

CAPITOLO 6

IL FILESYSTEM DI LINUX

Pietro Buffa

Uno degli scogli con cui un novello utente Linux deve fare i conti è indubbiamente la gestione dei files. La prima cosa che “salta agli occhi” è infatti la presenza di un gran numero di directory già dal primo accesso al sistema. Mentre una tipica installazione di Windows crea poche cartelle (\Windows, \Programmi, \Documenti), in Linux ci troviamo di fronte ad almeno una decina di directory dai nomi abbastanza criptici, che analizzeremo in dettaglio nel corso di questo capitolo.

Il fatto che vi siano così tante directory è del tutto normale in questo tipo di sistemi e presto vedremo come tutto ciò non sia affatto a scapito dell'utente ma in linea con un concetto molto importante, l'ordine: in Linux i files vengono solitamente disposti in base alla loro funzione. Ma andiamo per gradi e cerchiamo prima di chiarire cosa si intende per filesystem. Le periferiche di memorizzazione di massa: Floppy, CD-Rom, DVD ed Hard Disk, sono significativamente “stupide”. Esse si limitano in genere a supportare la lettura e la scrittura di blocchi di dati dalla dimensione prefissata, la gestione di questo spazio di memorizzazione viene effettuata da una parte fondamentale del sistema operativo: il filesystem, che fornisce quindi il modo logico per memorizzare, leggere e organizzare i dati sui vari dispositivi. Il file è l'unità fondamentale di un filesystem, altro non è che una sequenza di byte alla quale si può accedere in maniera sequenziale dall'inizio alla fine. Per gestire meglio la complessità e consentire di trovare i files più facilmente, i filesystem implementano dei contenitori chiamati directory (o meno propriamente, cartelle), ogni directory può contenere uno o più files e/o altre directory. Come conseguenza, l'organizzazione del filesystem si presenta in un sistema GNU/Linux come una struttura gerarchica simile ad un albero rovesciato, la cui radice (root, in Inglese) è la directory base, indicata con il simbolo “/”, da dove si dipartono altre directory che a loro volta si possono ulteriormente dividere in altre directory e così via.

6.1 LA “STRUTTURA GERARCHICA AD ALBERO”

Andiamo subito a vedere più da vicino l'organizzazione di un ipotetico filesystem Linux standard. Una tipica installazione del sistema GNU/Linux crea generalmente tutte le directory rappresentate nello schema sottostante. Accade spesso però che alcune distribuzioni personalizzino un pò le cose rispetto allo standard, per cui alcune directory o alcuni files qui elencati potrebbero non esistere sul vostro sistema o presentarsi in maniera un pò diversa (ad esempio la directory */etc*, modifica il suo contenuto in relazione ai pacchetti che avete installato sul vostro sistema). Cercate dunque di leggere questo capitolo afferrandone la filosofia che sta alla base e non i dettagli.

Il filesystem schematizzato nella **Fig. 6.1** mostra la tipica organizzazione delle directory standard in un sistema GNU/Linux. Esso viene generalmente detto **filesystem principale** e viene “montato” all'avvio del sistema in maniera automatica (vedremo più avanti, nel corso di questo manuale, di approfondire i termini: “montare” e “smontare”, concetti molto importanti nei sistemi GNU/Linux). A questo è possibile collegare **filesystem secondari**, sempre mediante una particolare operazione definita di “montaggio” che può essere svolta soltanto dall'amministratore di sistema. Anche se ne parleremo in maniera dettagliata più avanti, vediamo comunque di chiarire, almeno nelle linee principali, l'importante concetto di “montaggio”: quando si inserisce un floppy nel lettore di un PC con sistema Windows, il contenuto dello stesso è automaticamente disponibile nel file manager, di norma basta fare un doppio click sull'apposita icona (A:). Quando inserite un floppy nel lettore di un PC con sistema GNU/Linux, è necessario che avvenga una particolare operazione che ha lo scopo di collegare il filesystem del dispositivo floppy al filesystem principale del sistema. Questa operazione, definita di “montaggio”, viene svolta in automatico non appena un qualunque supporto di memorizzazione di massa viene collegato al computer o inserito in esso, è bene comunque sapere cosa fare se l'automatismo manca o dovesse non funzionare, in questi casi, la riga di comando ci viene ancora in aiuto.

TIPICO FILESYSTEM GNU/LINUX:

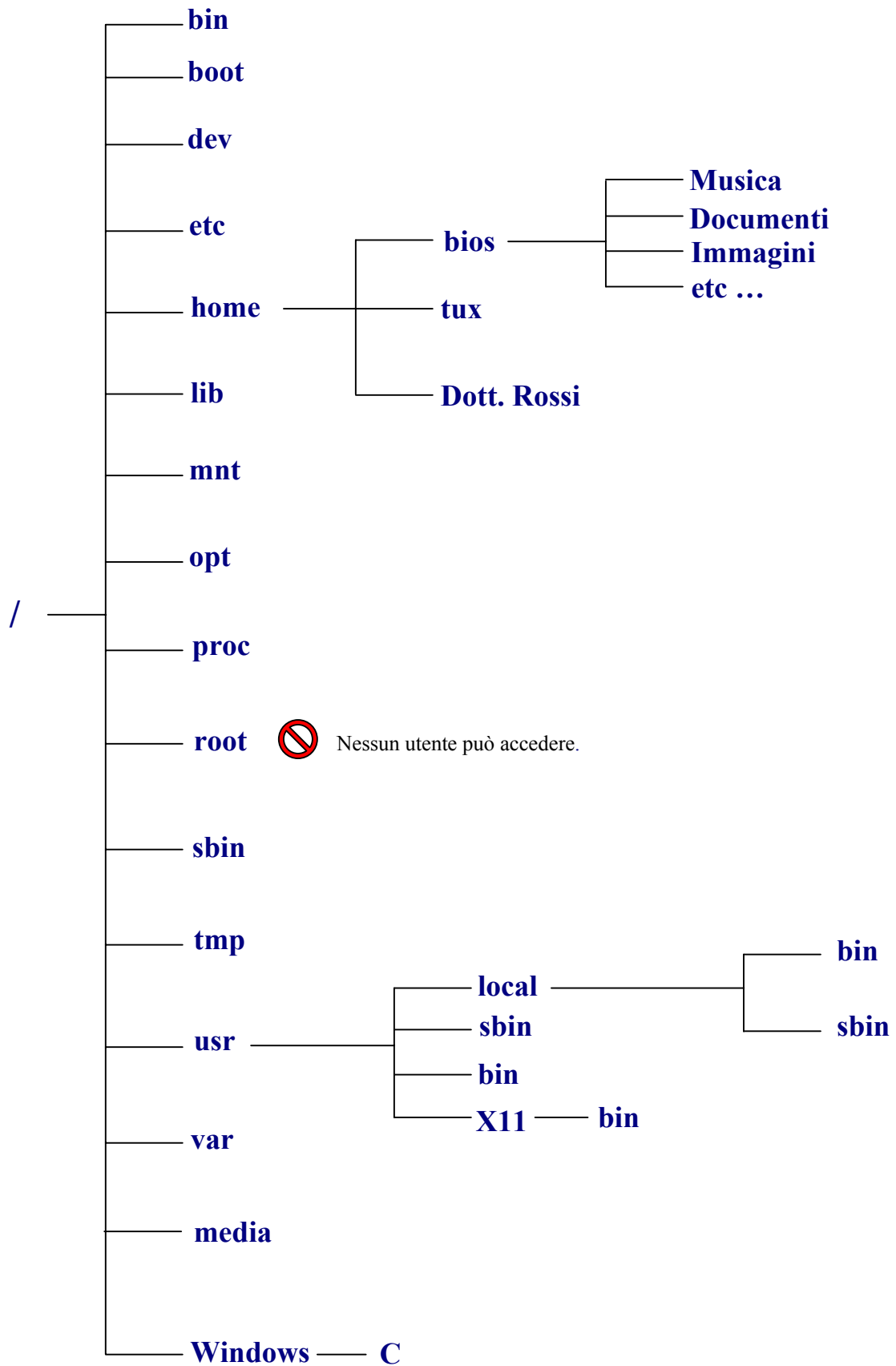


Fig. 6.1 Struttura gerarchica ad albero.

6.2 ANALISI DELLE PRINCIPALI DIRECTORY

La directory */bin*

La directory */bin* contiene i file binari (ovvero eseguibili), sia dell'utente che di sistema, da qui il suo nome. La maggior parte dei comandi GNU/Linux che andremo a studiare nel corso di questo manuale richiamano il relativo programma eseguibile posto generalmente all'interno di questa directory. Tuttavia non tutti i files presenti sono in formato binario poiché nei moderni sistemi alcune funzioni sono svolte da script shell.

La directory */boot*

Contiene file necessari per il processo di avvio (bootstrap) del sistema e l'immagine del kernel (chiamata *vmlinuz*). Modificare qualcosa in questa directory senza una conoscenza nell'ambito dell'amministrazione è decisamente una cattiva idea. Per questo motivo di default la directory */boot* è di proprietà del Super Utente *root*.

La directory */dev*

Questa è la directory che contiene le definizioni di tutti i device (dispositivi hardware) installati ed installabili sul sistema: hard disk, mouse, stampanti, consolle, modem, memorie, unità floppy, CD-ROM, porte USB, etc. I file contenuti in questa directory sono davvero tanti in modo da accogliere quante più periferiche possibili, anche di differenti marche, che possono essere quindi collegate al computer.

I file dei dispositivi sono suddivisi in due tipologie principali:

- Device a blocchi
- Device a caratteri

Il tipo indica il metodo di trasmissione e ricezione dei dati. Senza entrare troppo in particolari, possono essere considerati "Device a blocchi" tutti quei dispositivi che memorizzano dati, mentre "Device a caratteri" sono tutti quei dispositivi che trasmettono dati. Facciamo un esempio: i dispositivi floppy, i dischi rigidi, i CD-ROM, sono tutti device a blocchi mentre le porte parallele, le porte seriali, i mouse, i modem o le stampanti sono tutti device a caratteri.

La directory */etc*

Avrebbe potuto essere chiamata più adeguatamente "config" poiché contiene quasi tutti i file di configurazione e le directory subordinate ai dati di configurazione del sistema. La maggior parte dei file in questa directory sono quindi riservati all'amministratore e sono di importanza cruciale nelle operazioni di amministrazione e manutenzione del sistema. Esistono oggi in ogni distribuzione GNU/Linux, diverse utility grafiche in grado di andare a modificare molti di questi file di configurazione, tuttavia è importante notare che queste utility sono soltanto destinate a rendere più semplice la manutenzione del sistema. Il modo migliore per comprendere questi file, è aprirli e modificarli in maniera diretta, utilizzando un semplice editor di testo.

ALCUNI FILE DI CONFIGURAZIONE ABITUALMENTE MODIFICATI:

- **hostname** - Questo file contiene il nome di dominio qualificato (FQDN) del sistema. Per esempio, se la macchina è hook ed il dominio aziendale è fish.net, il vostro FQDN sarà hook.fish.net.
- **X11** - Questo file contiene tutte le informazioni necessarie che permettono al server X di funzionare.
- **Crontab** - Questo file è la tabella cron impostata per l'esecuzione automatica di routine di sistema.
- **fstab** - Questa è la tabella dei filesystem montati automaticamente all'avvio del sistema. Le voci di questo file vengono create per definire device specifici, filesystem, mount point ed operazioni di montaggio (vedremo più avanti di cosa si tratta).
- **group** - Il database dei gruppi di utenti.
- **grub.conf** - GRUB e LILO sono i due "bootmanager" più famosi in Linux. Essi si escludono reciprocamente. Questo file contiene la configurazione di GRUB.
- **lilo.conf** - Questo file contiene la configurazione di LILO.
- **host** - Questo file viene utilizzato per definire un nome di sistema ed un dominio in combinazione con un indirizzo IP.
- **inittab** - E' la tabella che contiene le definizioni per i processi del programma "init", ancora utilizzato da tante distribuzioni, che svolge importanti operazioni all'avvio del sistema.
- **issue** - Questo file contiene il testo che viene visualizzato prima del login; in altre parole, il messaggio che viene mostrato subito sopra il prompt di login.

- **motd** - File del “messaggio del giorno”. L’amministratore può inserire in questo file qualsiasi testo che deve essere mostrato a tutti gli utenti in sede di login nel sistema.
- **passwd** - E’ il database degli utenti. Un tempo le password degli utenti erano memorizzate in questo file, da qui il suo nome. Adesso le password sono codificate e memorizzate nel file */etc/shadow*.
- **printcap** - In questo file vengono memorizzate le definizioni per le stampanti di sistema, locali o remote.
- **profile** e **bashrc** - File letti ed eseguiti dalla shell all’accesso che permettono all’amministratore di impostare opzioni predefinite per tutti gli utenti.
- **shadow** - In questo file sono memorizzate le password codificate per gli account utente.
- **apt** - In alcune famose distribuzioni (Debian e Ubuntu), esiste un file chiamato *source.list* contenente la lista dei repository su cui queste distribuzioni si poggiano per l’installazione dei software direttamente dalla rete.

La directory */home*

La directory più importante per ogni utente Linux è la sua home directory. Ogni nuovo utente deve avere una directory personale all’interno della directory home (vedi schema). In essa vengono memorizzati i files personali e tutti i files di configurazione personali utilizzati dalle applicazioni.

Per default ogni utente effettua il login nella sua specifica directory home. Provate a dargli un’occhiata, vi accorgete che di files ne contiene veramente pochini. In realtà i files che avete nella vostra home sono molti di più, specialmente se state utilizzando il vostro sistema da qualche tempo ed avete già installato e configurato qualche applicazione.

Ma come! Vi chiederete, «I files di configurazione non avevamo detto che si trovano nella directory */etc*?».

In effetti è vero che i files di configurazione si trovano normalmente in quella directory, ma si tratta dei files contenenti le impostazioni generali del sistema e dei programmi. Nella vostra directory home, si trovano invece i files di configurazione personali, che registrano le vostre personalizzazioni senza che gli altri utenti debbano soffrirne. Quando lanciate un’applicazione, questa controllerà prima le impostazioni generali nel suo file di configurazione posto nella directory */etc* e poi quelle particolari sotto la vostra directory */home*.

I files di configurazione personali vengono solitamente memorizzati nella vostra home come file nascosti (il loro nome comincia con un puntino), per tale motivo sono invisibili all’utente, abbiate pazienza ed impareremo più avanti come far elencare alla shell bash file normali e nascosti.

La directory */lib*

Uno dei problemi che gli utenti di Windows hanno è il fatto di avere a che fare con le DLL (Dynamic Link Library). Le famiglie DLL sono essenziali per la stabilità del sistema, ma alcuni programmatori hanno deciso di lasciare a chiunque o a qualunque applicazione la facoltà di aggiornarle o di sovrascriverle. L’equivalente Linux della DLL sono le cosiddette “librerie condivise” (Shared Library). Esse sono di norma contraddistinte dall’estensione *.so* e giacciono nella directory */lib*. La presenza di queste librerie all’interno del sistema operativo, permette ai programmi di essere più piccoli e veloci, in quanto fanno riferimento a funzioni presenti già in queste librerie. Esse sono quindi essenziali per il corretto funzionamento del sistema.

La directory */mnt*

Questa directory viene comunemente utilizzata per effettuare il mount (montaggio) di Filesystem Secondari al Filesystem Principale da parte dell’amministratore di sistema. Bisogna dire che non è obbligatorio utilizzare questa directory come punto di montaggio, ma essa viene fornita per comodità.

La directory */opt*

La directory */opt* (options) è la directory in cui sono installate le applicazioni aggiuntive. Solitamente tutte le applicazioni di terze parti vengono installate in questa directory.

La directory */proc*

La directory */proc* non occupa alcuno spazio sul disco, in quanto si tratta di un filesystem virtuale allocato soltanto in memoria. I files di */proc* contengono informazioni dettagliate in merito ai vari processi in esecuzione sul sistema.

La directory */root*

E' la home directory del Super Utente root o amministratore. La directory */root* non deve essere confusa con la directory */*, che viene comunemente indicata come la directory di livello massimo all'interno del filesystem.

La directory */sbin*

La directory */sbin* (Superuser Bin) contiene i files binari di amministrazione. In genere contiene programmi necessari all'avvio del sistema ed utility per il partizionamento del disco rigido, per la creazione, il controllo e la riparazione di un filesystem etc.

La directory */tmp*

Molti programmi hanno bisogno di uno spazio temporaneo per memorizzare i dati durante i cicli di elaborazione. La collocazione tradizionale di questi files è la directory */tmp*.

La directory */usr*

Essa costituisce la sezione principale del filesystem, qui si trova gran parte del software di sistema, la maggior parte della documentazione (inclusa nella directory */usr/doc*), varie librerie, l'albero sorgente completo per il kernel Linux (directory */usr/src*) e via dicendo. Non è raro che su un sistema medio, */usr* tenda a divenire la directory più grande.

La directory */var*

La directory */var* è impostata in modo da contenere i dati variabili e la sua dimensione varia continuamente. Essa contiene i files di spooling per la posta e la stampa, le news, informazioni sui crash ed altri files.

La directory */Windows*

Nelle moderne distribuzioni non è raro trovare una directory del genere. Con l'aumentare delle dimensioni degli Hard disk, sono molte le persone che preferiscono far coesistere sul proprio computer due sistemi operativi, solitamente Windows e linux. Questa directory ha come unico scopo quello di servire come punto di montaggio per un eventuale filesystem Windows.

6.3 IL FILESYSTEM NATIVO DI LINUX

Fino ad ora ci siamo limitati ad osservare come GNU/Linux si sia creato un sistema gerarchico sul disco tale che tutto quello che salviamo o installiamo, abbia una collocazione ordinata a lui congeniale. GNU/Linux supporta almeno una dozzina di filesystem diversi, ma quello usato di default e che si può considerare in qualche modo il filesystem "nativo" è il **Second Extended Filesystem** o **ext2** introdotto nell'ormai lontano 1994, solido, affidabile ed in grado di gestire files su partizioni estese fino a 4 TB (TeraByte), da qui il suo nome. Anche se ext2 ha dimostrato negli anni di essere robusto, presenta oggi un importante difetto: non da garanzie sullo stato del filesystem in caso di crash del sistema. Se ciò succede, il filesystem rimane in uno stato di inconsistenza e deve essere verificato prima di poter essere usato nuovamente. Tutti i filesystem hanno i loro strumenti di controllo ma visto che il sistema non ha idea su quale area è avvenuto il crollo, non ha alternative se non quella di controllare l'intera struttura che, nel caso di filesystem voluminosi, si traduce in decine di minuti di attesa. La soluzione più comune per aggirare il lungo controllo sul filesystem dopo un crash è quella di conservare una traccia delle attività recenti. Una simile traccia è chiamata journal e l'operazione di scansione di un journal ed il ripristino del filesystem è detta "replay". Dal momento che il journal è enormemente più piccolo dell'intero filesystem, l'operazione di replay risulta sempre veloce. Nasce così il filesystem **ext3**, che offendo il supporto al journaling, ha preso di diritto il posto dell'ormai vecchio ext2. Il filesystem **ext4** è in corso di sviluppo.

SUPPORTO A FILESYSTEM ESTERNI

Linux ha la grande capacità di supportare e quindi riconoscere le strutture di differenti filesystem, molti dei quali utilizzati dai sistemi operativi comuni, analizziamone qualcuno:

- **MsDOS, VFAT** Linux implementa la FAT12 usata nelle prime versioni di MsDOS e le nuove varianti FAT16 e FAT32. Questo è molto importante in quanto tale tipologia di filesystem è ampiamente utilizzata nei dispositivi attuali di memorizzazione esterni come HardDisk o memorie flash USB.

- **NTFS** è stato creato da Microsoft nel tentativo di offrire un file system superiore per classi Enterprise ed oggi divenuto il filesystem di default di WindowsXP. Il supporto a Linux è storicamente limitato all'accesso in sola lettura, ma recentemente sta iniziando a consolidarsi anche il pieno supporto delle operazioni di scrittura.
- **HFS** (Hierarchical File System) è il filesystem dei vecchi computer Macintosh di Apple. Extended HFS (o **HFS+**) è oggi usato dal moderno OSX. Ha guadagnato il supporto di Linux solo recentemente.
- **ISO9660** Filesystem standard per CD-Rom e DVD-Rom.