

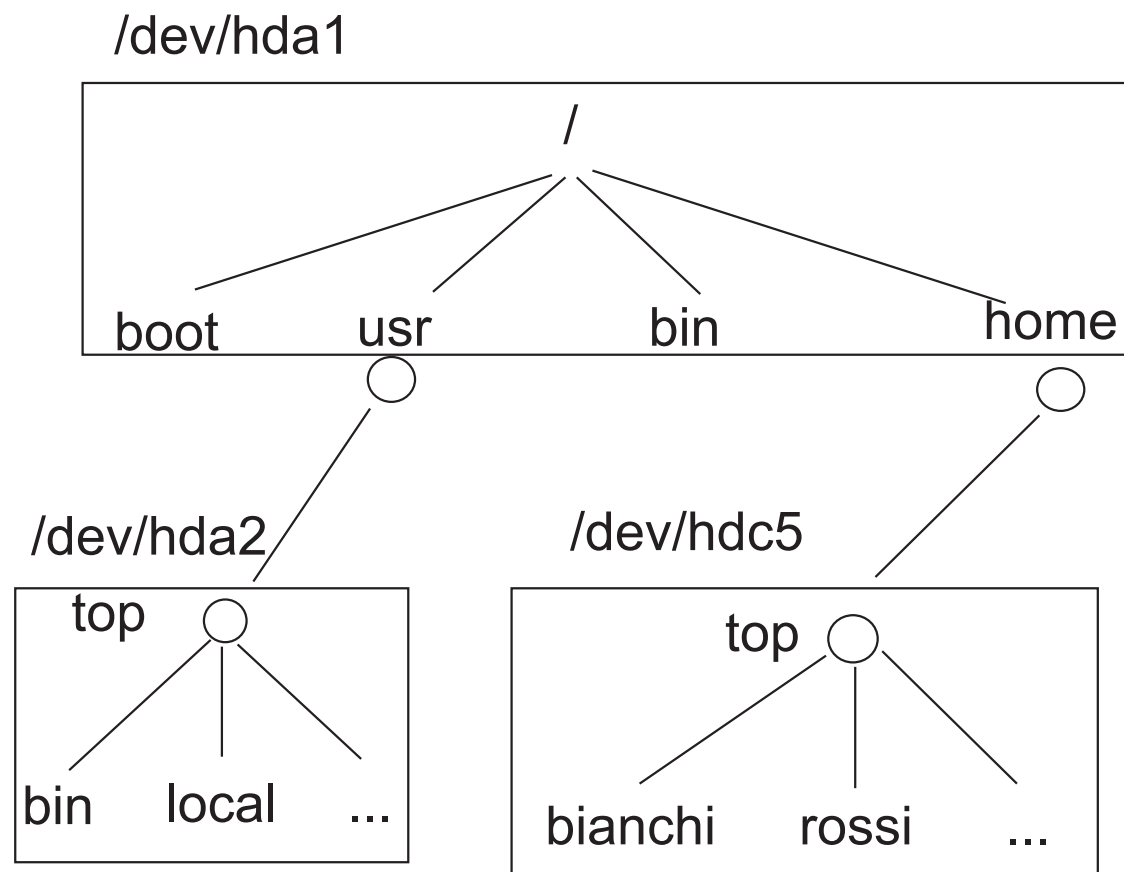
Il filesystem di Unix/Linux (I)

- Comunemente, in un elaboratore l'informazione è memorizzata in modo permanente nei **dischi fissi**.
- Ogni disco fisso può essere suddiviso in **partizioni**.
- Ogni partizione può contenere un filesystem con una propria **top level directory**.

Sorge quindi il problema di come permettere agli utenti di accedere ai vari filesystem contenuti nelle differenti partizioni:

- la prima possibilità consiste nell'avere **root directory distinte** (e.g. in Windows: C:\, D:\ ecc.): una per ogni partizione; quindi per riferirsi ad un file, bisogna usare un pathname che parta dalla root directory giusta (e.g. D:\Doc\bilancio.xls).
- Unix/Linux invece fa in modo che i diversi filesystem vengano combinati in un'**unica struttura gerarchica**, "montando" la top level directory di una partizione come foglia del filesystem di un'altra partizione.

Il filesystem di Unix/Linux (II)



dove `/dev/hda1`, `/dev/hda2` e `/dev/hdc5` sono partizioni differenti. Le informazioni su quali filesystem montare al boot ed in che modo sono contenute nel file `/etc/fstab`. Il comando per montare i filesystem è `mount <file speciale> <mount point>` e, solitamente, solo l'utente `root` può utilizzarlo. `mount` senza argomenti elenca i filesystem in uso nel sistema.

Controllo dello spazio su disco

Per controllare la quantità di spazio su disco in uso:

```
user> df
/                (/dev/dsk/c0t0d0s0 ): 2231020 blocks    304297 files
/proc           (/proc                ):          0 blocks     11692 files
/opt            (/dev/dsk/c0t0d0s3 ): 2486076 blocks    325130 files
/usr/local     (/dev/dsk/c0t0d0s7 ): 4067088 blocks    430227 files
/opt/solaris2  (apphost:/opt/solaris2): 1995968 blocks    286013 files
```

Legenda: il primo campo contiene il nome del file system; il secondo il device corrispondente (eventualmente virtuale); il terzo il numero di blocchi occupati; il quarto il numero di inode.

Per controllare la quantità di spazio su disco utilizzata da una directory (in blocchi):

```
user> du LABORATORIO_S0

8      LABORATORIO_S0/LABSO/CVS
16884  LABORATORIO_S0/LABSO
16     LABORATORIO_S0/scriptColonne
14     LABORATORIO_S0/linguaggio_c
17342  LABORATORIO_S0
```

Ulteriori comandi sui file

- Confronto tra file:

1. `> cmp file1 file2`

restituisce il primo byte ed il numero di linea in cui `file1` e `file2` differiscono (se sono uguali, non viene stampato nulla a video).

2. `> diff file1 file2`

restituisce la lista di cambiamenti da apportare a `file1` per renderlo uguale a `file2`.

- Ricerca di file:

`> find <pathnames> <expression>`

attraversa ricorsivamente le directory specificate in `<pathnames>` applicando le regole specificate in `<expression>` a tutti i file e sottodirectory trovati.

`<expression>` può essere una fra le seguenti:

1. opzione,
2. condizione,
3. azione.

Esempi d'uso di find

- `> find . -name '*.c' -print`
cerca ricorsivamente a partire dalla directory corrente tutti i file con estensione `c` e li stampa a video.
- `> find . -name '*.bak' -ls -exec rm {} \;`
cerca ricorsivamente a partire dalla directory corrente tutti i file con estensione `bak`, li stampa a video con i relativi attributi (`-ls`) e li cancella (`-exec rm {} \;`; Il carattere `\` serve per fare il “quote” del `;`).
- `> find /etc -type d -print`
cerca ricorsivamente a partire dalla directory `/etc` tutte e solo le sottodirectory, stampandole a video.

I comandi filtro

I **filtri** sono una particolare classe di comandi che possiedono i seguenti requisiti:

- prendono l'input dallo **standard input device**,
- effettuano delle operazioni sull'input ricevuto,
- inviano il risultato delle operazioni allo **standard output device**.

Tali comandi risultano quindi degli ottimi strumenti per costruire pipeline che svolgano compiti complessi.

Ad esempio:

```
> uniq file
```

restituisce in output il contenuto del file `file`, sostituendo le linee adiacenti uguali con un'unica linea.

Comandi filtro: grep, fgrep, egrep

I comandi:

- grep: General Regular Expression Parser,
- fgrep: Fixed General Regular Expression Parser,
- egrep: Extended General Regular Expression Parser,

restituiscono solo le linee dell'input fornito che contengono un **pattern** specificato tramite espressione regolare o stringa fissata.

Sintassi:

```
grep [options] pattern [filename...]
```

```
fgrep [options] string [filename...]
```

```
egrep [options] pattern [filename...]
```

Opzioni:

-i: ignora la distinzione fra lettere maiuscole e minuscole,

-l: fornisce la lista dei file che contengono il pattern/string,

-n: le linee in output sono precedute dal numero di linea,

-v: vengono restituite solo le linee che **non** contengono il pattern/string,

-w: vengono restituite solo le linee che contengono il pattern/string come parola completa,

-x: vengono restituite solo le linee che coincidono esattamente con pattern/string.

I metacaratteri delle espressioni regolari

metacarattere	tipo	significato
<code>^</code>	B	inizio della linea
<code>\$</code>	B	fine della linea
<code>\<</code>	B	inizio di una parola
<code>\></code>	B	fine di una parola
<code>.</code>	B	un singolo carattere (qualsiasi)
<code>[str]</code>	B	un qualunque carattere in <code>str</code>
<code>[^str]</code>	B	un qualunque carattere non in <code>str</code>
<code>[a-z]</code>	B	un qualunque carattere tra a e z
<code>\</code>	B	inibisce l'interpretazione del metacarattere che segue
<code>*</code>	B	zero o più ripetizioni dell'elemento precedente
<code>+</code>	E	una o più ripetizioni dell'elemento precedente
<code>?</code>	E	zero od una ripetizione dell'elemento precedente
<code>{j,k}</code>	E	un numero di ripetizioni compreso tra j e k dell'elemento precedente
<code>s t</code>	E	l'elemento s oppure l'elemento t
<code>(exp)</code>	E	raggruppamento di <code>exp</code> come singolo elemento

dove B (basic) indica che la sequenza di caratteri è utilizzabile sia in `grep` che in `egrep`, mentre E (extended) indica che la sequenza di caratteri è utilizzabile solo in `egrep` (o in `grep` usando l'opzione `-E`).

Esempi d'uso di grep, fgrep, egrep

- `> fgrep rossi /etc/passwd`
fornisce in output le linee del file `/etc/passwd` che contengono la stringa **fissata** `rossi`.
- `> egrep -nv '[agt]+' relazione.txt`
fornisce in output le linee del file `relazione.txt` che **non** contengono stringhe composte dai caratteri `a`, `g`, `t` (ogni linea è preceduta dal suo numero).
- `> grep -w print *.c`
fornisce in output le linee di tutti i file con estensione `c` che contengono la parola **intera** `print`.
- `> ls -al . | grep '^d.....w.'`
fornisce in output le sottodirectory della directory corrente che sono modificabili dagli utenti ordinari.
- `> egrep '[a-c]+z' doc.txt`
fornisce in output le linee del file `doc.txt` che contengono una stringa che ha un prefisso di lunghezza non nulla, costituito solo da lettere `a`, `b`, `c`, seguito da una `z`.

Comandi filtro: sort

Il comando `sort` prende in input delle linee di testo, le **ordina** (tenendo conto delle opzioni specificate dall'utente) e le invia in output.

- `sort` tratta ogni linea come una collezione di vari campi separati da delimitatori (default: spazi, tab ecc.).
- l'ordinamento di default avviene in base al **primo** campo ed è **alfabetico**.

Il comportamento di default si può cambiare tramite le opzioni:

- b ignora eventuali spazi presenti nelle chiavi di ordinamento,
- f ignora le distinzioni fra lettere maiuscole e minuscole,
- n considera numerica (invece che testuale) la chiave di ordinamento
- r ordina in modo decrescente,
- o *file* invia l'output al file *file* invece che allo standard output,
- t *s* usa *s* come separatore di campo,
- k *s1*,*s2* usa i campi da *s1* a *s2*-1 come chiavi di ordinamento.

Esempi d'uso di sort

Volendo ordinare le righe del file `/etc/passwd` in base al terzo campo (user ID), il comando

```
> sort -t: -k3,4 /etc/passwd
root:x:0:1:Super-User:/:/sbin/sh
guest:x:1001:120:Guest User:/home/guest:/usr/local/bin/bash
daemon:x:1:1::/:
...
```

non dà il risultato voluto in quanto di default l'ordinamento è alfabetico, mentre il campo user ID è un numero; quindi si rende necessaria l'opzione `-n`:

```
> sort -t: -k3,4 -n /etc/passwd
root:x:0:1:Super-User:/:/sbin/sh
daemon:x:1:1::/:
bin:x:2:2::/usr/bin:
...
guest:x:1001:120:Guest User:/home/guest:/usr/local/bin/bash
...
```

Si noti che in entrambi gli esempi il separatore (`:`) è stato impostato con l'opzione `-t:`.

Comandi filtro: tr

Character translation: `tr` è un semplice comando che permette di eseguire operazioni come la conversione di lettere minuscole in maiuscole, cancellazione della punteggiatura ecc. Siccome può prendere input soltanto dallo standard input e stampare soltanto sullo standard output, bisogna usare delle pipe o delle ridirezioni di input/output per farlo leggere/scrivere su file.

Sintassi di base: `> tr str1 str2`

(i caratteri in `str1` vengono sostituiti con i caratteri in posizione corrispondente della stringa `str2`)

Esempi:

- `> tr a-z A-Z`
converte le minuscole in maiuscole.
- `> tr -c A-Za-z0-9 ' '`
sostituisce tutti i caratteri **non** (opzione `-c`: complemento) alfanumerici con degli spazi.
- `> tr -cs A-Za-z0-9 ' '`
come nell'esempio precedente, ma comprime gli spazi adiacenti in un'unico spazio (opzione `-s`: *squeeze*).
- `> tr -d str`
cancella i caratteri contenuti nella stringa `str`.

Cut and paste

- Il comando `cut` serve ad estrarre delle colonne specifiche dalle linee di testo che riceve in input:

```
> cut -d: -f1 /etc/passwd  
root  
daemon  
bin  
...
```

il separatore si specifica con l'opzione `-d` (*delimiter*), il campo da estrarre con l'opzione `-f` (*field*).

- Il comando `paste` combina le righe corrispondenti di due file, inserendo un delimitatore fra esse (default: <Tab>):

```
> cd; cut -d: -f1 /etc/passwd > p1.txt; cut -d: -f6 /etc/passwd > p6.txt  
> paste p1.txt p6.txt  
root      /  
daemon    /  
bin       /usr/bin  
...
```

Comandi filtro: sed

Il nome del comando `sed` sta per **S**tream **E**ditor e la sua funzione è quella di permettere di editare il testo passato da un comando ad un altro in una pipeline. Ciò è molto utile perché tale testo non risiede fisicamente in un file su disco e quindi non è editabile con un editor tradizionale (e.g. `vi`). Tuttavia `sed` può anche prendere in input dei file, quindi la sua sintassi è la seguente:

```
sed actions files
```

- Se non si specificano azioni, `sed` stampa sullo standard output le linee in input, lasciandole inalterate.
- Se vi è più di un'azione, esse possono essere specificate sulla riga di comando precedendo ognuna con l'opzione `-e`,
oppure possono essere lette da un file esterno specificato sulla linea di comando con l'opzione `-f`.
- Se non viene specificato un **indirizzo** o un intervallo di indirizzi di linea su cui eseguire l'azione, quest'ultima viene applicata a tutte le linee in input. Gli indirizzi di linea si possono specificare come **numeri** o **espressioni regolari**.

Esempi d'uso di sed

- > `sed '4,$d' /etc/passwd`

stampa a video soltanto le prime 3 righe del file `/etc/passwd`: `d` è il comando di cancellazione che elimina dall'output tutte le righe a partire dalla quarta (`$` sta per l'ultima riga del file).

- > `sed 3q /etc/passwd`

stesso effetto del precedente comando: in questo caso `sed` esce dopo aver elaborato la terza riga (`3q`).

- > `sed /sh/y/:0/_%/ /etc/passwd`

sostituisce in tutte le righe che contengono la stringa `sh` il carattere `:` con il carattere `_` ed il carattere `0` con il carattere `%`.

- > `sed '/sh/!y/:0/_%/' /etc/passwd`

sostituisce in tutte le righe che **non** contengono la stringa `sh` il carattere `:` con il carattere `_` ed il carattere `0` con il carattere `%`. Si noti l'uso del quoting per impedire che la shell interpreti il metacarattere `!`.

Sostituzione del testo con sed

Il formato dell'azione di sostituzione in sed è il seguente:

`s/expr/new/flags`

dove:

- `expr` è l'espressione da cercare,
- `new` è la stringa da sostituire al posto di `expr`,
- `flags` è uno degli elementi seguenti:
 - `num`: un numero da 0 a 9 che specifica quale occorrenza di `expr` deve essere sostituita (di default è la prima),
 - `g`: ogni occorrenza di `expr` viene sostituita,
 - `p`: la linea corrente viene stampata sullo standard output nel caso vi sia stata una sostituzione,
 - `w file`: la linea corrente viene accodata nel file `file` nel caso vi sia stata una sostituzione.

Esempi di sostituzioni con sed

- `sed '/^root/,/^bin/s/:.....:/::/w disabled.txt' /etc/passwd`
sostituisce la password criptata (lunga 13 caratteri) con la stringa vuota nelle righe in input comprese fra quella che inizia con `root` e quella che inizia con `bin`; tali righe sono poi accodate nel file `disabled.txt`.
- `cat /etc/passwd | sed 's?/bin/. *sh$/usr/local&?'`
cerca tutte le righe in input in cui compare la stringa corrispondente all'espressione regolare `/bin/. *sh$` (ad esempio `/bin/bash`) e sostituisce quest'ultima con la stringa corrispondente a `/usr/local/bin/. *sh$` (ad esempio `/usr/local/bin/bash`). Si noti che, siccome il carattere separatore di `sed` compare nella stringa da cercare, si è usato il carattere `?` come separatore. Inoltre il carattere `&` viene rimpiazzato automaticamente da `sed` con la stringa cercata (corrispondente a `/bin/. *sh$`).

Esercizi

- Scoprire quanto spazio occupa il contenuto della propria home directory. Esiste un modo per ottenere in output soltanto il numero di blocchi (evitando di visualizzare informazioni ulteriori)?
- Qual è l'effetto del comando `sort file >file`, dove `file` è il nome di un file?
- Fare alcuni esperimenti per scoprire qual è l'effetto del comando `tr str1 str2` se le stringhe `str1` e `str2` hanno lunghezze diverse.
- Scrivere un comando per sostituire tutti i caratteri alfanumerici nell'input con un carattere `<Tab>`, in modo che non compaiano più `<Tab>` consecutivi.
- Il comando `date` fornisce data e ora su standard output. Scrivere una pipeline per estrarre soltanto i minuti.
- Scrivere una pipeline che permetta di scoprire se ci sono linee ripetute in un file.
- Visualizzare su standard output, senza ripetizioni, lo user ID di tutti gli utenti che hanno almeno un processo attivo nel sistema.