

Filesystems & Disks

CIS 68C1

UNIX System Administration

The Filesystem

- Filesystem Components
 - ✗ A convention for naming & arranging things
 - ✗ A **namespace**
 - ✗ A well-defined set of programming functions
 - ✗ **API**
 - ✗ Security Model
 - ✗ The code that implements these

The Filesystem

- The UNIX file tree may be comprised of one or more separate filesystems
 - ✗ Attaching a filesystem is called **mounting**
 - ✗ The **mount** command maps a filesystem **onto** an existing directory
 - ✗ Filesystems can reside
 - ✗ on a local disk
 - ✗ on a network
 - ✗ in RAM

The Filesystem

- The / File Tree Hierarchy
 - ✗ Historically not well organized
 - ✗ Standard directories
 - ✗ Review directory structure on pg. 65 of the text
 - ✗ /bin, /sbin, /dev, /etc, /tmp
 - ✗ Typically part of the root filesystem
 - ✗ Required during boot stage
 - ✗ /lib, /usr, /tmp, /var
 - ✗ Often separate filesystems

The Filesystem

□ File Types

- ✗ Regular File
- ✗ Directory
- ✗ Symbolic Link
- ✗ Device
 - ✗ Character and Block
- ✗ Named Pipe (FIFO)
- ✗ UNIX Domain Sockets
 - ✗ Printing subsystem, X Windows, syslog facility

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

5

The Filesystem

□ Directories

- ✗ Directories provide namespace access points to files
- ✗ A file's name is stored in its parent directory
- ✗ Several names can reference a single file
- ✗ A file can be referenced by any number of directories
- ✗ Each reference is called a **link**
- ✗ Cannot create hard links to directories

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

6

The Filesystem

□ Links

- ✗ Also called hard link
- ✗ Reference to an inode
- ✗ No distinction between multiple links to a file
- ✗ Commonly used to give multiple names to a file
 - ✗ Some programs behave differently when invoked with a particular name

```
$ ls -li /usr/bin/ypch* /usr/bin/yppass*
97807 /usr/bin/ypchfn
97807 /usr/bin/ypchsh
97807 /usr/bin/yppasswd
```

inode number

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

7

The Filesystem

□ Symbolic Links

- ✗ Also called soft link or symlink
- ✗ Reference to a file by name
- ✗ Distinct from the file to which it references
- ✗ Commonly used to create alternate access points or rearrange a subtree of the filesystem

```
$ ls -ld /bin /lib /usr /usr/bin /usr/lib
lrwxrwxrwx 1 root root 9 Aug 27 1998 /bin -> ./usr/bin
lrwxrwxrwx 1 root root 9 Aug 27 1998 /lib -> ./usr/lib
drwxrwxr-x 30 root sys 1024 Jan 1 2000 /usr
drwxrwxr-x 3 root bin 7680 Sep 15 2000 /usr/bin
drwxrwxr-x 42 root bin 10240 Oct 9 20:29 /usr/lib
```

File Type Field
l=symlink d=directory

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

8

The Filesystem

□ Devices

- ✗ Live in /dev
- ✗ Are access points to a device via the filesystem namespace
 - ✗ Share a common access interface (API) with files
- ✗ Two types of devices
 - ✗ Block
 - ✗ Kernel buffers I/O to increase performance
 - ✗ Character / raw
 - ✗ Kernel does not buffer I/O
 - ✗ Some devices offer both interfaces

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

9

The Filesystem

□ Devices

- ✗ Entries in /dev are access points to their respective drivers
 - ✗ They are neither the device nor the device driver
- ✗ Major and Minor device numbers
 - ✗ Major number tells kernel which driver to use
 - ✗ Minor device indicates which device unit
 - ✗ Also used in non-standard ways to indicate certain mode of operation for device

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

10

The Filesystem

□ Devices

- ✗ Types, major numbers, and minor numbers

```
$ ls -l /dev/hda2 /dev/hdc /dev/hde /dev/tty
brw-rw---- 1 root   disk    3,  2 Aug 24  2000 /dev/hda2
brw----- 1 cappella disk   22,  0 Aug 24  2000 /dev/hdc
brw-rw---- 1 root   disk   33,  0 Aug 24  2000 /dev/hde
crw-rw-rw- 1 root   root    5,  6 Aug 24  2000 /dev/tty6
```

File Type Field
c=character, b=block

Major Device Number

Minor Device Number

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

11

The Filesystem

□ File Permissions

- ✗ User (rwx), Group (rwx), Other (rwx)
- ✗ Permissions on directories
 - ✗ r + x required for directory listing
 - ✗ w + x allows file creation, deletion, and rename
 - ✗ x allows using directory as part of a path
- ✗ Permissions on files
 - ✗ r + x required to execute scripts

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

12

The Filesystem

□ File Permissions

✗ Setuid

- ✗ Process runs with EUID = program's owner ID
- ✗ Allows process to have program file owner's access
- ✗ Used by programs such as **su**, **passwd**, and **sendmail**
- ✗ Security risk if set on un-trusted program
- ✗ Never set on shell scripts

```
$ ls -lF /usr/bin/passwd
-r-s--x--x 1 root  root    13536 Jul 12  2000 /usr/bin/passwd*
```

s=setuid in owner execute field

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

13

The Filesystem

□ File Permissions

✗ Setgid

- ✗ Process runs with EGID = program's group ID
- ✗ On directories, causes files to inherit group ID of the parent directory

```
$ ls -ldF /usr/games/freecell /var/ftp/pub
-r-xr-s--x 1 root  games  39692 May 23 11:16 /usr/games/freecell*
drwxr-sr-x 2 root  ftp    4096 Aug 17  2000 /var/ftp/pub/
```

s=setgid in group execute field

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

14

The Filesystem

□ File Permissions

✗ Sticky Bit

- ✗ Obsolete for files
- ✗ On directories, only allows owner of parent directory to rename or delete a file

```
$ ls -ldF /var/tmp
drwxrwxrwt 2 root  root    4096 Oct 13 19:53 /var/tmp/
```

t=sticky in other execute field

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

15

The Filesystem

□ File Permissions

✗ umask

- ✗ Affects permissions set on new files / directories
- ✗ The permissions you do not want to give
- ✗ Represented in octal notation
- ✗ Set a reasonable **umask** value in startup files such as /etc/csh.login or /etc/profile
 - ✗ 077 no permission for group / others
 - ✗ 066 no read/write permission for group/others
 - ✗ 022 no write permission for group / others
 - ✗ 002 no write permission for others

Updated: 10/14/2002

CIS 68C1 Unix System Administration
Copyright 2002 - Mike Cappella

16

Disks

□ Disk Interfaces - IDE

- ✗ ATA-2 (EIDE), ATA-3, Ultra-ATA, ATA-4, ATA-5
- ✗ 2 disks per channel, 2 channels per system
 - ✗ Primary master/slave, Secondary master/slave
- ✗ ATAPI Interface
 - ✗ Allows non-ATA devices using EIDE channel
 - ✗ CD-ROM, CD-Recordable, CD-RW, DVD, tape drives, super-floppy drives (ZIP, LS-120)
 - ✗ Requires ATAPI driver

Disks

□ Disk Interfaces

- ✗ SCSI
 - ✗ SCSI-1, SCSI-2 (fast, wide), SCSI-3
 - ✗ Controller is not usually built into typical PCs
 - ✗ Requires external SCSI controller
 - ✗ Known as SCSI bus adapter or host bus adapter
 - ✗ Often BIOS does not know about SCSI
 - ✗ Requires **host bus adapter** driver to boot from SCSI device
 - ✗ Technically superior to IDE
 - ✗ But is more expensive

Disks

□ Adding a Disk

- ✗ Install and connect hardware
- ✗ Create device nodes in /dev
- ✗ Format disk
- ✗ Partition disk
- ✗ Create filesystem within partitions
- ✗ Setup mounting
- ✗ Setup swapping

Disks

□ Setting up Device Entries

- ✗ Adding new device nodes
 - ✗ **mknod** utility
 - ✗ Must know the **major** / **minor** device numbers
 - ✗ MAKEDEV script/utility
 - ✗ Very OS-specific
- ✗ Most device entries are already made for you
- ✗ Check the permissions on /dev entries

Disks

□ Formatting

- ✗ Writes address & timing information to disk
- ✗ Maps bad sectors
- ✗ Also called **low-level formatting**
 - ✗ Not the same as the Window's "Format" command
- ✗ Today disks are factory pre-formatted
 - ✗ Avoid reformatting unless absolutely required
 - ✗ IDE drives not designed for re-formatting
 - ✗ Software may be available from vendor to re-format
 - ✗ SCSI drives format themselves in response to a command

Disks

□ Partitioning

- ✗ Disks must be **partitioned** before use
- ✗ Sub-divides a disk into smaller pieces
- ✗ IDE drives can have at most either
 - ✗ 4 **primary** partitions, or
 - ✗ 3 **primary** partitions + 1 **extended** partition
 - ✗ **Extended** partition can further reference many **logical** partitions
- ✗ A partition contains a filesystem

□ Partitioning

- ✗ The **fdisk** utility is used to partition a disk
 - ✗ The **cdisk** utility is also available on Linux
- ✗ Common Linux partitions
 - ✗ **root** Minimally contains files necessary to boot
 - ✗ **swap** Required for virtual memory
 - ✗ **user** Home directories, data files, source code
 - ✗ **home** Sometimes used for home directories
 - ✗ **boot** Contains boot kernel(s), etc.

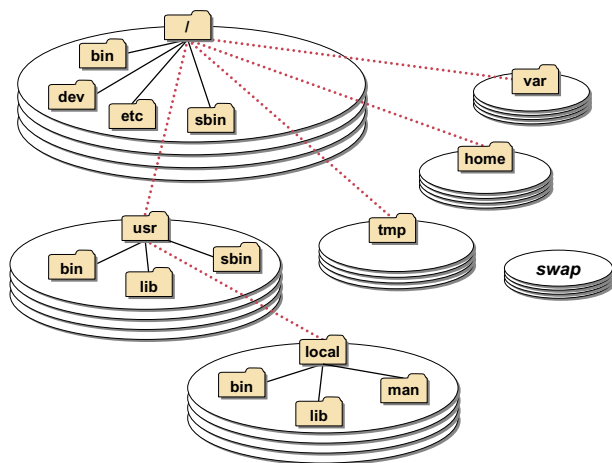
Disks

□ Partitioning

- ✗ Careful partition planning is essential ...
 - ✗ for optimal performance
 - ✗ to provide ample storage space
 - ✗ for performing backups
 - ✗ to avoiding filling the root filesystem
 - ✗ allow for redundancy
 - ✗ for storing crash dumps if system panics

Disks

Partitioning



Updated: 10/14/2002

CIS 6801 Unix System Administration
Copyright 2002 - Mike Cappella

25

Disks

Creating a Filesystem

- ✗ The **ext2** filesystem
 - ✗ Very efficient - based on BSD Fast Filesystem
 - ✗ **Inode table**
 - ✗ Table that keeps track of files
 - ✗ Fixed size - cannot be changed once created
 - ✗ **Superblocks**
 - ✗ Superblock is master record for filesystem
 - ✗ Backup superblocks are created for safety
 - ✗ Always a backup at block 32

Updated: 10/14/2002

CIS 6801 Unix System Administration
Copyright 2002 - Mike Cappella

26

Disks

Creating a Filesystem

- ✗ **mkfs** creates a filesystem within a partition
 - ✗ Creates **lost+found** directory used by **fsck**
- ✗ Linux device nodes for hard disks
 - ✗ IDE
 - ✗ `/dev/hd[a-h][partition_number]`
 - ✗ SCSI
 - ✗ `/dev/sd[a-p][partition_number]`

Updated: 10/14/2002

CIS 6801 Unix System Administration
Copyright 2002 - Mike Cappella

27

Disks

Setting up automatic mounting

- ✗ Mount filesystems to make them available
 - ✗ Manually mount new filesystem first to test
- ✗ **/etc/fstab**
 - ✗ Contains information about filesystems
 - ✗ Device node, mount point, file system type, mount options, dump frequency, and fsck order
 - ✗ Add entry to cause automatic mounting at boot
 - ✗ Read by **mount**, **umount**, **swapon**, and **fsck**

Updated: 10/14/2002

CIS 6801 Unix System Administration
Copyright 2002 - Mike Cappella

28

Disks

- Setting up Swap Partition
 - ✗ Must initialize a swap partition with **mkswap** prior to use
 - ✗ Add entry into **/etc/fstab**
 - ✗ Do not run **fsck** on swap partitions
- Checking a Filesystem
 - ✗ **fsck** checks filesystem integrity / fixes inconsistencies
 - ✗ It runs automatically at boot time
 - ✗ Skips filesystem check if it is marked **clean**
 - ✗ No guarantee the filesystem is clean however
 - ✗ Should periodically force a clean

Disks

- Checking a Filesystem
 - ✗ Cannot always automatically fix all problems
 - ✗ Requires manual execution
 - ✗ Places unfixable files in **lost+found**
 - ✗ Always check root filesystem
 - ✗ in single-user mode
 - ✗ before others are checked or mounted
 - ✗ Rerun **fsck** until there are no errors
 - ✗ Never boot into multi-user with corrupted root filesystem