Backup and Restore (Unix and AFS)

This page has been transferred from the previous website as-is. The information is still relevant.

Unix File System Backups at SLAC

Unix File Systems

For Unix, there are several network file systems managed by the Computing Division, but we will broadly refer to their backups as NFS/GPFS/Lustre and AFS backups. NFS/GPFS/Lustre and AFS are backed up in different ways and have different backup schedules. However, there are a few underlying policies that were followed in setting up both backup systems. **NOTE: not all file systems are backed up.** Home directories are backed up, but most other directories are opt-in only. See below for more information.

For file systems that are backed up, NFS/GPFS/Lustre and AFS backups are performed automatically on a daily basis and should be viewed primarily as a way of recovering from hardware failures, not as an archival system. This means that the backups are not retained forever. Please see the retention period information below for each backup type.

NFS/GPFS/Lustre Backups and Recovering Your Own Files

Lustre home directories in the /sdf/home file system are backed up nightly. By default, we do not backup other NFS, GPFS, or Lustre file systems unless requested to do so (opt-in). In those cases, backups run nightly and tapes must be provided by the groups who own the data. Please submit a backup request to unix-admin@slac.stanford.edu.

Those NFS/GPFS/Lustre file systems that are backed up are done so via IBM Spectrum Protect (formerly Tivoli Storage Manager or TSM) software. We currently support RHEL 6/7 and Centos 7 x86/x64 clients. Solaris SPARC/x86/x64 clients will be supported until existing hosts are retired.

To request a file restore for an NFS/GPFS/Lustre file/directory that you have access to, please send email to unix-admin@slac.stanford.edu. Include an explanation, the full path to the file/directory, and from what point in time you need a restore.

Spectrum Protect is an incremental backup system. It backs up only the files that changed since the last backup, and maintains information on the state of the client file system. It is possible to restore the file system to the last backup state, and to restore some older versions of deleted files. Spectrum Protect is not configured to restore the file system to the state it had at a specific point in time, i.e., it may not be possible to restore a directory to the way it looked 4 weeks ago at 12pm or any other particular day or time. Such a policy would use significantly more tape space since the backup server would be forced to keep a copy of every file version going back to that date.

Spectrum Protect Schedule and Retention Policy

The Spectrum Protect backup runs each night, usually starting sometime between 12:00AM-5:00AM on client machines.

Spectrum Protect maintains backup data for both active and inactive file versions. An active version of a file is the most recent backup copy of a file stored in Spectrum Protect for a file that currently exists on a file server or workstation. An active version remains active on tape and exempt from tape deletion until: 1) replaced by a new backup version or 2) Spectrum Protect detects, during an incremental backup, that the user has deleted the original file from a file server or workstation. An inactive version of a file is a copy of a backup file in Spectrum Protect that either is not the most recent version, or the corresponding original file was deleted from the client file system.

Unless otherwise stated, the STANDARD retention policy is as follows:

- Up to 31 copies of a particular file are kept on tape as long as the file exists on the client's file system.
- Only the most recently backed up version on tape is active. All other versions on tape are inactive but still recoverable.
- Once a file on tape goes inactive, it expires after 31 days and gets deleted off tape.
- If a file is deleted from a client's disk, several things happen during the next full incremental backup: 1) the active version of that file on tape will be marked inactive, 2) all inactive copies start expiring off tape as they reach 31 days of age, and 3) the last remaining inactive copy (which is also the most current backup copy) will be kept for 31 days, after which it expires too.
- Note that as long as a file remains on a client's disk, its latest backup copy will remain active on tape and not expire.

What does this all mean? Basically, while a file still exists on disk, the last 31 days of backups are also on tape. But once a file is deleted from disk, all of its copies on tape will expire as they each reach 31 days of age. So, for example, if a file is changing every day and then gets deleted, there will be 31 copies on tape. Each day, the current oldest copy will expire. After 31 days, no more copies will remain on tape.

We will notify file owners in advance if their backups are not using the STANDARD retention policy.

AFS Backups and Recovering Your Own Files

Legacy home directories are located in the AFS file system and are backed up nightly, as are most (but not all) AFS group directories. AFS backups are provided by TiBS software from Teradactyl. The unit of AFS file storage and backup is the volume. Typically, each user's home directory is a single volume. For the first level of backup, TiBS creates a copy of each volume at midnight each night. This copy is called a "backup volume". You can find this backup volume from the .backup link in each home directory. If you have just deleted or damaged a file that existed at midnight the previous night, type "cd ~/.backup" to find a version of it from the previous day and copy it back into your home directory.

Note that backup volumes are automatically "mounted" for home directory volumes only. This means that you must manually mount backup volumes for group volumes or user sub-volumes if you need to recover files from the midnight copy. To do this you will need the volume name. The easiest way to get this is to execute the "fs listquota" command on the directory in question. For example, if you had accidentally removed a file from the directory /afs/slac/g /babar/data/data01, you would type

```
> fs listquota /afs/slac/g/babar/data/data01

Volume Name                    Quota      Used %Used   Partition
g.babar.data.01               500000    223503   45%         32%
```

The first column lists the name of this volume as "g.babar.data.01". To get the name of the backup volume, append ".backup", then mount it in your home directory with the command:

```
> fs mkmount ~/bdata01 g.babar.data.01.backup
```

and reference it at ~/bdata01. (You may pick any name in place of bdata01 as long as the directory doesn't already exist.) The only privileges you need for fs mkmount are insert and administer for the directory you are mounting in (such as your home directory).

We recommend doing such mounts in your home directory to avoid creating directory "loops". For example, it is tempting to mount the .backup volume in the volume you're dealing with, because that is frequently your current directory. However, if you mount a volume's .backup volume within itself, and you leave the mount there, then tomorrow and thereafter, .backup and .backup/.backup and .backup/.backup/.backup etc. will exist. This causes real problems to recursive commands like "ls -lR", "find", and "du". We also recommend that you remove the mount when you are done with it, because you won't really like seeing it under your home directory. You can remove it with

```
> fs rmmount ~/bdata01
```

AFS Backup Schedule and Retention Policy

AFS backups are a series of full and incremental backups, designed to provide complete coverage of recent changes, and sparser coverage going back in time. A level 0 backup is a full backup of the AFS file system. A level 1 backup is an incremental backup of all changes since the previous level 0 backup. A level 2 backup is an incremental backup of all changes since the previous level 1 backup. The schedule of AFS backups is as follows:

Level 0: A full backup is performed starting at midnight on the first Sunday of each month. This backup is retained for six months. After six months, only the quarterly (January, April, July, October) backups are kept. The quarterly backups are retained for one year.

Level 1: An incremental backup is performed starting at midnight every Sunday morning (except for the first Sunday of each month). These backups are retained for two months.

Level 2: An incremental backup is performed starting at midnight Monday through Saturday. These backups are retained for two weeks.

The result of that schedule is that a volume can be retrieved from the daily backups for the first two weeks, then from the weeklies for the first two months, then from the monthlies for the first six months, and then from the quarterlies for one year.

AFS backups are not yet retrievable by users with the exception of those files that are located in the user's .backup subdirectory created around midnight. See the AFS Restore web page or send email to unix-admin@slac.stanford.edu to request the retrieval of a file from backup tape.

Related articles

- Installing YFS on Ubuntu Desktop
- Backup and Restore (Unix and AFS)
- Restoring files using TSM