

# Dischi e Filesystem

Stefano Sanfilippo

POuL – Politecnico Open unix Labs, Corsi Linux Avanzati

27 marzo 2013

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

## 1 Definizione

## 2 Struttura

- Generale
- Eseguibili
- Speciali
- Deviazioni dallo standard

*FHS* è lo standard per la struttura delle directory nei sistemi Linux.

<http://www.pathname.com/fhs/>

- Estensione della gerarchia BSD
- Mantenuto dalla Linux Foundation
- Versione attuale è la 2.3 (29 gennaio 2004)
- Successore di FSSTND e `hier`

In pratica: definisce *dove* mettere *cosa*.

**Regola I** tutti i files sono collocati sotto /, anche se memorizzati su dispositivi fisici o virtuali differenti.

**Regola II** non tutte le directory specificate devono essere presenti sul sistema.

**Regola III** tutto è un file e se non è un file allora è un processo. Anche i dispositivi di memoria (cfr. C:\, D:\ di Windows) sono files speciali.

# FHS III

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Le principali distribuzioni sono FHS-compliant, ma presentano alcune piccole deviazioni.

Distribuzione	Riferimento
Debian	<a href="http://debian.org">debian.org</a>
SuSe	<a href="http://novell.com">novell.com</a>
RHEL/Fedora	<a href="http://redhat.com">redhat.com</a>
Ubuntu	<a href="http://ubuntu.com">ubuntu.com</a>

Tabella: Compliance delle principali distribuzioni

Ma non tutte!!

Ad esempio, [GoboLinux](http://gobolinux.org)

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

## 1 Definizione

## 2 Struttura

- Generale
- Eseguibili
- Speciali
- Deviazioni dallo standard

# Gerarchia

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Percorso	Contenuto
/	cartella <i>root</i> , la <i>radice</i> del filesystem.
/boot	files di avvio: bootloader, kernel e initrd.
/etc	files di configurazione del sistema
/home	cartelle personali degli utenti
/media	punto di mount per i filesystem esterni (cd...)
/mnt	punto di mount per i filesystem interni temporanei
/opt	pacchetti software <i>opzionali</i> (concetto elastico)
/root	home dell'utente root
/usr	eseguibili e librerie per le applicazioni utente
/var	dati (mutabili), log e cache di sistema

/usr e /var contengono delle sottogerarchie (vedremo)

# Gerarchia per eseguibili e librerie

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

FHS specifica una gerarchia a *tre livelli*: eseguibili e librerie possono comparire in tre punti del filesystem:

/	essenziali al funzionamento del sistema
/usr	applicazioni utente
/usr/local	applicazioni locali (installate a mano...)

In ciascun percorso, i binari sono suddivisi in 3 cartelle:

.../bin	binari
.../lib	librerie
.../sbin	binari di sistema

## Perché separare bin e sbin?

Di solito, i binari in .../sbin richiedono privilegi di amministratore per funzionare.



# Gerarchia /usr

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Oltre a /sbin, /bin e /lib, contiene

/usr/include	Include files (cfr. *.h del C/C++)
/usr/share	Dati delle applicazioni (immagini, suoni. . .)
/usr/src	Codice sorgente (es. kernel)

A sua volta, /usr/share contiene cartelle standard (consuetudine, non coperte da FHS):

- .../icons
- .../wallpapers
- .../sounds
- .../pixmap (icone e immagini dei programmi)
- .../doc

E altre cartelle (specifiche per i singoli programmi, es. .../amarok/)

# FS speciali

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Non sono vere «cartelle», ma rappresentazioni virtuali di alcuni aspetti del sistema, nello spirito della Regola III (tutto è un file).

/dev	riferimenti ai dispositivi
/proc	informazioni sui processi e sul kernel

Sono il punto di mount di FS opportuni.

# Files temporanei

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

## Regola

se vanno persi, il sistema non ne sentirà la mancanza.  
Sono memorizzati in due posizioni diverse, a seconda del  
tempo di persistenza

/tmp	Cancellata ad ogni riavvio (o anche in corsa)
/var/tmp	Persistente

# Files variabili

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Dati condivisi delle applicazioni durante il normale funzionamento.

/var/cache	Cache di sistema (velocizzano applicazioni)
/var/lib	Informazioni di stato
/var/lock	Lock files
/var/log	Log di sistema (demoni, firewall, accessi. . .)
/var/run	Informazioni sul sistema (utenti connessi. . .)
/var/spool	Coda per le operazioni lunghe (stampa, mail. . .)

## Problema

/var spesso è montato come una partizione separata, quindi potrebbe non essere accessibile nelle prime fasi di avvio.

# Deviazioni dallo standard

Dischi e  
Filesystem

Stefano  
Sanfilippo

Definizione

Struttura

Generale

Eseguibili

Speciali

Deviazioni dallo  
standard

Le moderne distribuzioni Linux creano altre due directory, non standard:

`/run` come `/var/run`, ma montata come filesystem virtuale, **subito accessibile**.

`/sys` contiene i parametri del kernel e le opzioni per configurarne il funzionamento. **NOTA:** sui vecchi unix, era un collegamento al codice sorgente del kernel.

Alcune distribuzioni (cfr. Archlinux e Fedora) appiattiscono la gerarchia a tre livelli degli eseguibili:

- `/lib`, `/sbin` e `/bin` potrebbero essere collegamenti a `/usr/...`

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

3 Dispositivi

4 Partizioni

5 mount e umount

6 fuser

7 Notazione

8 Loop

9 UDev

10 Hard e soft link

11 fstab

12 Organizzazione

13 S.M.A.R.T.

# Dispositivi (file speciali)

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Distinguiamo tre tipi diversi di dispositivo:

**A blocchi** operazioni di I/O per *blocchi* di dimensione predeterminata (**es.** hard disk, che si leggono a settori – p.e. di 4KB)

**A caratteri** operazioni di I/O per singoli byte (o, comunque, senza blocchi di dimensione fissata) (**es.** standard output)

**Speciali** sono «astrazioni», non necessariamente legate ai dispositivi di I/O.

Per ragioni di sicurezza, i dispositivi possono essere creati solo da root e non sono direttamente accessibili ai normali utenti.

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

3 Dispositivi

4 Partizioni

5 mount e umount

6 fuser

7 Notazione

8 Loop

9 UDev

10 Hard e soft link

11 fstab

12 Organizzazione

13 S.M.A.R.T.



# Partizioni

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Sono le suddivisioni del disco. Ogni disco deve avere almeno una partizione, sulla quale si può creare il FS. Esistono due schemi di partizioni, che tengono traccia del punto in cui ogni p. inizia e del tipo di FS che contengono:

**MBR** legacy, la struttura dati precede la prima partizione. La tabella delle partizioni descrive al massimo 4 partizioni fisiche, di 2TiB l'una al massimo. Sono state aggiunte le partizioni logiche, memorizzate in una lista concatenata (se si danneggia il record di una, si perdono i dati delle seguenti)

**GPT** usata su UEFI e su alcuni sistemi BIOS. Innalza i limiti di MBR (128 partizioni, max 8ZiB)

Notare che i due schemi possono coesistere

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount**
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.

# mount

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- Ogni dispositivo ha un file speciale corrispondente nella cartella /dev.
- I dispositivi sono direttamente leggibili e scrivibili (ad esempio, standard input), ma a volte sono disponibili interfacce di livello superiore, specialmente se la lettura/scrittura avviene secondo un protocollo.
- Per accedere ai dati contenuti in un dispositivo a blocchi (i files), bisogna «montarlo».

Il comando per montare i filesystem è **mount**:

```
# mount $nome_device $punto_di_mount
```

Eseguire mount senza argomenti mostrerà i FS attualmente montati, con le relative opzioni.

# umount

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

È il comando simmetrico: serve a «smontare» un dispositivo, in modo da effettuare una rimozione sicura e indolore:

```
# umount $punto_di_mount
```

```
# umount $nome_device
```

**Non è un errore, il comando è umount, non *un*mount!**

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser**
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.

# fuser

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Non è possibile smontare un filesystem occupato (es. c'è un trasferimento di file in corso), in tal caso si usa **fuser** per mostrare il PID dei programmi che stanno usando il dispositivo.

**# fuser /nome/del/file** PID dei processi che hanno aperto *il file* specificato

**# fuser -c \$punto\_di\_mount** PID dei processi che stanno usando un *qualsiasi* file nel dispositivo montato alla posizione specificata.

## Attenzione!

I comandi precedenti vanno eseguiti da root, altrimenti non avranno effetto!

# df

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Indica quanto spazio libero rimane su ogni partizione. Viene spesso usato con l'opzione `-h`, che riporta le misure in unità «*human-readable*»

```
$ df
```

```
$ df -h
```

Lo spazio libero non corrisponde alla differenza tra la capacità massima e la somma della dimensione dei files! (una parte potrebbe essere riservata).

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione**
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.



# Notazione per i dispositivi I

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

I dispositivi compaiono nella cartella `/dev` e hanno *nomi convenzionali*.

Le notazioni riguardano *solo* Linux! (es. BSD usa altre convenzioni).

Dispositivo	Input	Output
<code>/dev/null</code>	Scarta l'input «Disco pieno»	N/A
<code>/dev/zero</code>		Flusso continuo di NULL
<code>/dev/full</code>		
<code>/dev/random</code>	N/A	Flusso <i>truly random</i>
<code>/dev/urandom</code>		Flusso pseudorandom

Tabella: Dispositivi speciali

Altri dispositivi speciali: `/dev/stdin`, `/dev/stdout` e `/dev/stderr`

# Notazione per i dispositivi II

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

I dischi PATA/SATA (la quasi totalità di quelli usati attualmente) usano una prefisso *sd* e una convenzione a due livelli:

`/dev/sda1`



Numero: m-sima partizione sull'n-simo disco

Lettera: n-simo dispositivo (~disco)

`/dev/sda` ← Senza numero finale, n-simo dispositivo

Su MBR, la prima partizione logica è sempre `sdX5`

# Notazione per i dispositivi III

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Framebuffer	/dev/fb
CD-ROM	/dev/cdromX
Lettori Floppy Disk	/dev/fdX
Terminali virtuali	/dev/ptX
Terminali	/dev/ttyX, ttyS (seriale), ttyUSBX
Loop	/dev/loopX

**Tabella:** Altri nomi, con *X* *numero progressivo* del dispositivo

Esistono ancora altri nomi, meno usati (legacy o rari), non li tratteremo (e Google è vostro amico ☺)

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop**
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.

# Loop

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Pseudo-device che rende un file accessibile come se fosse un dispositivo a blocchi.

- Utile per montare immagini di dischi o chiavette senza doverle masterizzare.
- Su Linux, mount si occupa di tutto, deve solo essere informato con l'opzione loop:

```
■ # mount -o loop nome_file punto_di_mount
```

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev**
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.

- Su unix classico, i dispositivi vengono creati con `mknod`.
- Sui sistemi Linux moderni i dispositivi a blocchi sono creati e gestiti automaticamente da un demone, UDev.
  - Udev ascolta gli eventi di sistema (uevent, appunto) e reagisce opportunamente.
  - È un componente articolato, ma non ce ne occuperemo. *Sappiate che esiste e funziona.*

Quando attaccate una chiavetta, viene creato automaticamente `/dev/sdb1`

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link**
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.



# Hard e soft link

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 1 Nel FS, i files sono rappresentati da inode. Un file nel FS è un link ad un inode.
- 2 Un **hard link** crea un altro link allo stesso inode.
  - Quando cancellate un file, viene rimosso un link all'inode. L'inode viene marcato come cancellabile.
- 3 Un **soft link** è un collegamento ad un altro nome nel FS.
- 4 Spostare o rinominare un file non influenzerà un hard link, perché è un collegamento all'inode (e i cambiamenti all'inode si riflettono su tutti gli hard link).
- 5 Gli hard link sono validi solo nello stesso FS, mentre i soft link possono espandersi ad altri FS (essendo semplici riferimenti al nome del file).

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab**
- 12 Organizzazione
- 13 S.M.A.R.T.

# fstab

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

**fstab**

Organizzazione

S.M.A.R.T.

- fstab elenca tutti i dispositivi disponibili sul sistema e indica come vanno inizializzati o integrati nel FS del sistema.
- È usato ancora per la configurazione dell'hard disk principale e i FS di startup, ma **è stato soppiantato** da UDev per i dispositivi esterni.

## Linea di /etc/fstab

```
# device name mpoint fs-type options dump-freq  
pass-num
```

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione**
- 13 S.M.A.R.T.

# Organizzazione

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

Uno schema su disco usuale prevede più partizioni per gestire meglio i dati. Non ci sono regole assolute – dipende dall'uso che la distribuzione fa di ciascuna partizione.

Percorso	Dimensioni consigliate	Motivo primo
/	10~20GB	Applicazioni installate
/boot	50MB~100MB	Numero di kernel
/home	?	Utente
/var	5~10G	Cache varie

**1** «Partizionate» l'uso del disco! Se il server riceve molto spam, mettere /var/mail su una partizione separata aiuta.

**2** Se possibile, lasciare un po' di spazio libero per futuri accostamenti



# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

S.M.A.R.T.

- 3 Dispositivi
- 4 Partizioni
- 5 mount e umount
- 6 fuser
- 7 Notazione
- 8 Loop
- 9 UDev
- 10 Hard e soft link
- 11 fstab
- 12 Organizzazione
- 13 S.M.A.R.T.**

# S.M.A.R.T.

Dischi e  
Filesystem

Stefano  
Sanfilippo

Dispositivi

Partizioni

mount e  
umount

fuser

Notazione

Loop

UDev

Hard e soft  
link

fstab

Organizzazione

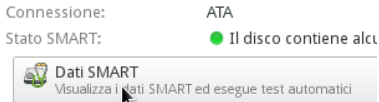
S.M.A.R.T.

I dischi rigidi (quasi tutti) hanno un sistema di diagnostica interna, detto S.M.A.R.T., che fornisce diversi indicatori di affidabilità per anticipare i guasti e salvaguardare i dati.

Alcuni dati forniti da S.M.A.R.T. sono:

- Temperatura
- Numero di settori danneggiati
- Ore di funzionamento

Possono essere visualizzati con strumenti appositi: **smartctl** (CLI), **gsmartcontrol** e **palimpsest** (grafici).



# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**14** FS overview

15 FS virtuali



# Concetti di base I

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**Checksumming** calcolo di codici di controllo per dati e metadati, dà una certa resistenza alle «silent corruptions»

**Deduplicazione** mappatura di copie identiche della stessa risorsa sullo stesso blocco (cfr. COW)

**Copy on Write (COW)** creazione «al volo» di una copia delle risorse condivise, in caso di modifica unilaterale (duplico solo se necessario).

# Concetti di base II

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**Frammentazione** Suddivisione di un file in più blocchi o gruppi di blocchi non contigui. Rallenta I/O su dischi a rotazione (obbliga ad un seek), sfrutta meglio lo spazio (tappa i buchi).

**Deframmentazione** Rimozione della f., ottenuta muovendo i blocchi per rendere contiguo il più grande numero possibile (100% limite teorico).

**Extent** area di memorizzazione contigua (blocchi contigui) riservata per un file. Riduce la frammentazione.

# Concetti di base III

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**Journaling** scrittura su disco per transazioni (come nei DB) per garantire correttezza e consistenza del FS in caso di cadute di tensione.

- Le operazioni (metadati) vengono inserite nel **log** (un'area riservata del FS), poi applicate. Ogni successo/fallimento viene registrato. In caso di spegnimento, Linux analizza il log e prende opportune misure.

# Concetti di base IV

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**Mirroring** copia esatta del contenuto del FS: in caso di rottura del disco, potrà essere sostituito e ripristinato.

**Quota** quantità massima di spazio sul FS usabile da un utente, gruppo, processo (o una combinazione).

**Wear leveling (TRIM)** distribuzione dell'uso delle celle di memoria su un SSD, per bilanciare l'usura della flash. Viene effettuata dall'hardware, ma ha bisogno della collaborazione del SO (che deve sapere e indicare quali blocchi non sono più in uso).

# Principali FS

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

In uso attuale, sono diffusi 4 filesystem:

Nome	Stato	Journal	COW	C'sum	Dedupe	Defrag	Mirror	Quota
ext4	stabile	✓	~	✗	✗	✓	✗	✓
btrfs	«testing»	✓	✓	✓	✓	~	✓	~
zfs	stabile	✓	✓	✓	✓	✓	✓	✓
vfat	legacy	✗	✗	✗	✗	✗	✗	✗
ntfs	legacy	✓	✗	?	✗	~	~	✓

**VFAT DEVE essere EVITATO** su qualsiasi disco fisso, ma è utile per chiavette (compatibile con Windows): no Journal!!

# Note

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

- 1 ZFS non può essere distribuito assieme al Kernel di Linux per questione di licenza. Va installato a parte.
- 2 EXT4 è stabile. BTRFS (che lo rimpiazzerà nel futuro prossimo) è ancora dichiarato sperimentale, ma il formato su disco è stabile nella serie 3.x.x (salvo «bug che rendano la modifica assolutamente necessaria»). Se i dati memorizzati non sono essenziali (es. backup), è consigliato provarlo.
- 3 BTRFS e ZFS offrono circa le stesse caratteristiche, ma ZFS è nato nel mondo OpenSolaris, BTRFS in quello Linux.
- 4 Un volume EXT4 può essere facilmente convertito in BTRFS: è possibile installare EXT4 e, in futuro, convertire.
- 5 BTRFS e ZFS supportano la compressione trasparente, ma è consigliata solo in sistemi

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

14 FS overview

15 FS virtuali

# FS speciali

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

- procfs** montato in `/proc`, fornisce informazioni sui processi attivi, sotto forma di pseudo-FS (es. `/proc/$PID`)
- sysfs** montato in `/sys`, permette di modificare i parametri del kernel e apprendere la configurazione del sistema.
- tmpfs** FS creato nella RAM, utile per `/run`, `/var/lock` e altre strutture «piccole» (RAM limitata) che hanno bisogno di velocità ma non devono resistere al reboot
- aufs** raggruppa più FS, facendoli apparire come un solo FS virtuale. Successore di `unionfs`.



# Altri FS

Dischi e  
Filesystem

Stefano  
Sanfilippo

FS overview

FS virtuali

**squashfs** FS compresso in sola lettura. Viene spesso usato nei Live CD, facendo risparmiare spazio.

**ISO 9660 (e UDF)** FS in sola lettura usato nei CD dati. UDF è il successore (usato nei CD-RW e nei DVD)

**eCryptfs** overlay che supporta la cifratura trasparente a livello di singoli file e path (es. si può abilitare solo per una cartella).

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

16 LUKS

17 LVM2

18 RAID

# LUKS

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

Permette la cifratura del disco. Una partizione LUKS può contenere qualsiasi filesystem, la cifratura avviene al livello dei singoli blocchi (al contrario di eCryptFS).

La modalità di funzionamento consigliata è AES256 (512 bit), XTS-plain64.

# Perché cifrare?

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

- 1 In caso di furto del laptop, nessuno (CIA?) potrà leggerne il contenuto
  - fotografie, email, etc. sono al sicuro
  - account e password salvate pure!
- 2 Niente bisogno di cancellare il disco se dovesse andare in «manutenzione» o essere rivenduto.
- 3 Riduce la possibilità di manomissione criminale
  - injecting limitato al bootloader/immagine del kernel.
- 4 In alcuni casi (studi medici, etc.) è un obbligo di Legge.
- 5 Protegge i dati sui vostri server (probabilmente ospitati in una server farm) da sguardi indiscreti.

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

16 LUKS

17 LVM2

18 RAID

# Logical Volume Mapper (LVM2)

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

Mappa i blocchi fisici di uno o più partizioni fisiche (di tipo LVM, appunto) in un unico spazio di blocchi di partizioni «virtuali». La traduzione fisico–virtuale avviene a livello del kernel. Permette quindi di:

- Ridimensionare le partizioni a sistema acceso (dipende anche dal FS)
- Fondere più dischi (partizioni) fisici, creandone uno virtuale di dimensione pari alla somma
- Fare striping, ovvero distribuire i blocchi dei files su più dischi, aumentando il throughput
- Mirroring (snapshotting), indipendentemente dal FS sulle partizioni LVM2.

BTRFS e ZFS offrono nativamente queste caratteristiche, in tal caso l'uso di LVM diventa opzionale.

# Outline

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

16 LUKS

17 LVM2

18 RAID

# RAID

Dischi e  
Filesystem

Stefano  
Sanfilippo

LUKS

LVM2

RAID

*Redundant Array of Independent Disks*: distribuisce i dati su più dischi facendoli figurare come un unico disco virtuale, in modo che il guasto di un dispositivo danneggi minimamente il contenuto del disco virtuale. Permette di realizzare uno storage persistente.

Esistono più configurazioni RAID, identificate da un numero (RAID-1, RAID-3, RAID-5...). La distribuzione dei dati viene effettuata a livello di byte o blocco su un numero (minimo) variabile di dischi, dipendente dalla configurazione RAID.

I dischi devono essere della stessa dimensione – altrimenti viene presa a riferimento la dimensione del più piccolo e lo spazio in eccesso sugli altri è ignorato: spreco.



# Licenza

Dischi e  
Filesystem

Stefano  
Sanfilippo

Questa opera è licenziata sotto i termini della Creative Commons, Attribuzione – Non commerciale – Condividi allo stesso modo 3.0 Italia License.

Per una copia della licenza, consultare

<http://creativecommons.org/licenses/by-nc-sa/3.0/it/>.

