Personal data storage

Background

In the course of time most users tend to accumulate a lot of data which they don't have the time to sort and delete *right now* or which they simply don't want to delete at all. Who knows, it might turn out to be usefull some day. For this kind of rarely accessed data we now have a dedicated storage system that every user can use for his personal data collection.

Your home directory is not such a good place for this, because home directories are very important for the whole system to properly function: one can't log in or read email without them. The more data lying around in homedirs, the slower the whole systems performs. In case of problems the ZEDV has to bring home directories back online as fast as possible. Just copying the data takes approximately eight hours (nov 2006), due to a lot of old data stored in home directories.

If you think of your home directory as being your desk, the personal data storage system is something like your office cabinet. Please move some of the less important data from your desk to your cabinet!

Access Permissions

Your personal data storage is protected in the same way your home directory is: it's only accessible by you.

How to access the system

Windows
> Using Windows the personal data storage is accessible using the network path **\\storage\\<username>** where <username> denotes your Unix login name. We recommend to assign this network path to a drive letter. For machines in the PHYSIK domain this is automatically done (letter S:).

Linux
> On all machines in the Linux cluster you personal data satorage is available at **/storage/<username>**.

SECNET
> From within the SECNET you can access the personal data storage at **\\zerberus\storage\<username>.**

It is also possible to access the data from outside the faculty via sftp. Please use the host name **storage.physik.fu-berlin.de**, your unix user name and your unix password. The path is also **/storage/<username>.**.

Operating system	Network path
Windows 2000/XP	**\\storage\<username>**
Windows im SECNET	**\\zerberus\storage\<username>**
Linux (ZEDV-Cluster)	**/storage/<username>**

Data security

Your data is stored on a 2 Terabyte filesystem on a HP Enterprise Virtual Array 5000 SAN machine. It's fully redundant and protected against multiple drive failures. Additionally your data is transfered to a second machine every night for backup. Data loss due to hardware problems is extremely unlikely.

On the other hand, there is always the danger of users *accidently* deleting data. To protect against this kind of data loss is much more difficult, as the system is unable to detect whether a file was deleted accidently or on purpose. Also we can impossibly retain copies of all deleted files, just for the case that they will be needed again some time later.

The backup scheme currently implemented ensures that a deleted file can be restored, if it already existed before the nightly backup. Files created the same day as they are deleted are not protected against accidental removal.

If a user creates a file named 'hello', deletes this file some days later and creates a new file named 'hello', the previous version can be restored. If he also deletes the new file 'hello', the first version is lost, unless he contacts the ZEDV immediately (i.e. before the nightly backup).

This kind of backup works nicely for nearly static data. It is important to not move or rename files often. So it pays to spend some time thinking about the directory structure before copying data to the system.

Exceptions from the backup

If while creating the data it is already known that these files will be deleted soon, or if the files should not be included in the nightly backup, the files should be stored in a directory named **nobackup**. It does not matter where this directory is placed: a file will not be backed up if any of its parent directories is named nobackup. These files are still protect against hardware failure, just not against user mistakes.

We ask you to use this possibility especially for extremly big files (for example images of complete disks, that should be stored for a short amount of time), as the backup concept requires that all files still remain stored somewhere for at least some time after deleting.

Costs

For the time being the space is free of charge. As soon as the current space is used up, an expasion must possibly be installed. The costs of the expansion might be charged to the groups according to the ration of the space they use.

The ZEDV reserves the right to exclude single users from the access to the personal data storage system in obvious cases of abuse.