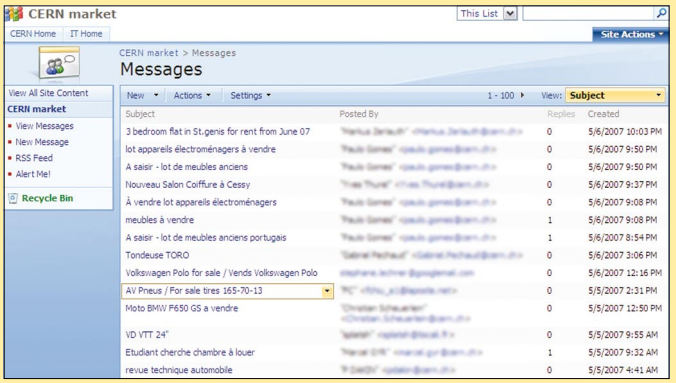


Fig. 6. The newsgroup cern.market as a web discussion forum.

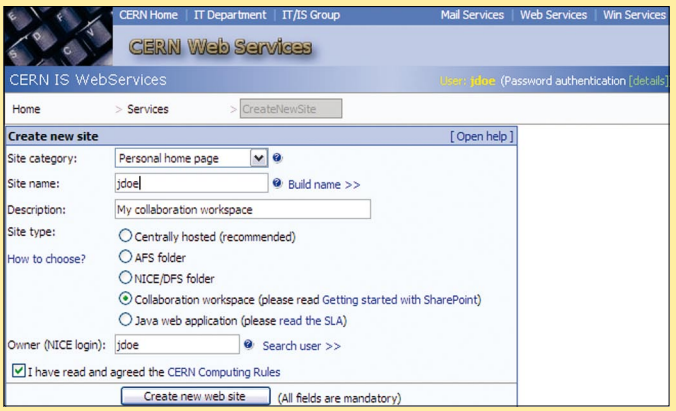


Fig. 7. Workspaces are created from the Web Services site.

Stay connected with RSS feeds and alerts

All lists, document libraries and discussion forums expose RSS feeds, which is a standard notification protocol. Users can subscribe with their favourite RSS feed reader (readers include Internet Explorer, Outlook 2007 and Mozilla Thunderbird) and receive notifications about updates without having to connect to the collaboration workspace. RSS feeds combined with web discussion forums are an interesting alternative to newsgroups (figure 6).

Users can also register to receive a mail alert when something is changed in a list or document library that interests them.

Work with people outside CERN

In order to provide collaboration features, collaboration workspaces need to identify individual users. CERN users can of course authenticate themselves with their computer account, but external users can also authenticate themselves by attaching a password to their e-mail address: this is the "CERN external account" option.

The same "external account" registration is shared with other CERN applications such as Indico, CDS and the Simba mailing-list service, so that external users can use a single password to access any of these CERN applications as well as collaboration workspaces.New users can register at http://cern.ch/Externals.

When configuring the access permissions for a collaboration workspace, you can specify external users using their e-mail address or with Simba mailing lists (go to http://cern.ch/simba).

Getting started with collaboration workspaces

Collaboration workspaces are created and managed in the same way as other types of website, from the central Web Services site at http://cern.ch/web (figure 7).

Read the guide "Getting started with SharePoint" (located in the section "Working with collaboration workspaces" on the Web Services site) along with the SharePoint online help. Support is provided as usual by the computing helpdesk at web.support@cern.ch. Seminars will be held after the summer period; watch the *CERN Bulletin* for announcements.

Alexandre Lossent, IT/IS

The deadline for submissions to the next issue of the CNL is

10 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

CERN is host to ROOT workshop

The seventh international ROOT workshop was held from 26–28 March at CERN. About 100 people attended, comprising primarily particle physicists but also scientists from other fields and some representatives from commercial companies. The ROOT workshops are typically held every 18 months and previous events have been held at Fermilab, SLAC and CERN.

The aim of the workshop is for ROOT developers to present their latest work, and for users to report on how they are using the system and to tell the developers what new features they would like and what can be improved.

The programme consisted of 38 talks and seven posters. Most talks lasted 20–30 minutes, apart from the keynote talks from the ROOT developers.

The main topics were: the use of ROOT as a general framework, new features in the input–output system, 3D graphics, graphical user interface, progress with the merging of Reflex, C++ introspection, library with CINT, the C++ interpreter, and the status of the different language bindings like Python and Ruby.

Several talks were given about the new maths libraries and statistical tools, which will become important now that the LHC is reaching completion and the first data will soon have to be analysed. There were also presentations on distributed data analysis, PROOF, and the different Analysis Object Data (AOD) models of the LHC experiments.

The talks were generally of a high quality with detailed technical content. It is impossible to cover all of the talks in this article so I will mention some of the highlights. The full workshop agenda and all talks are available from the web http://indico.cern.ch/conferenceDisplay.py?confId=13356.

Around 100 people, from science and business, participated in the event for ROOT users.

Lightweight ROOT distribution

On the first day, René Brun gave a talk entitled "From ROOT to BOOT, from BOOT to ROOT". He outlined plans to make the ROOT distribution process more lightweight based on a small and stable BOOT kernel that at run-time will import, via the web, only the libraries that the user needs. To achieve this ROOT must be modularized into more components, and these components must be made smaller by reducing the high overhead of the CINT dictionaries and improving the C++ compiler performance to allow on-demand and on-the-fly compilation of the needed components. This development is driven by the observation that during a typical ROOT session a typical user needs only a few percent of the more than 1 million lines of ROOT code. These few thousand lines of code can then be custom compiled, using the compiler optimizations, to yield the best performance for the user's platform.

Statistical tools

On the final day there were several talks about the statistical analysis tools and frameworks that are integrated with ROOT.

- First there was a talk on the Toolkit for Multivariate Data Analysis (TMVA), which implements sophisticated classification algorithms based on neural networks. The toolkit is a welcome development as these machine-learning techniques to analyse high-energy physics data are being used more and more.
- Then there was an update on the RooFit package, which is an advanced data modelling language. RooFit represents modelling concepts such as observables, parameters and probability density functions as C++ objects, and provides operator objects for addition, multiplication, convolution and so on to build data models of arbitrary complexity.
- There was also a presentation on the RooStat package, which is built on top of RooFit. It provides a system that will enable the combining of the results of multiple measurements, including systematic uncertainties, from different LHC experiments. The systematic uncertainties will be evaluated with different techniques ranging from Bayesian to fully Frequentist.

Distributed data analysis

Finally there were several talks about distributed data analysis and the AOD models of the LHC experiments. A presentation of the latest developments on the Parallel ROOT Facility (PROOF) indicated that a new workload distribution method (packetizer) showed clear performance gains, especially with remotely accessed data.

Representatives of the ALICE experiment gave a presentation that showed the promising results of using PROOF to promptly analyse physics data on the CERN Analysis Facility (CAF). With PROOF the data is analysed in parallel on many CPUs, and this yields a higher throughput at a high level of efficiency than when the same job is split into many smaller jobs and submitted via a traditional batch system. The ALICE AOD model is tuned for use in stand-alone ROOT and PROOF. CMS is developing an AOD that can be used either in the full CMS Software (CMSSW) framework, or in stand-alone ROOT and PROOF. LHCb's AOD model is based on exclusive use via their Gaudi framework and to be run on a cluster or Grid via Ganga.

Taking ROOT to the next level

Overall this was an interesting ROOT workshop. It covered many topics and there was an emphasis on analysing LHC data in advance of the LHC start-up. During the question and answer sessions the ROOT team asked the audience if any areas needed to be developed further or what functionality was missing. At previous workshops there would be many suggestions for new functionality but this time the audience did not identify any major items that were missing. It is good to know that ROOT now covers the data-handling and analysis needs of the LHC users. Of course this does not mean that ROOT is finished. It is now up to the ROOT team to innovate and take the system to the next level.

Fons Rademakers, PH/SFT

Procedure for leaving

Members of staff who are about to leave CERN need to complete several computing tasks before they go. Here are some guidelines for what to do.

Register private e-mail address

We recommend that users who have a computer and an external e-mail provider at home register their private e-mail address before they leave CERN. The private e-mail address (or physical e-mail address, PEM) will replace the CERN e-mail for all further contact with CERN. The PEM can be registered in advance (using the field "Deferred Physical E-mail") and must be done at least before the end of the contract date. Users should simply connect to Computing Resources Administration (CRA) through http://cra.cern.ch. On the "My Details" page they should enter the PEM in the field "Deferred Physical E-mail", then in the "Date As Of" box complete the effective date as the contract end date, and finally click "Update".

Should users have any difficulties with CRA they can contact helpdesk@cern.ch (tel. 767 88 88).

Automatic account closure

We would also like to remind you about the rules regarding the automatic closure of accounts when users leave CERN. Depending on the individual, two leaving dates might be registered on the Human Resources database and they are not necessarily the same:

- the "last working day" (or last "physical date"); and
- the "end of contract".

These two dates can differ, especially in the case of early retirement or users who have accumulated lots of holiday.

The rules regarding the blocking of computer accounts when users leave CERN are as follows:

- all accounts (including NICE) are blocked on the "last working day" plus two months;
- the AIS and MAILSERV accounts are blocked only at the "end of contract" plus two months (this extension allows time for pay slips to be generated).

Users are notified about two months before either of the above actions are taken. If the "last working day" and "end of contract" are different, users will receive two messages. In this case they should read the message carefully to understand which account will be blocked and when.

CERN external account

This year, at the request of the finance department, we have made it possible for CERN pensioners to obtain their annual internal tax certificate via the web. To do this they have to create a CERN external computer account, which will enable them to access the CERN website. For the time being only former staff members require this external account to access their annual internal tax certificate.

An external account can be created using the PEM that is registered with CRA. First the user should connect to http://cern.ch/Externals (this url can be accessed from outside CERN); then click on "Online Registration" and complete the fields under "Account Information". The "E-mail" field must contain the same PEM as was recorded with CRA. Click "Register" to confirm.

To validate this account the member of staff will receive an e-mail that enables him/her to register a password. The user will need both the PEM and the password to access the new external account.

When the CERN external account has been created, users can access their internal annual tax certificate at http://aismedia-ext.cern.ch/aismedia/taxcert. The full procedure is described on the Human Resources website (at http://cern.ch/HumanResources/internal/generalinfo/impots/taxes/pages/Taxes-attestation-imposition. asp). Staff who do not have a home computer or an e-mail address can still obtain the tax certificate by post, by contacting the *bureau des salaires* in the finance department.

Nicole Crémel, IT/UDS; Nick Zogias, IT/AIS

IOP relaunches CERN Courier, CNL website

For over two years the *CNL* editorial team has collaborated with IOP Publishing (IOP) on the production of *CNL* as well as the bimonthly Computing News section in *CERN Courier*. This summer IOP will launch a new website for *CERN Courier* (still at http://cerncourier.com), which will include improvements to the *CNL* website.

The most significant changes on the *CNL* part of the site will be a user-friendly menu for browsing through the *CNL* archive and a more intuitive link for downloading a PDF of each issue. An RSS feed option for receiving a list of the contents of *CNL* will be made available, and *CNL* will appear on the *CERN Courier* navigation bar.

A new search function will make it easier for readers to find relevant content, from the entire *CERN Courier* site or by specifying a subset (such as *CNL*, Jobs or Products).

CNL editors

Calendar

June

25–28 **WorldComp'07**
Las Vegas, Nevada, USA
www.worldacademyofscience.org/worldcomp07/ws

25–29 **ICDCS 2007**
Toronto, Canada
www.eecg.utoronto.ca/icdcs07

26–29 **ISC2007**
Dresden, Germany
www.supercomp.de

27–29 **HPDC 07**
Monterey Bay California, USA
www.hpdc.org

July

5–7 **International Symposium on Parallel and Distributed Computing**
Hagenberg, Austria
www.gup.jku.at/ispdc

9–12 **2007 International Multi-Conference in Computer Science, Engineering, and Information Science**
Orlando, Florida, USA
www.PromoteResearch.org

22–26 **AAAI-07**
Vancouver, Canada
www.aaai.org

August

12–16 **International Wireless Communications and Mobile Computing Conference**
Honolulu, Hawaii, USA
http://dropzone.tamu.edu/~xizhang/IWCMC07/IWCMC07.htm

28–31 **Euro-Par 2007**
Rennes, France
http://europar2007.irisa.fr

29–31 **VIIP 2007**
Palma de Mallorca, Spain
www.iasted.org/conferences/cfp-583.html?viip

September

2–7 **CHEP'07**
Victoria, British Columbia, Canada
www.chep2007.com

October

1–5 **EGEE'07**
Budapest, Hungary
www.eu-egee.org/egee07

November

5–9 **HEPiX 2007 (autumn meeting)**
St. Louis, Missouri, USA
http://genome.wustl.edu/hepix2007