

# Backup

Cosa, Come, Quando e Perché effettuare il Backup di sistema

Andrea Grazioso  
grazioandre@gmail.com

Corsi Gnu/Linux Avanzati 2013/2014



# Cos'è il backup?

*There are two kinds of people: those who do regular backups and those who never had a hard drive failure — Unknown.*

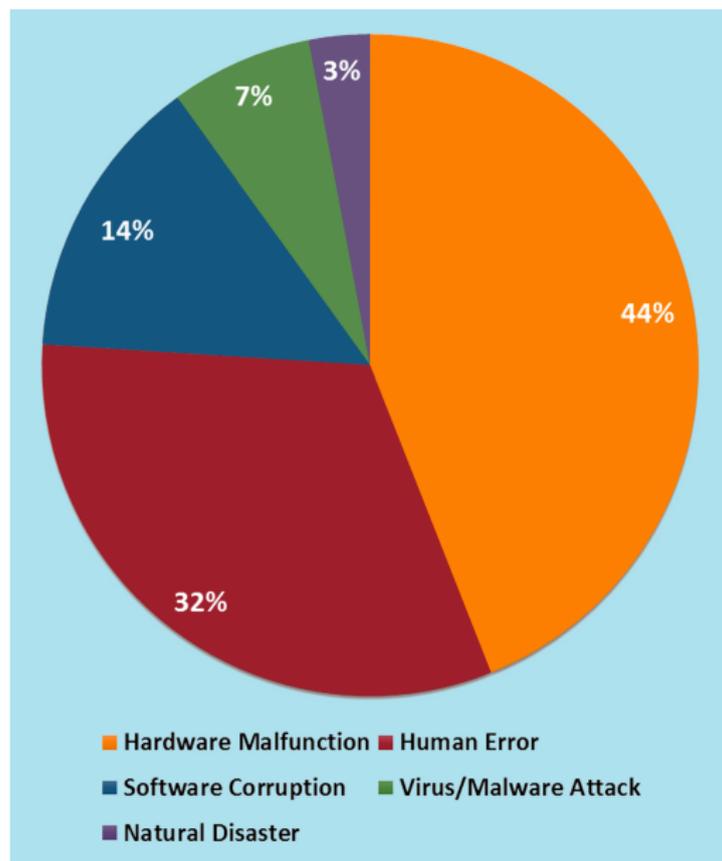
Con il termine «backup» si intende replicazione su supporti di memorizzazione di materiale informativo archiviato nella memoria di massa del PC al fine di prevenire perdite di dati a seguito di eventi malevoli o accidentali.

# Cos'è il backup?

*There are two kinds of people: those who do regular backups and those who never had a hard drive failure — Unknown.*

Con il termine «backup» si intende replicazione su supporti di memorizzazione di materiale informativo archiviato nella memoria di massa del PC al fine di prevenire perdite di dati a seguito di eventi malevoli o accidentali.

# Cosa minaccia i nostri dati?



# Supporti: CD, DVD e Blu-ray

Eredi dei floppy disk questi supporti ottici sono diventati i principali supporti per la vendita in campo audio, video, videoludico, software. Si presentano in vari formati ognuno con caratteristiche diverse a seconda del campo di applicazione per cui sono stati progettati

## Pro

Ottimali per il backup di dati

Disponibili in vari formati (CD, DVD Blu-ray)

Portatili

Resistenti ai danni superficiali, come al magnetismo e all'azione dell'acqua

Disponibili riscrivibili e non

Elevata longevità

## Contro

Necessitano di un drive apposito per essere utilizzati

Urti e scossoni durante la scrittura potrebbero rendere il drive illeggibile

# Optical Media Comparison

The logo for Compact Disc, featuring the word "COMPACT" in a small font above the word "disc" in a large, stylized, outlined font.The logo for DVD, featuring the letters "DVD" in a bold, blue, sans-serif font above a stylized blue disc icon.The logo for Blu-ray Disc, featuring a stylized blue "b" shape above the text "Blu-ray Disc" in a blue, sans-serif font.

	Dim Max	Lifetime	Min Write Speed	Max Write Speed
CD	700MB	10	1x=150Kib/s	56x=8400KiB/s
DVD	4.7GB	30	1x=1.32Mib/s	24x=32.46MiB/s
Blu-ray	200GB	50	1x=4.29Mib/s	16x=68.66MiB/s

## Nota Bene

$1\text{MiB} = 1,024\text{MB}$

# Supporti: Flash Drive

Dalle dimensioni ridotte e discreta capacità questi supporti hanno decisamente spopolato nel corso degli ultimi anni grazie alla diffusione dello standard USB; composti esclusivamente da materiale elettronico, è un peccato che siano scarsamente affidabili, ciò non le rende consigliate per il backup di dati.

## Pro

Non necessitano di driver per l'uso

Buona velocità di trasferimento

Estrema versatilità e portabilità

Resistenti agli urti

Buona longevità di circa 100'000 cicli di scrittura

## Contro

Costo per GB piuttosto elevato

Facili da perdere

Scarsamente affidabili

# Supporti: Hard Disk

Costituiti internamente da dischi magnetici in rapida rotazione attorno ad un asse comune, letti da una testina messa in moto da un attuatore questi sistemi sono diventati i principali supporti di memorizzazione di massa in PC, laptop, server e console. Ad oggi il costo di un GB negli hard disk è molto basso, circa \$0.05/GB e la loro durata è influenzata dal tipo di utilizzo che ne è stato fatto.

## Pro

Elevatissima capacità

Buona velocità di trasmissione

Non necessitano di driver esterni per l'uso

Compatibili con molte interfacce (PATA, SATA, USB, Firewire, eSATA)

## Contro

Necessitano di rimanere il più possibile fermi durante l'uso

Per essere utilizzati al meglio occorre avere conoscenze sulla gestione di dischi e di filesystem

# Supporti: Solid State Drive

Gli SSD sono gli ultimi tipi di memorie ad essere approdate sul mercato consumer. Composti da circuiti integrati che fungono da memoria per l'archiviazione dei dati in questi dispositivi non c'è nessuna parte in movimento, ciò, unito alle capacità dell'interfaccia SATA li rende estremamente performanti. Purtroppo sono anche maggiormente soggetti a guasti rispetto agli HDD, perciò questi ultimi sono ancora da preferire nel caso si voglia archiviare dati a lungo termine.

## Pro

- Grande capienza

- Elevatissima velocità di trasferimento

- Resistenti ad urti e movimenti durante l'uso

- Molto leggeri e silenziosi, adatti all'uso laptop

## Contro

- Costo elevato

- Bassa aspettativa di vita

In alternativa a device fisici di cui occuparsi personalmente si può optare per affidarsi a servizi di archiviazione/backup online che si caratterizzano per funzionalità aggiuntive come ridondanza dei dati in luoghi multipli, protezione dei dati mediante RAID dei dischi, crittografia, banda illimitata in upload, ed un servizio generalmente scalabile.

- Amazon Glacier
- OVH cloud storage
- Backblaze

Alla luce di quanto visto, quale supporto scegliere per il backup dei propri dati?

La risposta non è così banale...

Alla luce di quanto visto, quale supporto scegliere per il backup dei propri dati?

La risposta non è così banale...

E' il cosiddetto Snapshot che salva lo stato esatto del sistema in un particolare istante di tempo.

Lo scopo di questa tecnica è copiare bit per bit tutte le informazioni contenute nel disco sorgente ricostruendo così l'esatta struttura del sistema

- dd
- rsnapshot

Da semplice copia di file a backup strutturato, questo metodo è ottimo per gli utenti del sistema che vogliono salvare esclusivamente la propria home come per gli amministratori che vogliono un backup compatto dell'intero sistema.

- tar
- rsync
- duplicity

Possibili aperture a tecniche avanzate

Complementari ai metodi appena visti vi sono tre tecniche avanzate che è indispensabile conoscere

**Incrementale**, aggiunge all'ultimo backup, il materiale ha subito dei cambiamenti dall'ultimo backup o incremento. Più sicuro di quello differenziale ma necessita di attenzione nel ripristino

**Differenziale**, modifica di volta in volta il backup precedente con i file che hanno subito dei cambiamenti. Più veloce dell'incrementale e meno ingombrante, ma mantiene solo l'ultima versione dei file.

## Cose che dovrete sapere su

- /tmp
- /proc /sys
- /etc /var
- /dev /mnt /media

## Altre raccomandazioni

- Attenzione ai vostri file
- Ridurre gli sprechi
- Live media
- Cron

dd è un utility di copia e conversione dati. Se non altrimenti specificato dd si limita a copiare dati da una sorgente alla destinazione.

```
dd if=/dev/sdx of=/dev/sdy
```

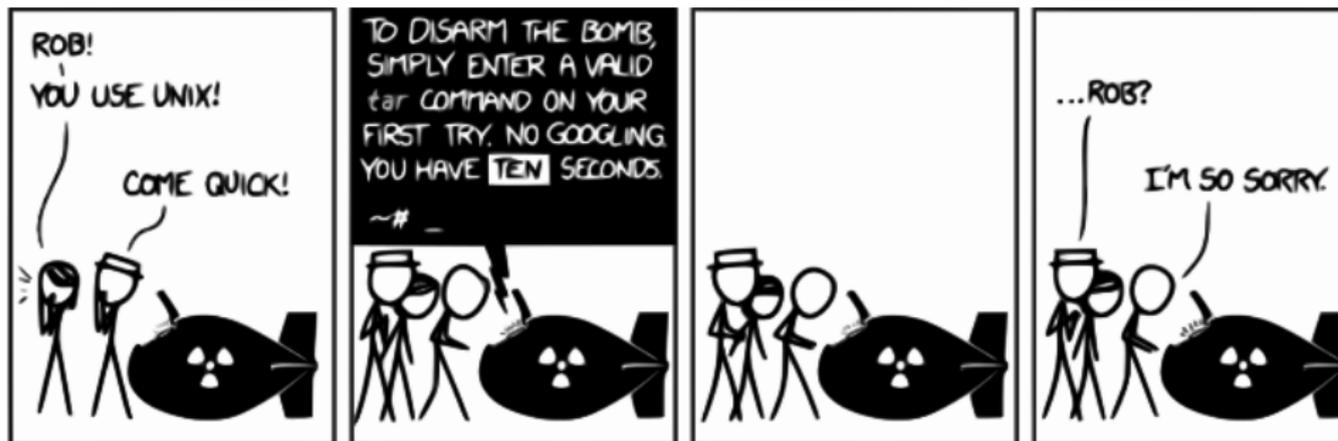
`if` - input file, cioè la sorgente

`of` - output file, cioè la destinazione

quando si usa dd è importante assicurarsi che si opera sulle giuste partizioni/file poichè questo comando è potenzialmente distruttivo.

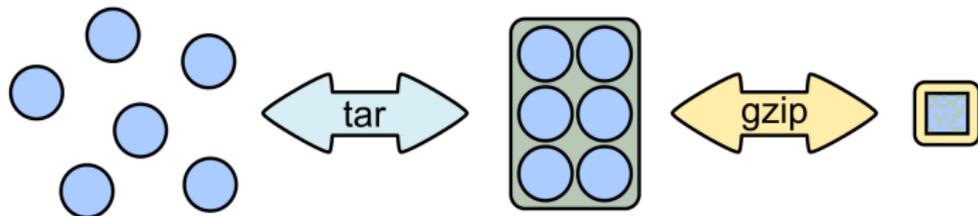
### nota

Per controllare che le partizioni siano quelle corrette ci si può servire di tools come fdisk o gparted.



Inizialmente sviluppato per scrivere dati sequenzialmente (per effettuare backup sui nastri magnetici), tar è ora un utility per riunire in un unico blocco tanti file per distribuzione/archiviazione, preservando allo stesso tempo informazioni sui permessi e sulla struttura delle cartelle.

# Tar: funzionamento



## pro

Permette di salvare i file in un unico blocco...

...ed eventualmente comprimerlo

Rispetta la struttura delle directory

Salva i permessi dei file

## contro

Non conserva informazioni sul filesystem di partenza

Non segue i symlink (Attenzione a /proc e simili)

Utilizzare algoritmi di compressione può richiedere molta potenza di calcolo

```
tar -cvpzf /media/BackupStorage/backup01-03-14.tar.gz  
--exclude=/media /
```

**c** - crea un nuovo archivio.

**v** - verbose mode.

**p** - preserva i permessi dei file.

**z** - comprime i file usando l'algoritmo di gzip.

**f** - Specifica il path assoluto nel quale salvare il file.

`/media/BackupStorage/backup01-03-14.tar.gz` - Specifica il path in cui salvare il backup.

`-exclude=/path` - Seleziona i percorsi da escludere.

`/` - directory di partenza per il backup, che include tutte le relative sottocartelle.

Rsync è un comando che permette la copia di file su una destinazione fisica eventualmente tentando di copiare solo le differenze. Noto anche per essere la versione avanzata di cp

```
rsync -Pr source destination
```

**P** - indica a rsync di mantenere i file parzialmente trasferiti e di mostrare la barra dell'avanzamento durante il processo di trasferimento

**r** - opzione per la ricorsività nelle cartelle

Altre opzioni comuni di utilizzo per rsync:

```
rsync -aAXv /* /backup_folder --exclude={...}
```

```
rsync -avz -e ssh
```

```
remoteuser@remotehost:/remote/dir /this/dir/
```

# Utilizzo di ssh

```
dd if=/dev/sda | ssh username@servername.net  
"dd of=/file"
```

```
tar -cvpz <...> / | ssh user@backuphost
```

```
"( cat > ssh_backup01-03-14.tar.gz )"
```

| - pipe, indica che l'output del comando a sinistra verrà reindirizzato al seguente

**ssh** - Secure Shell, tramite il quale ci si collega a "backuphost" come "user"

**cat** - utility di concatenamento di file

> - reindirizzamento dell'output

[ssh-backup01-03-14.tar.gz](#) - destinazione

## Nota

Ricordati di settare chiavi ssh e permessi sia su localhost che su remote host!

```
rdiff-backup dir1 user@system:~/dir2
rdiff-backup dir1 dir2
```

rdiff-backup è un utility in circolazione ormai da molto tempo che permette il backup in locale o anche in remoto utilizzando il metodo incrementale.

## Pro

preserva la struttura delle cartelle, links, permessi, attributi estesi, metadati.

ottimizzato per la rete

facile da utilizzare

permette backup incrementale

mantiene dei log

# Duplicity

In merito all' «eccessiva semplicità» di rsync sono nati dubbi e perplessità nel momento in cui si è cominciato ad archiviare sui propri PC dati sensibili, ed ecco quindi nascere duplicity, che integra perfettamente tutte le funzionalità di rsync aggiungendo opzioni per la sicurezza e la crittografia.

## Pro

- Molto semplice da utilizzare
- Permette il backup incrementale
- Ottimizzato per il backup
- Permette di cifrare i dati salvati

## Contro

- Pensato per il backup di singole macchine

# Duplicity Usage

```
duplicity /home/me sftp://uid@other.host//some_dir
```

`/home/me` - percorso da includere nel backup con tutte le relative sottocartelle

`sftp://uid@other.host//some_dir` - destinazione sottoforma di indirizzo sftp, nello specifico ci si collega ad “other.host” come user “uid”, e il backup avrà posto in “some\_dir”

```
duplicity restore sftp://uid@other.host//dir /home/me
```

# Duplicity Usage Extended

```
duplicity -t 3D restore
```

```
sftp://uid@other.host//some_dir /home/me
```

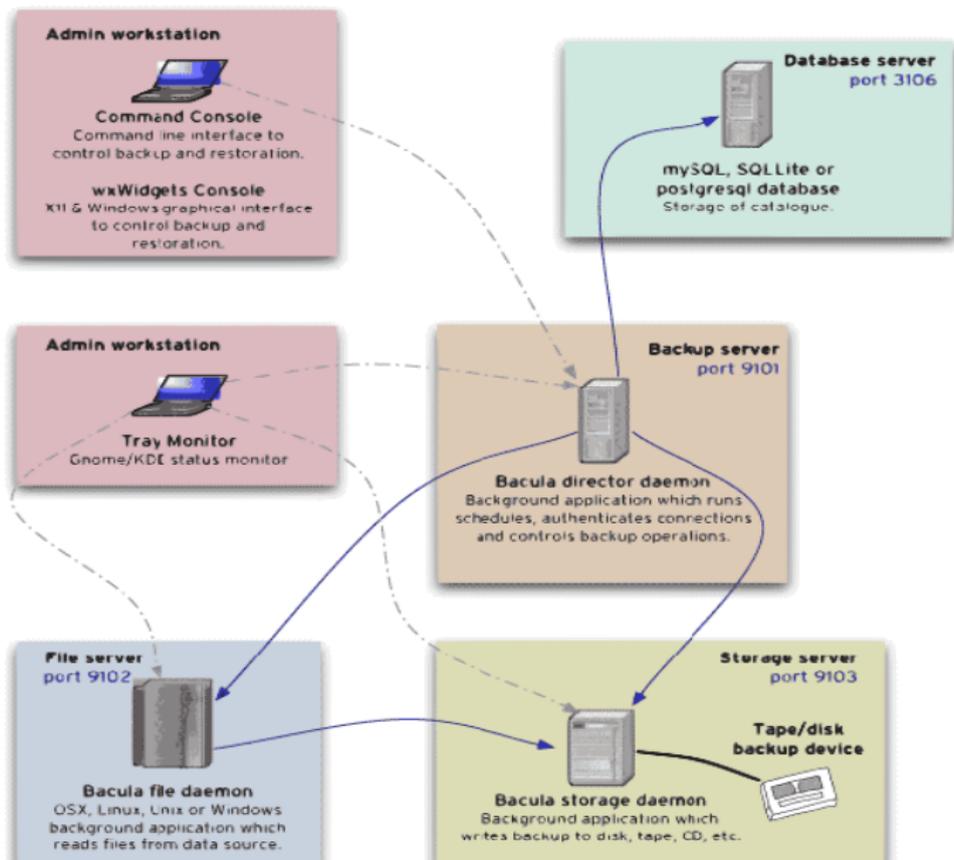
**xD** - specifica la necessità di recuperare dei dati di x giorni prima che, ad esempio, sono stati cancellati per errore e poi è stato lanciato il comando di duplicity per avviare il backup

Infine è possibile eliminare dati che oramai non servono più se avessimo bisogno di spazio per nuovi backup:

```
duplicity remove-older-than 30D
```

```
sftp://uid@other.host//some_dir
```

# Bacula e Amanda



- Backup Handbook
- <http://en.wikipedia.org/>
- <https://help.ubuntu.com/community/BackupYourSystem/TAR>
- <https://wiki.archlinux.org/index.php/>
- <http://www.tomshardware.com/reviews/usb-3-uas-turbo,3215-2.html>
- <http://data-backup-software-review.toptenreviews.com/removable-backup-media.html>
- <http://www.nongnu.org/rdiff-backup>
- Giornalinux 2.0 N14 - [www.poul.org](http://www.poul.org)
- <http://www.bacula.org/en/>
- <http://amanda.org/>
- <http://aws.amazon.com>
- Unix and Linux system administration handbook, fourth edition, Evi Nemeth, Garth Snyder, Trent R. Hein, Ben Whaley

# Grazie per l'attenzione!

## Grazie per l'attenzione!



Queste slides sono licenziate Creative Commons Attribution-ShareAlike 4.0

<http://www.poul.org>