

CERN COMPUTER NEWSLETTER

Volume XL, issue 2 April–May 2005

Contents

Editorial

IT Department plans busy year 1



IT Department management at the programme of work meeting.

Announcements & news

Computing featured in this month's *CERN Courier* 2

Openlab calls for students to work on Grid projects 2

Personal use of Internet Relay Chat is banned 2

CERN Linux migrates to SLC3 3

CERN School of Computing to be held in France 3

Desktop computing

Why folder redirections should replace roaming user profiles 4

Shadow copies for restoring files 5

Technical briefs

Wiki eases Web collaboration 7

Latest news on fight against spam 8

Information corner

Questions and answers from the Helpdesk 9

Recent acquisitions at IT bookshop 9

Recent changes to IT services 10

New release of LaTeX at CERN 10

WBTechT offers online training 10

Comedy corner: sorting facts from user imagination 10

Calendar 10

Editors Nicole Crémel and Hannelore Hämmerle, CERN IT Department, 1211 Geneva 23, Switzerland. E-mail: cnl.editor@cern.ch. Fax: +41 (22) 7677155. Web: cerncourier.com (link CNL).

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

Produced for CERN by Institute of Physics Publishing

Dirac House, Temple Back, Bristol BS1 6BE, UK. Tel: +44 (0)117 929 7481. E-mail: chris.thomas@iop.org. Fax: +44 (0)117 920 0733. Web: iop.org.

Published by CERN IT Department



IOP

©2005 CERN

The contents of this newsletter do not necessarily represent the views of CERN management.

IT Department plans busy year

The IT Department managers spent three days in February/March planning the department's programme of work for 2005. The head of the department, Wolfgang von Rüdén, talked to CNL about the plans for the coming year and beyond.

You have made IT security a top priority for 2005. What concrete actions will this result in?

A reorganization within the department towards the end of last year allowed us to add six staff to the security activities. IT security is an issue that affects all departments, but the most urgent area to be addressed is the network for the technical services, the accelerator and the experiments' systems at the pits. For this, the interdepartmental Computing and Network Infrastructure for Controls (CNIC) project was set up last year. It is coordinated by Uwe Epting from TS and reports via the Controls Board to me. IT contributes the required tools and the infrastructure, but all departments concerned will have to provide their own effort to implement the solutions.

An additional four people were funded jointly by the DG, AB and IT. Two people will work on global-security issues such as connectivity management – providing controlled access to resources and protecting machines from attacks from infected computers. Another two will work on NiceFC and LinuxFC (for controls), providing tools and procedures to manage Windows and Linux machines in production environments. In parallel, we are pursuing options to strengthen our on-site security with the support

of industry and the EU, which should lead to new projects in this area during 2005 and beyond. We have identified site security as a weak element in Grid security, so this effort is complementary to the security work being done by the LHC Computing Grid (LCG) and Enabling Grids for E-science (EGEE) projects.

How is the department balancing the ever-increasing expectations from users, with tight budgets and recruitment limitations?

This is always a difficult issue. We have a long list of requests from users to support many types of hardware and software, and we have to continuously evaluate where we can apply our limited resources most effectively. We continue to work with the user community in 2005 to define what is "nice to have" and "need to have" on their wish lists. With the LHC start-up just around the corner, the priority must be those services that are critical for the success of the project. Providing new services generally means phasing out old ones – for example, we will phase out SUNDEV in 2005. It is also important to realize that supporting, for example, an operating system means supporting three versions of it – previous, current and next. Adding a new platform is therefore a heavy investment. Shortening phase-in and phase-out times of some services may be one way to free up resources, but we depend on our users to migrate their applications. While some always insist on the latest version of the OS and the compilers, others want stability

Announcements & news

first of all. Such contradictory requests often come from the same experiment. The compromise we have to make sometimes is to delay the upgrade of services. On the positive side, open-source projects, such as CDSWare and InDiCo, have managed to get substantial external support. This is a positive response to internal budget limitations, which we will surely encourage where relevant.

The department has quite a few externally funded activities, both through EU projects and industry partnership. How will this evolve? 2005 is an important year for the major Grid projects that the department is managing. The EGEE project was successfully reviewed in February, and we are confident that the EU will fund a second phase of the project starting in April next year, provided we submit a strong proposal in September. The CERN openlab industrial partnership will complete its first three-year phase at the end of the year, and we expect to define the future for this successful collaboration with the partners in early July, at the annual openlab Board of Sponsors

meeting. The LCG project has received voluntary contributions from several EU member states, and discussions are ongoing for the funding of the second phase of this project. The IT Department hired more than 90 people last year, mostly owing to external funding from projects like EGEE. It is great to get an influx of young, highly motivated staff this way, although having so many people on short-term contracts is a challenge on the human resources planning front and a very heavy training load.

Looking further ahead, what major changes do you see on the horizon for the IT Department and the services it provides?

Well, our major goal is to get LHC off the ground, which keeps us more than busy. Although we have just passed the latest data challenge, reaching 500 megabytes/s to tape sustained during one week, we have to go about eight times higher to reach the level needed when LHC will be in full swing. This is just one area – many others face similar challenges. So, there will be no major changes in the near future unless they are needed to reach the goal.

The department is always investing a small effort in a number of IT areas where changes are anticipated. For example, this year we will be evaluating different computer platforms for the LCG, to identify the best cost:performance options. Public computing, such as the department's LHC@home screensaver, provides essentially free CPU power and is in many ways complementary to the Grid, so we will be investigating other HEP-related programmes that could benefit from this.

In the area of communications technology we are planning for a major upgrade of the campus network, both at the core routing and the star-point level where the ageing equipment needs to be replaced. The external networking will probably see a more drastic change because we will need to ramp up to very high bandwidth to the Tier1 centres. Certificate-based security, inspired by the work going on in the Grid community, is also being studied for more general applications. Again, the list of interesting technologies is almost endless, and the challenge is to identify those that are most valuable for CERN.

Openlab calls for students to work on Grid projects

For the third year in a row, CERN's IT Department will be hosting 12 students this summer to work on Grid-related projects. The openlab programme, which was sponsored by Hewlett-Packard (HP) and Enterasys last year, has been a great success, enabling bachelor, masters and even PhD students to work on cutting-edge hardware and software projects in the area of Grid computing.

The scheme aims to build long-term relationships with the students, many of whom have returned to do bachelor or masters projects, or have worked on related projects after returning to their home institute.

An emphasis on team-based work, and special study tours to places such as IBM Zurich Labs and HP Grenoble, are part of what makes the openlab programme a hit with the students. The scheme is also open to students from non-member states, the only requirement being that the home institute should cover half of the costs.

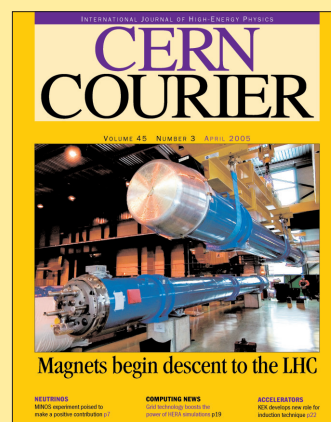
To find out more about the programme and some of the student projects planned, see <http://cern.ch/openlab-mu-internal>. Candidates should send their CV with a supporting letter from a home-institute supervisor to francois.grey@cern.ch to arrive by 30 April.

Computing featured in this month's CERN Courier

The articles listed below have been published in the April 2005 issue of *CERN Courier*. Full-text articles can also be found in HTML format on the *CERN Courier* website at www.cerncourier.com, together with the remainder of the issue's contents.

Computing News

- UNOSAT tackles tsunami challenge
- W3C's character model promotes universal access
- Semantic Web boosts access to information
- DOAR opens the way to research data



● GridX1 project joins LHC Computing Grid

● NSF release backs cyberinfrastructure

● @home with Einstein and the LHC

IT products and calendar

Feature article

● DESY becomes hub for Grid-based HERA Events

Viewpoint

In need of the human touch

Software development is more than engineering and still needs the human factor to be successful, says Federico Carminati.

● Tim Berners-Lee takes up chair...

● ...while David Williams gets professorship

Personal use of Internet Relay Chat is banned

Owing to the many computer-security incidents at CERN where intruders have used Internet Relay Chat (IRC) software to control computers, IRC connections have been blocked in the CERN firewall. Personal use of IRC software is not permitted at CERN.

Details of the restrictions on personal and professional use of IRC are documented at <http://cern.ch/security/IRC>.

CERN Security Team
(computer.security@cern.ch)

CERN Linux migrates to SLC3

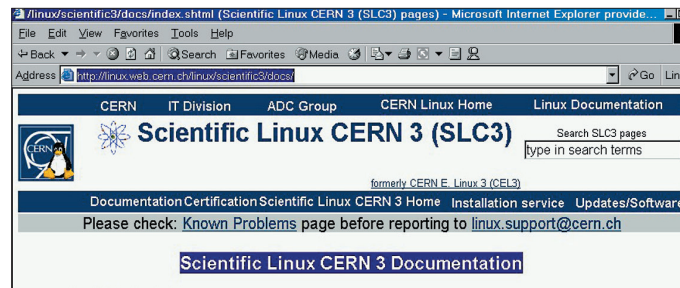
LXPLUS migration

As most CERN users will know, all the central Linux services (LXPLUS and LXBATCH machines) were migrated to Scientific Linux CERN 3 (SLC3) in January. SLC3 is a Linux distribution build within the framework of Scientific Linux, which in turn is rebuilt from freely available Red Hat Enterprise Linux 3 product sources under terms and conditions of this product's End-User Licence Agreement. SLC3 was certified by representatives from all major CERN Linux user communities for general use at CERN on 1 November 2004.

The migration to SLC3 was mostly successful. We are aware of some issues with SLC3 on LXPLUS and LXBATCH, which are fully described at <http://cern.ch/plus/issues.html>.

In particular, this page lists software and commands previously available on LXPLUS that are now being replaced by official SLC3 tools or software. The LXPLUS administrators (in IT/FIO) may consider adding software if it is requested by many users, providing the software is in the SLC3 repository and the Linux Support team (in IT/ADC) agrees to it. Such software will be installed as is, and if it causes any problems it will be removed.

Another difference that users will have noticed as soon as they connected to the new LXPLUS service is that the window manager is no longer the Motif-



based fwm2 (which is not supported on SLC3) but instead one of the more standard KDE or GNOME window managers. Users will get GNOME by default, but they can switch to KDE at login if they wish.

General 7.3 phase-out at CERN

After this successful switch of central Linux services to SLC3, the CERN Linux team strongly recommends that all Linux machines are upgraded to this version. Where possible this should be done by April, as discussed in the Desktop Forum meeting on 28 October 2004. By the time this CNL is published, machine administrators will have been contacted via automated mails, and migration is well under way in some of CERN's Linux user communities.

It is necessary to upgrade from 7.3 owing to the limited support lifetime available for that release through the outsourced support contract (progeny.com). Earlier versions are already considered a security threat because of the lack of security updates.

In most cases a full SLC3 reinstallation will be required (no update is possible), owing to the age difference between these distributions (for example, multiple changes in the package sets and different configuration file syntax). The documentation for installation is available at <http://cern.ch/linux/scientific3/docs/install.shtml>, and the Computing Helpdesk (via Linux.Support@cern.ch) may be contacted for assistance. Support for installing multiple machines with the same configuration is also provided.

Machines running 7.3 that cannot be upgraded by April must be upgraded before the end of 2005 (this is a hard end-date). Security updates will be unavailable after that date.

If a 7.3 system cannot be upgraded within the above timeframe, it will be considered a security threat because of the lack of security updates. Additional protection, such as a firewall or other means of access control, will need to be implemented until the system can be upgraded.

For machines running even older CERN (for example, 6.1.1 and 7.2.x) or unsupported Linux vendor releases, a supported version needs to be installed urgently. These machines need to be treated as a security threat since they have not received any updates (security or otherwise) from IT since 1 July 2003. Such machines may already have been compromised given the long exposure, and many break-in attempts have already been detected on the CERN site (more than 300 computers were compromised during 2004 compared with around 30 in the previous years).

Further information

- News for change of default operating system on LXPLUS at <http://cern.ch/it-div/news/LXPLUS-SLC3-050113.htm>.
 - Desktop Forum minutes: Meeting on 28 October 2004 at <http://agenda.cern.ch/fullAgenda.php?id=a044567>. Meeting on 25 November 2004 at <http://agenda.cern.ch/fullAgenda.php?id=a045119>.
 - "LINUX upgrade required to keep operating systems up to date" in CNL November–December 2004, at <http://cern.ch/cnlart/2004/002/8>.
 - LXPLUS and LXBATCH running SLC3 at <http://cern.ch/plus/SLC3.html>.
 - CERN home page on SLC3 at <http://cern.ch/linux/scientific3>.
- Nicole Crémel (IT/UDS), and Jan Iven (IT/ADC) for the Linux Support team**

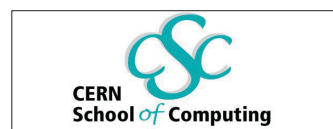
CERN School of Computing to be held in France

The 27th CERN School of Computing (CSC2005) will be held in Saint Malo, France, from Sunday 4 to Saturday 17 September. It will be organized by CERN and the Commissariat à l'Énergie Atomique (CEA)/Saclay.

The computing school is aimed at postgraduate students and research workers who have a few years' experience in particle physics, computing or related fields.

The main themes this year are:

- The GRID track, to deliver unique theoretical and hands-on



education on some of the most advanced GRID topics.

- The Software track, to address the most relevant modern techniques and tools for large-scale distributed software development and handling, as well as for computer security. The lecturers will come from the US and CERN to teach theory

and organize practical work, as well as lead discussions on lessons learned.

- The Physics Computing track, to focus on informatics topics specific to the HEP community. After setting-the-scene lectures, it will address experiment simulation and visualization. Physics Computing will also offer a range of practical exercises.

Grants are available from the EU's Sixth Framework (FP6) programme. These may cover part or all of the school fees and

living expenses, as well as part or all of travel expenses.

The closing date for full applications to arrive at CERN is 15 May. You can apply via the Web at http://cern.ch/CSC/2005/How_to_Apply_2005/How-to-apply.asp.

The CERN School of Computing website can be found at www.cern.ch/CSC/, and you will find more information about this year's event at http://cern.ch/CSC/2005/This_year_school/This_year_school.asp.

Jackie Franco-Turner, IT/DI

Why folder redirections should replace roaming user profiles

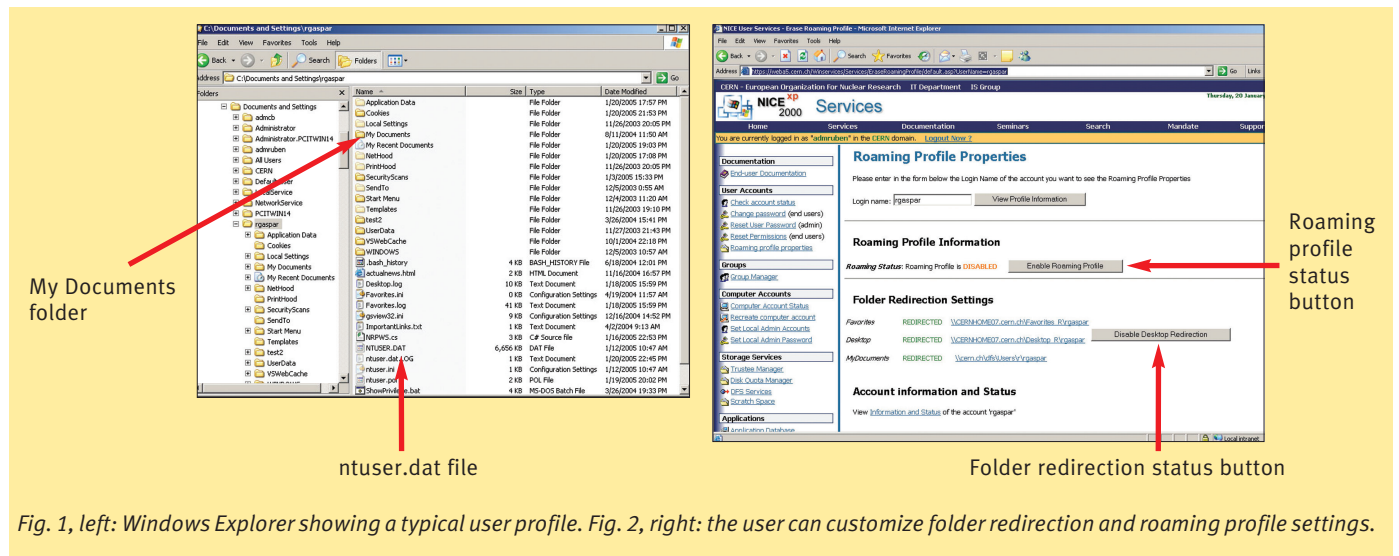


Fig. 1, left: Windows Explorer showing a typical user profile. Fig. 2, right: the user can customize folder redirection and roaming profile settings.

What is a user profile?

A user profile is a set of files and folders that stores a user's personal preferences for his or her Windows workstation. Everything that is user-specific about Windows or the applications running on the workstation is stored here.

By default, the Windows 2000 and Windows XP platforms store user profiles at %SystemRoot%\Documents and Settings\%Username% (where SystemRoot is the system drive, normally C:).

In this location, the Documents portion of the user profile is sorted into categories such as My Documents (the default location where many Windows applications such as Microsoft Office store user-produced documents); the Favorites folder, which contains links providing fast access to Web pages or documents; the Desktop folder, which holds the contents of the user's desktop; and Application Data, which contains application-specific documents.

Note that Application Data is a hidden folder, and although we do not recommend that you normally work with this setting enabled, you can view the hidden folders on your workstation by clicking

Start\Run\Explorer.exe\Tools\Folder Options\File Types>Show Hidden Files and Folders.

The Settings portion of the user profile is stored in the ntuser.dat file. Windows writes changes made during your session to this file, such as your favourite screensaver or any modifications made to your Outlook application. This file contains the HKEY_CURRENT_USER part of the Windows registry.

An example of a user profile for the user "rgaspar" can be seen in figure 1. In this case, the Desktop and Favorites folders are not listed. This is explained in the Folder Redirection section of this article.

User profiles come in three flavours: local, roaming and mandatory.

A local profile is only kept on the machine locally. This means that modifications to the profile will not be propagated either to a server or to any other machines on which you initiate a Windows session. This is the default Windows behaviour (for example, if you have a computer at home, the profiles will be left on the local disk). It also means that it is the user's responsibility to take care (by backing up) of documents that might be lost in

the case of disk failure or computer reinstallation.

A roaming profile is a profile that is copied to a central server at log-off and copied back onto the machine at log-on. In this configuration, the profile follows the user to any workstation he or she logs onto and it can be backed up on the central server.

Mandatory profiles are profiles that are copied from a central server at every log-on, but never copied back. User modifications are lost between sessions. This is a special configuration that can be used, for example, in computers available for public use on which the settings need to be restored at each start-up.

Issues with roaming profiles

The following problems were detected for users working with roaming profiles:

- Long log-on/log-off times because the profiles were copied in their entirety at log-on, and again at log-off. For example, we had one user with a profile larger than 9 gigabytes who could no longer properly log on or off from the system.
- Due to this upload/download mechanism we had situations in which users having simultaneous sessions could overwrite a final

version of a document.

- Users misunderstanding which folders are backed up on the server side and which remain local to the workstation (not all folders in a roaming profile roam to the server e.g. TEMP, Local Settings, History and Temporary Internet Files remain locally on the user's desktop).

If a user does not regularly log off they will not benefit from roaming profile features, because the profile will behave more like a local profile.

At CERN we have been using roaming profiles for several years. However, since June 2004 our policy has changed, and we now suggest using local profiles to overcome these issues.

Folder redirection instead of roaming profiles

The concept of folder redirection has already been used for many years with regard to the My Documents folder. All user accounts at CERN have this folder of their profile redirected to their NICE Home Directory in the CERN Distributed File System (DFS) share. Having a folder redirected retains the advantages of the roaming profile but eliminates its drawbacks. For instance:

- No need to upload/download

documents because they are stored on the server. The log-on/log-off time is faster.

- User data (such as the Home Directory) are available from any computer.
- Offline files will be available to work on when the user is disconnected from the CERN network (at home, travelling, etc). More information on the WinServices site is available at <http://cern.ch/win/docs/PortableComputers/#h-04.4>.
- Increased security and availability of user data – as the folder is kept on a server, it benefits from automatic backups even if the user never logs off.
- In the case of local disk failure or reinstallation, the user's documents are safe.

From the end of September 2004, we started to progressively apply folder redirection to the Desktop and Favorites folders, in addition to My Documents. By the end of 2004, all user accounts at CERN had had all three folder redirections applied. We also

changed the process by which new accounts at CERN are created. Now newcomers will benefit from folder redirection on a local profile.

Note that the My Documents, Desktop and Favorites folders are automatically made available offline. However, the My Documents folder is not automatically available offline in Windows 2000 SP3 or higher. You may want to configure the offline settings for this folder yourself – if so, please have a look at <http://cern.ch/win/docs/PortableComputers/#h-04.4>.

Also note that the Desktop folder that was formerly located at C:\Documents and Settings\%username% is now located at \\%your_homeserver%.cern.ch\Desktop_%first_letter_of_username%\%Username%, e.g. the user rgaspar can find the folder at \\CERNHOME07.cern.ch\Desktop_R\rgaspar.

Similarly, the Favorites folder is now located at \\%your_homeserver%.cern.ch\Favorites_

%first_letter_of_username%\Username%. For example, for the user rgaspar the location would be \\CERNHOME07.cern.ch\Favorites_R\rgaspar.

How can I customize this new configuration?

End-users can configure their own folder redirection and roaming profile behaviour at the WinServices website at <http://cern.ch/Win/Services/EraseRoamingProfile>. Access to account information and status is provided (figure 2 on p4).

Users really needing a roaming profile can enable this from the same place on the WinServices website.

What about my disk quota?

When more and more information is being stored on your Home Server, you may need to increase your quota on the server side so that you can accept more documents. You can carry out this operation on the WinServices site at <http://cern.ch/Win/Services/>

QuotaManager/.

Disabling roaming profiles

The final stage of the folder-redirection deployment was to disable progressively the roaming-profile attribute of entire CERN accounts. This phase of the project was completed in mid-February.

Parallel to this deployment, we e-mailed everyone who was still using the Outlook 2000 mail client (not supported since June 2004) to advise them that they should upgrade. Outlook 2000 stores user-related data (contacts, calendar items and tasks) on the client side. This information will be lost in the case of hardware failure, computer replacement or new installation.

It is highly recommended that such users migrate to Windows XP and Office 2003. More information on this topic can be found on the WinServices site at <http://cern.ch/Win/docs/Outlook/#PST>.

Ruben Gaspar, IT/IS

Shadow copies for restoring files

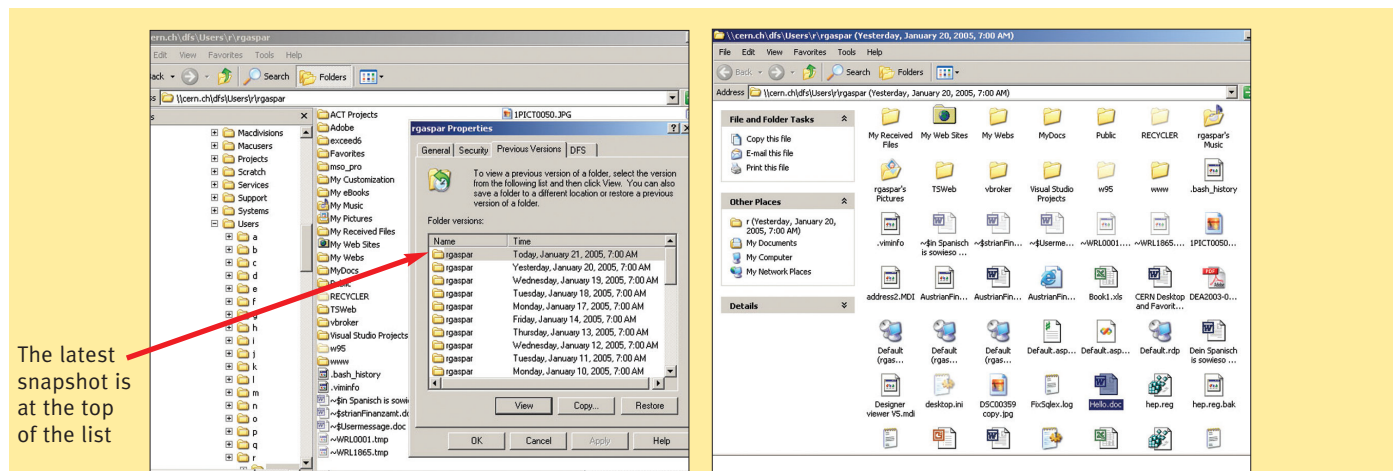


Fig. 1, left: shadow copies of your files are listed under the Previous Versions tab. Fig. 2, right: inside a previous My Documents folder.

At the beginning of January 2005, following the upgrade of Home Directories to the Windows 2003 Server, we introduced a new service to our user community. This service allows users to handle the recovery of deleted files or previous file versions on network shares without the need to

involve the Helpdesk. In particular, users are now able to:

- Recover files that were accidentally deleted, for instance in a network Recycle Bin. Everybody is now able to recover a deleted file on a network share, simply by taking it back out of the Recycle Bin.
- Recover from accidentally

overwriting a file by restoring a previous version of it.

All this is made possible by shadow copy technology.

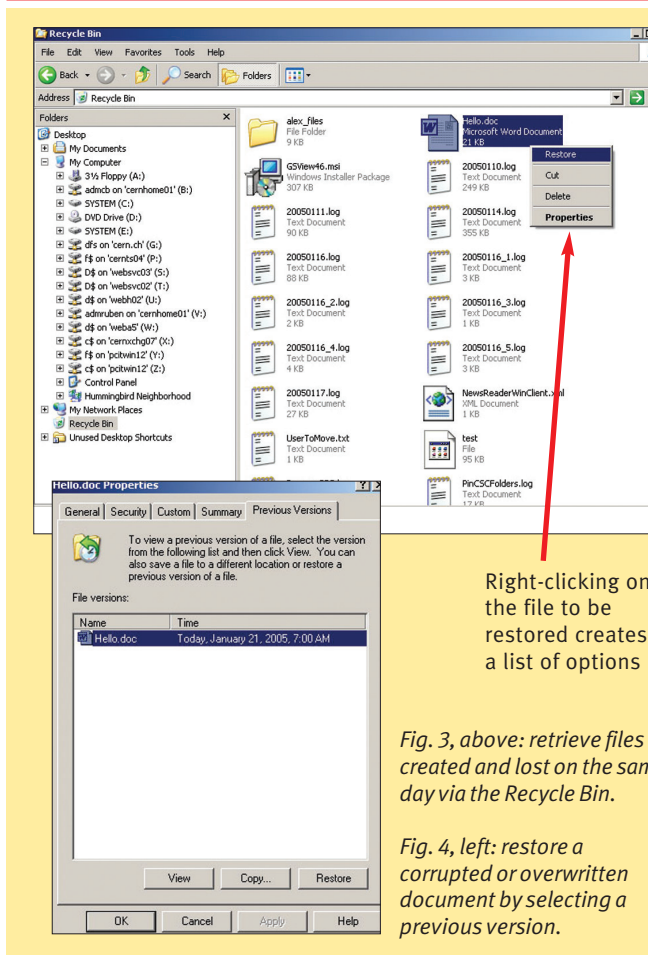
How do I recover a deleted file?

Imagine that the user rgaspar has deleted, in error, a file called Hello.doc from the My Documents folder. The My

Documents folder is redirected to a network location called cern.ch\dfs\users\r\rgaspar (your My Documents folder will always be a redirected folder pointing to a network share in DFS at \\cern.ch\dfs\users\%first_letter_of_username%\Username%).

In this example, the user

Desktop computing



Right-clicking on the file to be restored creates a list of options

Fig. 3, above: retrieve files created and lost on the same day via the Recycle Bin.

Fig. 4, left: restore a corrupted or overwritten document by selecting a previous version.

rgaspar deleted the file today, 21 January, at 8.00 p.m. He knows that the deleted file still existed yesterday. So:

- He opens Windows Explorer using Start\Run\Explorer.exe or Start\Programs\Accessories\Windows Explorer.
- He positions the cursor on a blank space and clicks the right mouse button. He selects the Properties tab and then the Previous Versions tab (see figure 1 on p5).
- He sees a list of available “snapshots” (shadow copies) of the Home Directory volume. Snapshots are taken at 7.00 a.m. daily. Users can restore the latest version of a file, or one of several prior versions. In this instance, rgaspar could open the volume from yesterday, 20 January, at 7.00 a.m. or from today, 21 January, at 7.00 a.m. As Hello.doc has not changed since yesterday morning, he picks the first version, selects View and sees the screen shown in figure 2 on p5.
- Now he selects a version of

Hello.doc from the volume, and can either drag and drop or cut and paste it to a local location.

Sticking with our example, rgaspar has also deleted a file that was created or modified today after 7.00 a.m. This document has no shadow copy. To restore it, he does as follows:

- He opens the Recycle Bin.
- He looks for the file he wants to recover (see figure 3).
- He right-clicks on it and selects Restore. The file goes directly to its previous location in the My Documents folder.

How do I recover an overwritten, corrupted or previous version of a file?

The same technique used for restoring deleted files can be used to recover previous versions of existing files, or to restore an overwritten or corrupted file. Just use the following procedure:

- Right-click on the overwritten or corrupted file and select Properties.
- Select Previous Version. Figure 4 shows that for the file

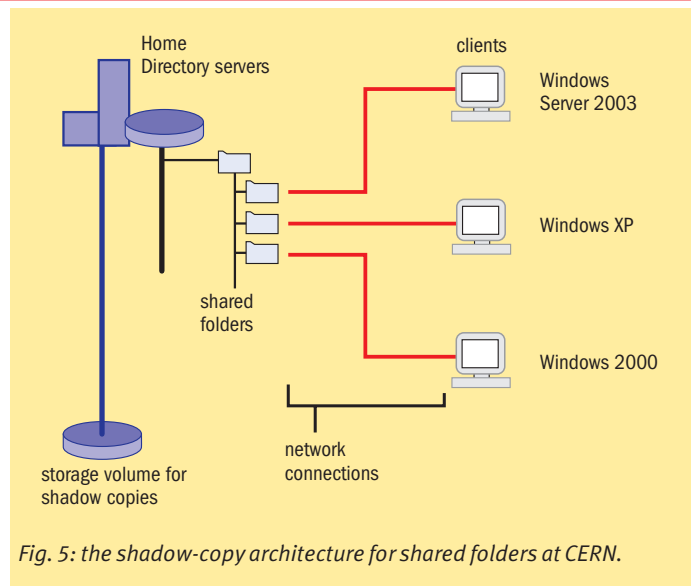


Fig. 5: the shadow-copy architecture for shared folders at CERN.

Hello.doc from the last example there is just one previous version, as the file’s content did not change until the moment it was accidentally overwritten.

- If you want to see the previous version, click View. To copy it to another location, click Copy. To replace the current version with the older one, click Restore.

What about permissions?

Permissions are always preserved. Shadow-copy technology always enforces the existing Access Control List, which defines the security descriptor of a system item, e.g. file or folder. A recovered file will contain the same permissions as the original file.

Note, however, that when you recover a deleted file, the restored file will have inherited permissions from the directory in which it is contained, as though the file was newly created.

On my computer I do not have the Previous Versions option

First of all, you must be running Windows XP, Windows 2000 SP3 or higher using the standard CERN desktop. Shadow copy is not supported on non-standard Windows installations.

Also, the Previous Versions tab is available only on network folders for which shadow copy has been enabled on the server. Currently, only NICE Home Directories offer this functionality, although there are plans to extend the service to the entire DFS file space.

Online information about the

shadow copy client can be found on the WinServices site at <http://cern.ch/Win/docs/ShadowCopyClient/>.

A technical insight: shadow-copy architecture at CERN

In the architecture chosen for shadow copies at CERN (see figure 5) three components can be identified:

- The Source Volume: the Home Directory servers are the volumes that are copied and made accessible via network shares. (A volume is a set of physical hard disks that form a logical unit or file system where information is stored – in our case, the format of our file system is NTFS).
- The Storage Volume: shadow copies, or snapshots, are kept here. Snapshots are taken once a day from Monday to Friday on our Home Directory servers.
- The shadow copy client (also called Previous Versions client): this software, on Windows XP and Windows 2000 SP3 or higher, gives users access to snapshots on the storage volume.

The number of snapshots that a Home server can hold depends on the space allocated for this purpose and also on the activity of the users who have their Home Directory there.

Once the maximum space allocated is reached, the oldest snapshot will be deleted to leave room for a new one. This is why shadow copies are not a replacement for backups.

Ruben Gaspar, IT/IS

Wiki eases Web collaboration

A Wiki (pronounced “Weekee”) is a Web collaboration platform, a Web-based post-it, a “blogging” tool or a document collection. In a Wiki you can directly edit Web pages from the browser. The original Wiki is the WikiWikiWeb of Ward Cunningham. “Wiki wiki” means “super quick” in a Hawaiian language. Wikis have become extremely popular over recent years. On the Internet you may participate in Wiki-driven forums on subjects ranging from AutoCAD development to how to stop spam. There is also a shared encyclopedia, the Wikipedia.

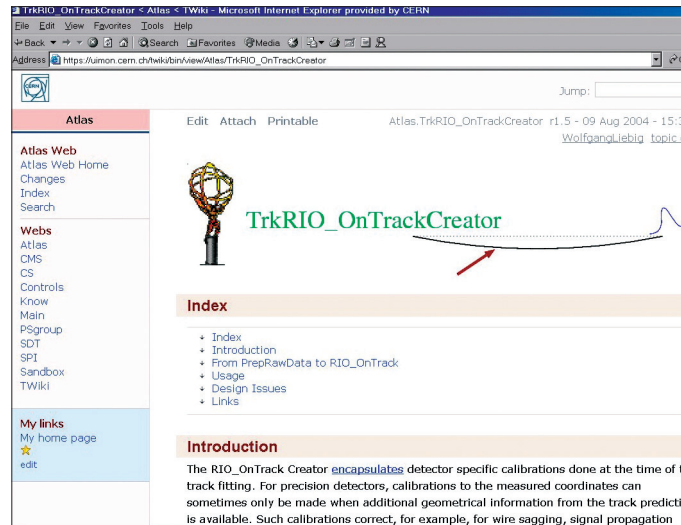
TWiki is an open-source Wiki implementation (using RCS, Perl and Apache) that is used at CERN and elsewhere to collaborate on Web documentation in software projects. The motivation for having TWiki at CERN was a user request from the Software Development Tools (SDT) community for a Wiki documentation tool.

The group also needed an internal area where it could collaborate on Web documentation and make quick project notes, informal minutes and notes about software installation. This is all live information that evolves and is available for editing by a community. The use of TWiki in this context was also inspired by the Finnish Grid team, which was using TWiki to document the activities of the Nordic Grid.

Easy to set up

TWiki enables direct editing of Web pages with revision control. The functional page layout is divided into webs and topics, with a choice of printable view for each Web page. TWiki is easy to set up and can run on standard Web servers without any extras. CERN’s SDT team has set up TWiki for intranet use with login via encrypted Kerberos authentication.

When editing documents with TWiki, hyperlinked documents are created directly on the server using simplified HTML syntax. Links and new pages are created with so-called “WikiWords”. Normal HTML is accepted as well as the upload



The OnTrackCreator Web page that was built with TWiki software.

of external documents and files. Automatic e-mail notification of changes on a subscription basis can be configured and access control is possible with groups, but restrictions are somewhat against the free, collaborative Wiki spirit. A login is needed to edit documents, and changes are traceable thanks to revision control. The system also includes a search engine.

Non-bureaucratic tool

TWiki provides powerful functionality for writing documentation “as you go along”. It is a simple and non-bureaucratic tool that can be used, for example, for developer notes and software-installation documentation. It provides an open and transparent approach, and only a Web browser is needed.

Notes can be edited and corrected without disturbing colleagues, and thanks to revision control it is possible to work on a shared document without erasing the changes made by others.

TWiki should be considered as a project-collaboration tool and not as a document-management or website-publishing system. TWiki may also be used in other areas; for example, as a shared-knowledge database for FAQs.

The use of TWiki in the SDT community has evolved into

wider use within the Atlas experiment. It is used for the online documentation of Atlas computing and Atlas data challenges. The complete installation documentation was created by a team of developers whereby the Web documentation evolves along with software development, and there is no need to interact with a Web master. TWiki is becoming increasingly popular in the Atlas computing community. An example of documentation in TWiki is OnTrackCreator (TrkRIO, see https://uimon.cern.ch/twiki/bin/view/Atlas/TrkRIO_OnTrackCreator).

Ed Moyses, a member of the Atlas computing team, said: “A major advantage of the Wiki over our conventional Web pages is that users can fix mistakes as they find them, whereas ‘normal’ Web pages can only be modified by a few selected people (who are usually too busy), and thus they quickly become out of date.”

Outside CERN, TWiki is a widely used Wiki implementation that mainly targets corporate Intranets. Some examples of TWiki installations are SAP AG (intranet for software documentation), British Telecom (internal development projects and the B2B – Business to Business – forum, see www.telcob2b.org.uk), the UK-based

Astrogrid Virtual Observatory (<http://wiki.astrogrid.org>), and the Web portal for the Helsinki Institute of Physics Technology Programme (<http://wikihp.cern.ch>).

At the moment TWiki is offered at CERN as a limited service to complement CVS and Savannah for the Web documentation of code and installation notes. Following the successful use of TWiki by the Atlas computing team, TWiki has been taken up by other groups at CERN and is being used for the Compact Muon Spectrometer computing Web documentation, and the DILIGENT (Digital Library Infrastructure on Grid Enabled Technology) project’s LCG Software Process and Infrastructure workbook. The Communications Systems group is now trying out TWiki.

The TWiki service at CERN is limited to server hosting using an AFS login (there is no extra user-registration mechanism), with a setting similar to the CVS service that is provided by IT for CVS servers. Support is limited to server maintenance (with version and security updates), and the possibility of creating new TWiki websites for major user projects.

Questions and answers about TWiki

- Q:** How is concurrent authoring handled?
A: There are edit locks to ensure that only one person at a time changes a document.
Q: What about the risk of abuse of the system?
A: In general there is a risk of abuse since in principle TWiki can be set up such that everyone can modify documents. However, the TWiki CERN application uses RCS and therefore all revisions of documents are kept and changes are fully traceable.
Q: What functionality is offered by the notification mechanism?
A: The notification works on a subscription basis and the URL that has changed is sent (not what has changed on the page).
Nils Haimyr (IT/CS), Pete Jones (IT/DES)

Latest news on fight against spam

The anti-spam architecture installed at CERN provides four levels of spam detection. Incoming mail first goes through a low-level spam filter at the mail gateway level, then through an anti-flood system. A mail is then analysed by the content-detection engine, and finally is scanned by antivirus software. Only after a mail has passed through these four filters will it be delivered into the user's mailbox.

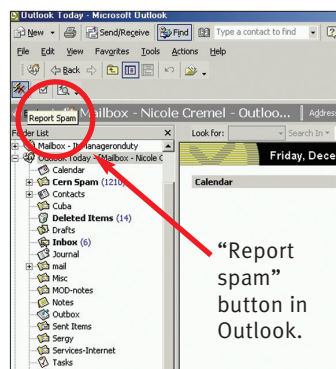
About 60 000 "good mails" are now delivered daily, compared with about 600 000 mails detected as spam. Since January, CERN's mail traffic has decreased by 50%. This may be related to legal action taken against the well known spammer Sanford Wallace around that time. However, last year it was also observed that content detection had become less accurate; regular expressions and Bayesian filters that are now used at CERN are probably well known by persons sending spam, and various techniques are used to circumvent them.

This is why it is necessary to move to a real-time spam detection system, based on regular user feedback. So we are repeating the message that appeared in the last edition of CNL: "CERN needs your help to tackle spam onslaught" (see www.cerncourier.com/articles/cnl/2/1/4/1).

Report spam

The automatic handling of spam that is reported via the Report Spam button in CERN Outlook enables the mail team to keep the rules up to date (the lifetime of spam is one or two weeks) and to extract URLs every 15 minutes in order to build a URL blacklist. This new system has been in production since January.

Please note that there is a problem accessing the Report Spam button in Outlook when connecting from outside the CERN firewall (for example, from home with a CERN laptop connected via ADSL and no VPN). Outlook tries "forever" to complete the command until it



Report Spam button in Outlook.

eventually times out. A new version is in development, which should correct this problem.

In the context of spam rejection at the gateway level, two rules were activated last year to verify a sender's authenticity. The so-called "reverse DNS" rule was activated on 15 June 2004 and checks that the remote server's IP has an associated Domain Name Server (DNS) name. Connections from anonymous servers are rejected. The external person who sent the mail is notified about this rejection, with an error message detailing the reverse DNS error and explaining which actions should be taken.

The "reverse SMTP connect" rule was activated on 15 October 2004, and verifies that the sender can receive mail. This is done by simulating a connection to the sender's domain MX (Mail Exchange). If the sender cannot receive mails the mail is rejected. After the first rule (reverse DNS) was implemented the rejection of spam increased from 55% to 85%; and after the second rule (reverse SMTP), the rejection increased again by 25%.

Emerging technologies

The next steps to fight the spam problem are to try to stop e-mail spoofing by using emerging technologies like Sender Policy Framework (SPF), Sender ID and/or DomainKeys.

- SPF was introduced by the open community. Domain owners identify in DNS their sending mail servers. SMTP

receivers verify the envelope sender address against this information and can distinguish authentic messages from forgeries before any message data is transmitted. SPF is widely used but it cannot reject mails without major modifications since forwarded mails cannot be verified.

- The Sender ID framework, developed jointly by Microsoft and industry partners, addresses a key part of the spam problem: the difficulty of verifying a sender's identity. It is based on SPF, adding header checks and addressing the problem of forwarded mails. One problem is that the status of Sender ID is still unclear due to the Microsoft patent and licensing scheme.

- DomainKeys is the spam-fighting technique proposed by Yahoo. The approach is to sign outgoing mails and to verify incoming mails: mails are signed at the gateway level, allowing later verification using a public key that can be found in the DNS. There is still a problem with mails sent to a mailing list that cannot always be verified.

DomainKeys seems the best technology today and looks promising. There is no centralized authority, there are no requirements for changes to existing protocols and the central concept is extremely easy to understand. It is already used by Yahoo (<http://antispam.yahoo.com/domainkeys>), Google Mail and several other providers of mail clients.

A DomainKeys library has been written at CERN in C# and is used with Exchange 2003 as SMTP Event Sink. It has been running on the CERN mail gateways (MX) since January. Mail rejection is now disabled until the problem with mailing lists is solved, which should happen in a few months' time. This new system is not yet very effective, but supporting it will help its worldwide development.

The CERN-written DomainKeys library is available for download under the CERN open-source licence (www.cern.ch/mmms/Antispam) to participate in the

worldwide anti-spam effort.

The last step (if the incoming mail has passed the previous filters) will be to implement some checks against local whitelists and blacklists to decide whether to accept or reject a mail. Thereby, the local default policy can be to "accept all and reject blacklist match" or to "reject all except whitelist match".

Targeting the spammers

It is likely that no consensus will be found and probably several techniques will have to be used in parallel. We all know that it is not possible to stop spam completely. However, we hope that these new steps to fight against the spam problem will help, for the following reasons:

- Increasing the costs for spammers. Correctly registered domains, correct DNS, valid SPF or DomainKeys records will cost more money to spammers. In the future they may prefer to find a cheaper method to spread messages.

- Legal action. As it will be easier to locate spammers, it will be easier to pursue them in the courts.

- Whitelists, blacklists and acceptance policy. It will be possible to identify some possible "spammer paradises" and then to blacklist all their domains and all IP addresses of that area. If the policy is to "reject all except whitelist match", unknown sources will not be able to send you mails.

Further information

Visit the following sites to find out more about the technologies mentioned in this article.

- SPF: <http://spf.pobox.com>
- Sender ID: <http://www.microsoft.com/mscorp/safety/technologies/senderid/default.aspx>
- DomainKeys: [yahoo.com/domainkeys](http://antispam.yahoo.com/domainkeys) <http://antispam.yahoo.com/domainkeys.sourceforge.net> <http://mmmservices.web.cern.ch/mmmservices/Antispam/DomainKeys.aspx>

Emmanuel Ormancey, IT/IS

Questions and answers from the Helpdesk

Edited by Nicole Crémel, IT/UDS

The User Assistance Team in IT/UDS maintains a database for Questions and Answers that have been dealt with by the Computing Helpdesk. This provides many tips on daily computing issues. You can search the database at <http://consult.cern.ch/qa>.

Below is an example of a Question and Answer (Q&A) that is frequently asked at the Helpdesk (<http://consult.cern.ch/qa/3772>).

System Administrator message for a mail I never sent.

Question

I often receive an e-mail from the System Administrator similar to the e-mail pasted below.

I do not know anybody of this name and have never sent him an e-mail. Is there anything I can/should do about it?

"From: System Administrator
To: ...
Subject: Undeliverable: Delivery Notification: Delivery has failed.

Your message did not reach some or all of the intended recipients.

Abdelhamid Halaouite on 21-Feb-05 16:40.

The e-mail account does not exist at the organization this message was sent to. Check the e-mail address, or contact the recipient directly to find out the correct address. <cernxchg17.cern.ch #5.1.1>"

or: "The message contains a content type that is not supported.

The MTS-ID of the original message is: 0ICE00GADR5Y1R@ndl1mx1-a-fixed.sancharnet.in <ndl1mx1-a-fixed.sancharnet.in (conversion-daemon) #5.600.500>"

Answer

This is a well known problem

linked to spam. Someone has used your e-mail address to send a mail to various addresses (good or bad). This is a trick used by spammers, and unfortunately the mail service can do nothing about it. After that, the mail (to the wrong address) is rejected and then you receive this message (addressed to the "From" address used by the spam mail).

You should just ignore the following type of message: "Your message did not reach some or all..." for mails that you did not send.

Related links

Regarding spam in general:
● <http://cern.ch/cnlart/2002/003/spam>
● www.cerncourier.com/articles/cnl/1/9/11/1

Other general-interest Q&As and their corresponding websites

Windows (NICE – Office) related

- <http://consult.cern.ch/qa/3671>
- <http://consult.cern.ch/qa/3724>
- <http://consult.cern.ch/qa/3729>

Favourites lost in Internet Explorer
Weekly virus-scan time schedule
Resize many pictures to equal dimensions in Windows

- <http://consult.cern.ch/qa/3743>
- <http://consult.cern.ch/qa/3751>
- <http://consult.cern.ch/qa/3775>

Creating PDF files doesn't work via Terminal Services
Increase quota on my secondary Mac directory
Creating a subfolder of an existing DFS workspace

Unix (AFS-Lxplus/Lxbatch) related

- <http://consult.cern.ch/qa/3274>
- <http://consult.cern.ch/qa/3762>
- <http://consult.cern.ch/qa/3738>

Gridftp: transfer files/data from outside institute
ccdbinfo and ccdbuser on Lxplus SLC3
Bad response time on Lxplus with SLC3: Gnome and KDE window manager
Ls command slow on Linux SLC3

- <http://consult.cern.ch/qa/3777>

Mail (Outlook-pine-mail services) related

- <http://consult.cern.ch/qa/3704>
- <http://consult.cern.ch/qa/3742>
- <http://consult.cern.ch/qa/3771>

Problem sending mail with Outlook: following recipient could not be reached or e-mail account does not exist
Pine on Lxplus (SLC3): font with accents (French) or German characters
List of e-mail addresses rejected because of the Reverse DNS Problem

Printing

- <http://consult.cern.ch/qa/3229>
- <http://consult.cern.ch/qa/2319>
- <http://consult.cern.ch/qa/3740>

Printing from a private Linux PC on CERN printers
Public plotter for drawings (A0 or A3 format)
a2ps and lpq on Lxplus (to print files)

Recent acquisitions at IT bookshop

The shop is situated in Building 513 R-047. CERN users can buy books and CDs there at discount prices. The service is open on weekdays from 8.30 a.m. to 12.30 p.m. and it can also be contacted by e-mail (Bookshop@cern.ch) or phone (74050).

The most recent acquisition in the physics field is:

● **Physics for Scientists and Engineers** by Fishbane *et al.* (third edition)

In the computing field:

- **C++ Template Metaprogramming** by Abrahams and Gurtovoy
- **C++ Coding Standards** by Sutter
- **Practical Guide to Red Hat**

Linux by Sobell (new edition)

- **Computer Networking** by Kurose
- **UML Reference Manual** by Runbaugh, Jacobson and Booch (new edition)
- **Visual Basic.Net** (Microsoft Press)
- **Linux SLC3.04** is now available in both CD and DVD format.

The complete list is integrated into the CERN library system. It is available on the CERN network at <http://cdsweb/tools/itbook.py>.

Books are purchased from some 15 publishing houses, and users are very welcome to suggest acquisitions.

Roger Woolnough and Jutta Megies, IT/UDS

The deadline for submissions to the next issue of CNL is

2 May

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>

Recent changes to IT services

Changes to services in the IT department are published on the Service Status Board (SSB)

which is located at <http://cern.ch/it-servicestatus>. The most recent changes and their dates

of posting are shown below. The SSB also includes service incidents, scheduled interventions, power cuts and the status of most services.

17 February	SLC3 and SL/RHE3 incompatibility (POOL-1.7 vs cxxabi)
10 February	accdb/cerndb1/jacowdb/rwdy connect string modification
10 February	Removal of (old) Oracle software in AFS for Linux clients (13 February)
7 February	Oracle 10g patch set 3 (10.1.0.3) available for Linux users on AFS
2 February	New batch system (lsf) shares for 2005 (3 February)
31 January	Stopping propagation of NIS/Yellow pages server discovery (14 February)

New release of LaTeX at CERN

TeXLive 2004

In January the latest release of TeXLive 2004 (TL2004) was made available at CERN. It was installed for general use on AFS for Linux and it can also be downloaded for installation onto a PC running Windows XP from DFS.

Linux

The default version of TeX on the "standard" Linux system (Scientific Linux CERN 3, SLC3) is a few years old (TeX is from 2001). It is fine for most common applications, but the more recent TL2004 distribution (compiled in November 2004) might prove more appropriate.

You can access TL2004 by adding the directory containing the binaries to your PATH

variable, as explained below.

```
If you are running a Bourne-like shell (such as sh, bash, ksh):> PATH=/afs/cern.ch/sw/XML/TL2004/bin/i386-linux:$PATH> export PATH
```

```
If you are running a C-type shell (such as csh, tcsh):> setenv PATH/afs/cern.ch/sw/XML/TL2004/bin/i386-linux:$PATH
```

Microsoft Windows XP

After verifying that you have deleted older versions of the TeX family, you will be able to install the 2004 version with the help of a Setup Wizard available in the TeXLive tree on DFS: \\cern.ch\dfs\Experiments\sw\TeXLive.

See *LaTeX at CERN* for detailed instructions (<http://cern.ch/XML/textproc.html>).

Recent publications about LaTeX

- Herbert Voß's *Math mode*

describes how LaTeX can be used to write publications that contain mathematical material (www.tug.org/tex-archive/info/math/voss/Voss-Mathmode.pdf).

- *The Comprehensive LaTeX Symbol List*. A document compiled by Scott Pakin that lists over 2800 symbols and the corresponding LaTeX commands to generate them (www.ctan.org/tex-archive/info/symbols/comprehensive/symbols-a4.pdf).
- *The LaTeX Companion* (second edition, 2004). An up-to-date description of more than 200 packages using tested examples (www.awprofessional.com/bookstore/product.asp?isbn=0201362996&rl=1). Written by Michel Goossens and Frank Mittelbach (available from the CERN IT Bookshop, <http://cdsweb.cern.ch/tools/itbook.py?subject=LaTeX>).

Michel Goossens, IT/UDS

WBTechT offers online training

The 2005 CERN Web-Based Technical Training portal is a computer-skills site offering multimedia learning. Visit

www.course-source.net/sites/cern/ to self-register, consult the available programmes or request a course via EDH.

A self-directed online course costs 50 SwFr for desktop applications and 90 SwFr for technical applications for three

months' unlimited access.

Visit the WBTechT portal or www.cern.ch/TechnicalTraining, and contact Technical.Training@cern.ch or your DTOs to find out more information. Use of the CERN WBTechT is restricted to CERN personnel.

Comedy corner: sorting facts from user imagination

Users may not be aware that Helpdesk work is not an easy task. The Computing Helpdesk is a link between the users of our complete computing infrastructure at CERN (you know how complex that is) and the service providers in the IT department. Recently we had a user who reported the following problem: "For one week my Web page hosted on AFS has not been available."

After contacting the AFS experts and service providers, who looked into all their logs for the previous week, their reply



and diagnostic was: "Yesterday and the day before there were two periods of 10 minutes when AFS was not available."

As you can imagine it is as difficult for you as it is for the Helpdesk to know who is right.

We take this opportunity to reproduce a well known story and joke on Helpdesk work, where this time it is obvious that the

user was not completely right.

User: "My monitor doesn't work."

Helpdesk: "Is it switched on?"

User: "Yes."

Helpdesk: "Okay. As a first measure, please switch it off."

User: "Oh, now it works."

However, we also know that users can be right and they sometimes notice and report problems before the Helpdesk knows about them. Life is, after all, about compromises and people helping each other!

An anecdote written by members of the User Assistance team in IT

Calendar

April

2–8 **HPC High Performance Computing Symposium**
San Diego, CA, US, www.caip.rutgers.edu/hpc2005

5–6 PVSS Users' Meeting

CERN, Switzerland, www.etm.at/cern.htm

7–9 HealthGRID 2005

Oxford, UK, <http://oxford2005.healthgrid.org>

18–22 3rd EGEE Conference

Athens, Greece, <http://eu-egee.org>

May

2–6 **GridAsia 2005** Singapore, <http://gridasia.ngp.org.sg>

9–12 Cluster Computing and Grid 2005

Cardiff, UK, www.cs.cf.ac.uk/ccgrid2005

9–12 2nd International Workshop on Grid Computing and Peer-to-Peer Systems

Singapore, <http://laurel.datsi.fi.upm.es/GPP05>

9–13 HEPiX 2005

Karlsruhe, Germany, <http://iwwwww1.fzk.de/hepix/hepix.html>

June

20–24 **Second International IEEE Symposium** Sardinia, Italy, www.globalstor.org

21–24 International Supercomputer Conference

Heidelberg, Germany, www.supercomp.de/index.php?s=default

26–29 14th Global Grid Forum

Chicago, IL, US, www.ggf.org

August

17–19 5th WSEAS International Conference on Simulation, Modeling and Optimization

Corfu, Greece, www.worldses.org/conferences/2005/corfu/smo/grid/index.html

September

18–22 **Oracle OpenWorld 2005**
San Francisco, CA, US, <http://www.oracle.com/openworld/sanfrancisco/conference/index.html>