

# Protezione del Filesystem

**Architetture SCSI**  
**Filesystem Journaling**  
**Sistemi RAID**  
**Storage Area Network**  
**Clustering**

Paolo PAVAN

Anno 2002

## Sistemi di Protezione dei dati

- **SCSI**: alte prestazioni, maggiore affidabilità e durata dei supporti Hardware (drive)
- **FS Journaling**: protezione dell'integrità dei dati dall'interruzione di corrente
- **RAID**: ridondanza dei dati e maggiori prestazioni
- **SAN**: maggiori volumi e maggiori prestazioni in accesso ai dati
- **Cluster**: ridondanza dei sistemi, grandi potenze di calcolo.

## Architetture SCSI

- SCSI sta per Small Computer System Interface è identifica uno standard, per l'interconnessione di più dispositivi (dischi, CDROM o DAT) ad un computer
- Lo SCSI nasce in effetti come un BUS (insieme di regole e protocolli) per far dialogare una componente centrale (detto Host Adapter ed a volte controller) con delle unità di memoria secondaria, cioè gli hard disk
- La connessione avviene su di una "catena" detta catena SCSI la quale deve necessariamente concludersi con un terminatore
- Fisicamente i dispositivi si collegano mediante attacchi idonei a cavi tipo flat specifici più larghi di quelli ide
- La catena è controllata da uno più controller SCSI o Wide SCSI

## Architetture SCSI come Periferiche per il RAID

- Si utilizzano per il RAID periferiche SCSI anche se si stanno diffondendo sistemi integrati sulla MB del sistema per il RAID IDE, di dubbia funzionalità
- Interfaccia locale, ad alta velocità per il collegamento di dispositivi periferici (dischi, stampanti, scanner,...) a server e workstation
  - Esistono diverse versioni
    - SCSI-1, SCSI-2, SCSI-3 (versioni dello SCSI)
    - UltraSCSI, WideSCSI, FastSCSI (incremento delle prestazioni delle singole versioni)

# Tabella delle Periferiche SCSI

	MaxBusSpeed	BusWidth	MaxBusLength (m)			MaxDevSupport
			SE	LVD	HVD	
<b>SCSI-1</b>	<b>5 MB/s</b>	<b>8 bit</b>	<b>6</b>	<b>-</b>	<b>25</b>	<b>8</b>
<b>Fast SCSI</b>	<b>10 MB/s</b>	<b>8 bit</b>	<b>3</b>	<b>-</b>	<b>25</b>	<b>8</b>
<b>Fast Wide SCSI</b>	<b>20 MB/s</b>	<b>16 bit</b>	<b>3</b>	<b>-</b>	<b>25</b>	<b>16</b>
<b>Ultra SCSI</b>	<b>20 MB/s</b>	<b>8 bit</b>	<b>1.5</b>	<b>-</b>	<b>25</b>	<b>8</b>
<b>Ultra SCSI</b>	<b>20 MB/s</b>	<b>8 bit</b>	<b>3</b>	<b>-</b>	<b>-</b>	<b>4</b>
<b>Wide Ultra SCSI</b>	<b>40 MB/s</b>	<b>16 bit</b>	<b>-</b>	<b>-</b>	<b>25</b>	<b>16</b>
<b>Wide Ultra SCSI</b>	<b>40 MB/s</b>	<b>16 bit</b>	<b>1.5</b>	<b>-</b>	<b>-</b>	<b>8</b>
<b>Wide Ultra SCSI</b>	<b>40 MB/s</b>	<b>16 bit</b>	<b>3</b>	<b>-</b>	<b>-</b>	<b>4</b>
<b>Ultra2 SCSI</b>	<b>40 MB/s</b>	<b>8 bit</b>	<b>-</b>	<b>12</b>	<b>25</b>	<b>8</b>
<b>Wide Ultra2 SCSI</b>	<b>80 MB/s</b>	<b>16 bit</b>	<b>-</b>	<b>12</b>	<b>25</b>	<b>16</b>
<b>Ultra3 SCSI</b>	<b>160 MB/s</b>	<b>16 bit</b>	<b>-</b>	<b>12</b>	<b>-</b>	<b>16</b>

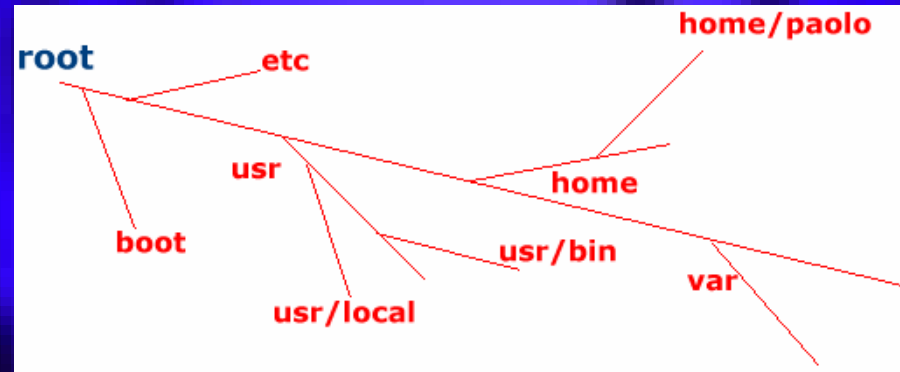
## Peculiarità dei sistemi SCSI

- Ogni dispositivo viene collegato ad una catena, che può ospitare anche fino a 16 dispositivi (contro i 4 al massimo dell'architettura IDE)
- Ogni dispositivo ha uno specifico ID
- I costi delle periferiche più il controller sono molto alti.

## Drive SCSI

- Prodotti un po' da tutte le case tra cui
  - IBM
  - Seagate
  - Quantum
- Dischi molto più capienti (80 MB) ma molto più costosi, un singolo disco può costare 400/1000 Euro.

# Filesystem



- Il **filesystem** rappresenta il modo con cui sono organizzati i dati all'interno di un disco o di una sua partizione. Nei sistemi operativi Unix non esiste la possibilità di distinguere tra un'unità di memorizzazione e un'altra, come avviene nel Dos, in cui ogni disco o partizione sono contrassegnati da una lettera dell'alfabeto (A:, B:, C:). Nei sistemi Unix, tutti i filesystem cui si vuole poter accedere **devono essere concatenati assieme**, in modo da formare un unico filesystem globale. Sotto unix si dice che un file system viene **montato o innestato** su quello principale.



## Data e Metadata

- Il filesystem esprime il metodo con cui i nostri dati vengono organizzati, letti e scritti sui supporti hardware (dischi).
- L'organizzazione di questi dati dipende dalla struttura interna del filesystem: la struttura interna di organizzazione dei dati (data) viene chiamata meta-data.
  - Data: i dati effettivamente contenuti (file)
  - Metadata: contenitori dei file, la struttura interna di organizzazione dei dati.

## Journaling Filesystem

- Il suo funzionamento è semplice prima di ogni operazione di scrittura vera e propria, un journaling filesystem si occupa di tenere un "**giornale**" dove vengono segnate le modifiche che si apprestano ad essere fatte sul filesystem.
- In caso di blocco del sistema il programma di controllo (fsck) controllerà solo il "giornale", tenuto dal sistema quindi **soltanto la parte di meta-data** riguardante le ultime modifiche presenti nel journal. Il tempo richiesto è talmente basso da non essere percepibile, se non per via di qualche scritta nei che appare al riavvio log di sistema.
- Tipi di filesystem journaling:
  - JFS
  - ext3
  - ReiserFS

# Funzionamento del Journaling sotto Linux (ext3)

- In caso di interruzione di corrente il sistema riparte verificando (fsck) solo il journal e non tutto il disco.
- L'operazione è impercettibile
- Partenza con supporto Journaling (ext3) dopo interruzione di corrente

## Partenza dopo shutdown "pulito"

**EXT3-fs: mounted filesystem with ordered data mode.**

VFS: Mounted root (ext3 filesystem) readonly.

## Partenza dopo shutdown "sporco"

**EXT3-fs: INFO: recovery required on readonly filesystem.**

**EXT3-fs: write access will be enabled during recovery.**

*(recovery.c, 253): journal\_recover: JBD: recovery, exit status 0, recovered transactions  
2 to 85*

*(recovery.c, 255): journal\_recover: JBD: Replayed 773 and revoked 0/0 blocks  
kjournald starting. Commit interval 5 seconds*

**EXT3-fs: recovery complete.**

**EXT3-fs: mounted filesystem with ordered data mode.**

## Redundant Array of Inexpensive disk Vantaggi del Raid

- Raid è l'acronimo di Redundant array of Inexpensive (o Independent) disk. In pratica, si tratta di una tecnologia che consente di amministrare un insieme di dischi rigidi in modo tale da aumentarne considerevolmente l'affidabilità e le prestazioni, sprecando solo un po' di spazio.
- Esistono sei "livelli" Raid base; ciascun livello costituisce una diversa implementazione della tecnologia e ha una diversa finalità. Ciò che accomuna ogni sistema Raid, di ogni livello, è che un insieme (array) di dischi è visto dall'utente come un singolo disco rigido, mentre la sofisticata tecnica di gestione rimane completamente trasparente.
- La sicurezza dati consiste nell'averli memorizzati su più dischi in modo da poterli ricostruire in caso di anomalie di uno dei dischi del sistema.

La velocità è data dai dati che vengono letti e scritti in più dispositivi.

## Origini del RAID

- Il RAID nasce nel 1987 quando i ricercatori Patterson, Gybson e Katz della Berkely university of California pubblicano il documento "A case for Redundant Array of Inexpensive Disks (RAID)" il cui scopo era di creare una tecnologia che permettesse di unire un gruppo di piccoli e poco costosi (inexpensive) dischi in una matrice logica (array) in grado di superare le performance di un unico disco più capiente, veloce e quindi costoso.
- Le cinque diverse architetture di array descritte nel progetto iniziale oltre al primario obiettivo prestazionale, includevano capacità di fault tolerance adatte alla protezione dei dati. Queste caratteristiche vengono riassunte nel termine ridondanza applicato all'insieme dei dischi.

## Cosa serve per fare un Raid?

- Raid Hardware
- Raid Software
- Controller RAID
  - SCSI
  - IDE
  - Dischi SCSI o IDE
  - Cavi adatti e terminatori
  - IDE solo mirroring

## Caratteristiche del RAID

- **REALIABILITY:** affidabilità. Esprime il momento o la frequenza con cui si manifesta un guasto ed è oltremodo conosciuta come MTBF (Mean Time Before Failure) la cui unità di misura è l'ora. L'affidabilità la si ottiene grazie alla ridondanza.
- **AVAILABILITY:** disponibilità. Indica l'accessibilità dei dati cioè la possibilità di leggerli e scriverli, e la disponibilità di farlo full time è un obiettivo importantissimo, definito dall'integrità dei dati e dalla capacità di fault tolerance.
- **SERVICEABILITY:** timbrare il cartellino. Rappresenta l'entrata in servizio in caso di guasto e viene valutata rispetto al tempo e alla fatica impiegati nell'individuare e riparare il guasto e riprendere le normali attività. Si distinguono tecniche di recovery e di monitoring dell'array.
- **PERFORMANCE:** grandi prestazioni in termini di lettura e scrittura con il metodo dello striping.

## Array, Stipes e Parità

- L'**array** è l'insieme di dischi che costituiscono il volume per lo storage, deve essere creato e inizializzato dal software di controllo del controller.
- Con il termine **stripe** si identifica una striscia ovvero una porzione in cui viene suddiviso un disco in cui vengono salvati dei dati
- **Parità**: le stringhe di parità rappresentano sono una sorta di "mappa" che permette di ricostruire i dati in caso di anomalia ad uno dei suddetti dischi.



# Architetture del RAID

## ARCHITETTURE RAID

Il Berkely paper definisce cinque tipologie di implementazione RAID a cui si aggiunge una tecnica di striping non ridondante detta RAID 0.

- RAID1
- RAID2
- RAID3
- RAID4
- RAID5

## Altri tipi di RAID

**RAID 0/1 (anche definito RAID 10).** E' l'insieme di RAID 0 e RAID 1, per cui i dischi in mirroring vengono uniti logicamente per dare un'unica unità più grande. RAID 0/1 offre buone performance e fault tolerance.

**RAID 0/5 (anche definito RAID 50).** In questa implementazione i dischi vengono uniti per sembrare un'unica unità dovuta all'unione di più volumi in Raid 5 in un'unica unità (striping - RAID 0)

**RAID 0/3 (anche definito RAID 30).** Non fa parte dello standard, ma è configurabile su alcuni controller (American megatrends, per esempio) e unisce in modo logico dischi in RAID 3.

## Livelli del RAID

- **RAID 0 - Striping**  
Disk Striping divide i dati in blocchi di 64k e li distribuisce in ugual misura su tutti i dischi nell'array.  
**Non è fault tolerant.**
- **RAID 1 - Mirroring e Duplexing**  
Disk Mirroring: Duplica una partizione in un altro disco fisico connesso allo stesso controller.  
Disk Duplexing: Duplica una partizione in un altro disco fisico che è connesso ad un altro controller.
- **RAID 2**  
Disk Striping w/ ECC. I blocchi di dati e le informazioni di parità vengono distribuiti per tutti gli hard-disk dell'array con il controllo degli errori (error checking).
- **RAID 3**  
Disk Striping w/ ECC I dati vengono distribuiti per tutti gli hard-disk dell'array ma usa un solo hard-disk per la memorizzazione delle informazioni di parità.
- **RAID 4**  
Disk Striping con blocchi grandi. I blocchi di dati vengono distribuiti per tutti gli hard-disk usando grandi blocchi.
- **RAID 5 - Striping con parità**  
Suddivide i dati e le informazioni di parità in modo uniforme su tutti i dischi.
- Windows NT supporta il RAID 0, 1 e 5.
- Unix supporta tutti i livelli di RAID

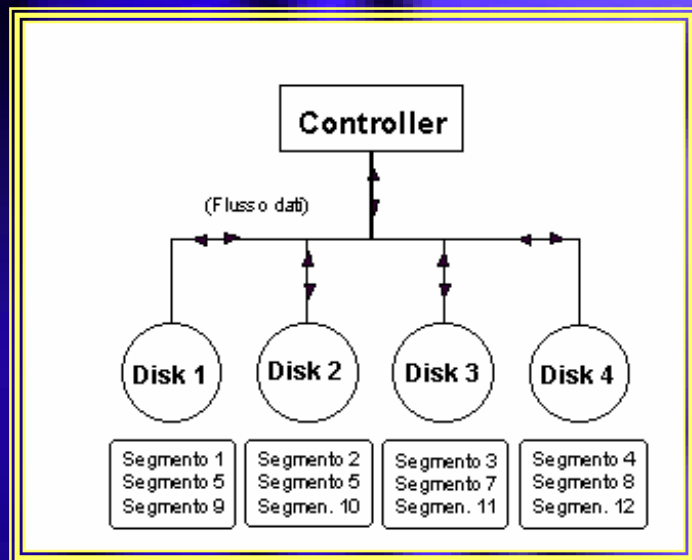
## RAID 0: Disk Striping

Il livello 0 della tecnologia Raid è lo "striping", un metodo per concatenare una serie di volumi fisici in una singola unità logica

Per ottenere le massime prestazioni, è necessario che il carico di I/O sia bilanciato su tutti i dischi di un sistema a unità multiple; lo striping consiste nel partizionare ogni disco in blocchi sequenziali (la dimensione dello stripe è determinante per le prestazioni) in modo tale da ripartire ciascun dato da memorizzare su tutti i dischi componenti lo stripe-set in modo equo. Il livello Raid 0 è l'implementazione pratica di questa tecnica, ma siccome non si registrano le informazioni di parità (Xor) per la ridondanza dei dati, è improprio definirla una vera e propria soluzione Raid: in parole povere, con questo metodo non è soddisfatta una delle prime peculiarità dei sistemi Raid; infatti, se uno qualsiasi dei dischi componenti lo stripe-set si rompe, si perdono irreparabilmente tutti i dati contenuti nei dischi.

Dunque, con il solo raid 0 non c'è alcun vantaggio circa l'affidabilità del sistema, anzi, il Mtbf (Mean time between failure, tempo medio prima di un guasto) di un array è uguale a quello del singolo disco diviso il numero dei dischi dell'array. Il grande pregio del Raid 0 è invece legato allo **strepitoso aumento delle prestazioni che ne consegue**. Per realizzare il raid 0 bastano due soli dischi, ma usandone di più si ottengono prestazioni migliori.

## RAID 0: Schema



I vari stripes vengono suddivisi sui diversi dischi in modo sequenziale.

Perdere un disco significa non poter più recuperare i segmenti contenuti e quindi significa perdere l'intero array.

**VANTAGGI:**

Aumenta le prestazioni del sistema di tante volte quanti sono i dischi dell'array

**SVANTAGGI:**

Nessuna protezione per i dati.

## RAID 1: Disk Mirroring

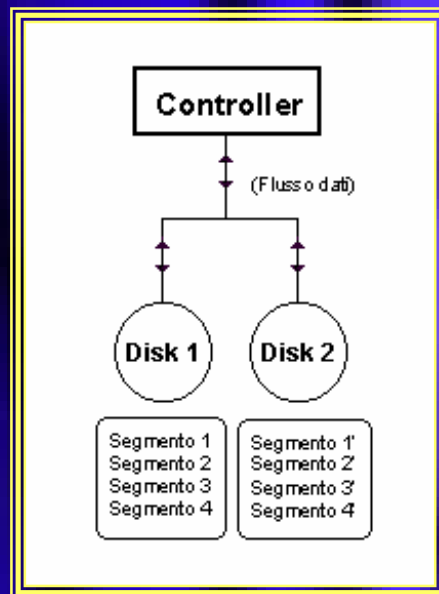
Definito anche "disk mirroring" è essenzialmente un metodo per duplicare un'operazione I/O effettuandola **contemporaneamente su due unità logiche distinte**; in altre parole, consente di avere in tempo reale due copie identiche degli stessi dati. È il livello Raid più sicuro in assoluto, ma ha un costo per megabyte piuttosto alto poiché la percentuale di spazio utilizzato è il doppio di quello che sarebbe necessario. Usato da solo, il disk mirroring non porta alcun vantaggio in termini di prestazioni, ma garantisce che per ogni file danneggiato, su un altro disco ci sia pronta una copia di riserva. Anche in questo caso sono necessari almeno due dischi.

Le performance migliorano in lettura dato che le richieste vengono soddisfatte simultaneamente da tutti i dischi dell'array, mentre in scrittura restano uguali, dovendo scrivere su entrambi i dischi.

Si conosce anche nelle varianti:

- **disk duplexing**: due dischi su due controller diversi
- **disk mirroring**: due dischi sullo stesso controller

## RAID 1: Schema



Ogni segmento stripe del disco 1 (Segmento 1) e duplicato sul disco 2 (Segmento 1').

Perdendo uno dei due dischi si avrà sempre una copia identica sull'altro.

**VANTAGGI:**

Aumento della velocità in lettura.  
Massima protezione dati.

**SVANTAGGI:**

Diminuiscono le prestazioni in scrittura.  
Dimezzamento della capacità dei dischi.

## RAID 1: Esempio

Mirroring software sotto Linux

2 drive uno per canale IDE

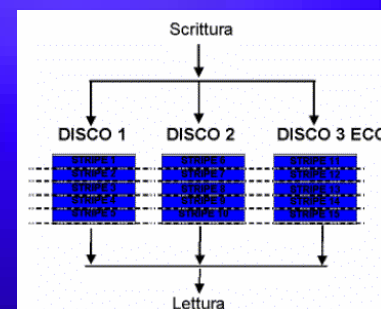
- hda
- hdc



Il sistema deve essere in grado di funzionare e ripartire anche con un solo disco a disposizione (Modalità Degraded)

## RAID 2

In questa vecchia configurazione di almeno tre dischi, due sono in striping e uno contiene informazioni ECC di fault tolerance. Dato che ormai ogni disco contiene un proprio codice di correzione degli errori questa implementazione è stata abbandonata. Inoltre i dati vengono disposti (cioè distribuiti con un procedimento di spanning) tra tutti i dischi e le operazioni di lettura e scrittura devono quindi continuamente accedere a dati disposti in posizioni diverse, muovendosi da un disco all'altro e rallentando il flusso di throughput.

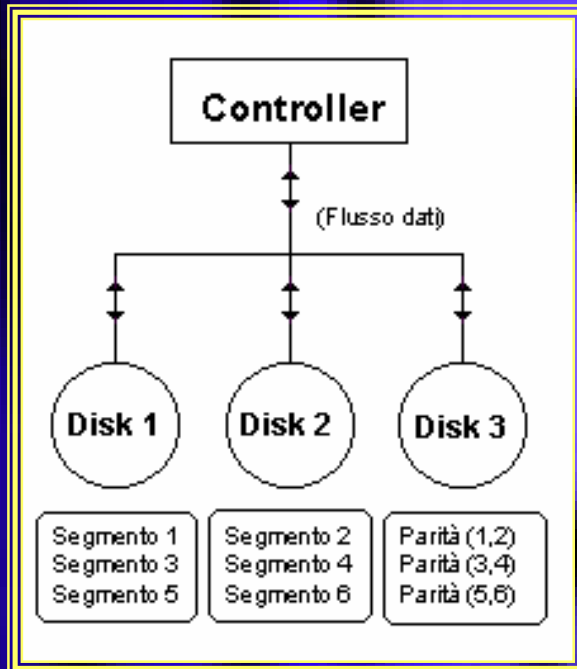




## RAID 3

- Si ottiene con un minimo di tre dischi, due usati come dati e uno riservato ai dati di parità. Il procedimento di frammentazione dei dati avviene in un modo simile allo striping, con la differenza che in un solo disco vengono memorizzate le stringhe di parità. Queste stringhe sono una sorta di "mappa" che permette di ricostruire i dati in caso di anomalia ad uno dei suddetti dischi.
- Questo sistema non è facilmente espandibile, inquanto il disco di parità non avrebbe capienza per le aggiuntive stinghe. Conseguenza è che se si vuole ampliare il sistema bisogna sostituire tutte e tre i dischi. La scrittura dei dati di parità sul disco dedicato causano un abbassamento delle prestazioni. Inoltre il malfunzionamento di un disco provoca un degrado delle prestazioni fino alla sostituzione di questo.
- Il sistema RAID 3 viene utilizzato nei sistemi monoutenti dove si accede a file di grosse dimensioni.

## RAID 3: Schema



Il Disco 3 contiene le informazioni di parità necessarie a ripristinare uno dei dischi andato perso.

**VANTAGGI:**

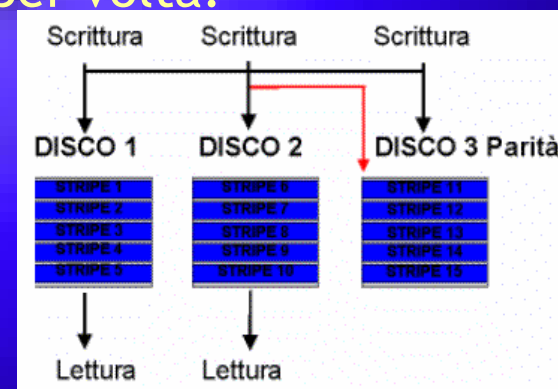
Alte prestazioni in accessi a grossi file.

**SVANTAGGI:**

Letture e scritture non simultanee.

## RAID 4

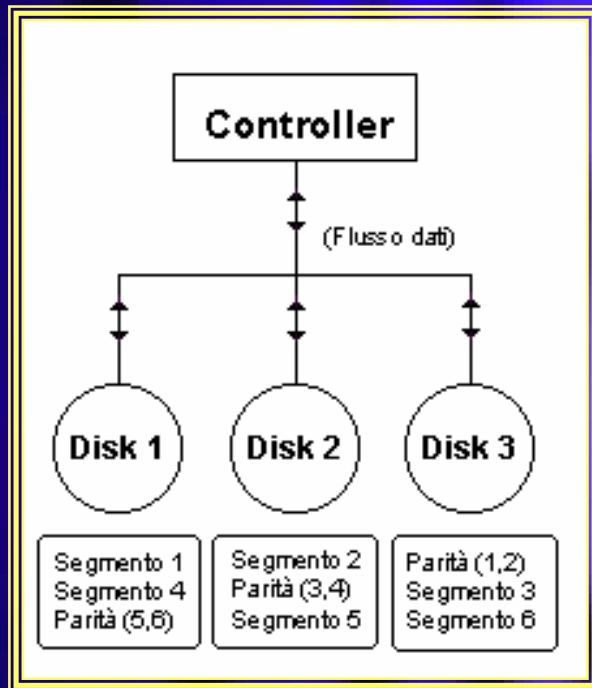
In questa implementazione di 3 drive minimo troviamo le stesse caratteristiche del RAID 3, solo che gli stripe sono più grandi di un record tipico, permettendo a quest'ultimo di risiedere interamente in un unico drive dell'array. Questa divisione dei record permette più operazioni di lettura simultaneamente, quindi un maggior flusso di dati. Tuttavia ogni scrittura deve aggiornare la parità sul drive specifico per cui non ne può avvenire più di una per volta.



## RAID 5

- Con una tecnica definita **parità distribuita**, il RAID 5 supera le limitazioni imposte al RAID 4 dal singolo drive destinato alla parità, dedicandone in ogni disco dei settori. Il totale dello spazio distribuito tra i dischi dedicato alla parità è uguale a quello di un singolo drive intero, permettendo che molteplici operazioni di scrittura avvengano simultaneamente, così come le letture.
- E nettamente superiore per prestazioni al RAID 3
- In questo modo si avrà la capacità pari a  $(n-1)*G$  Gbyte dove  $n$  è il numero di dischi presenti e  $G$  è la loro singola capacità. Questo dimostra che questo sistema è facilmente ampliabile rispetto al RAID 1. Il RAID 5 inoltre offre eccellenti prestazioni in lettura/scrittura. Questo sistema necessita di un minimo di 3 dischi ma consente l'ampliamento del RAID aggiungendo semplicemente dei dischi e riconfigurando il RAID.

## RAID 5: Schema

**VANTAGGI:**

Alte prestazioni.  
Alta protezione dati.  
Supporta più letture/scritture contemporanee.

**SVANTAGGI:**

Le prestazioni in scrittura sono inferiori ai sistemi RAID 0 o RAID 1.

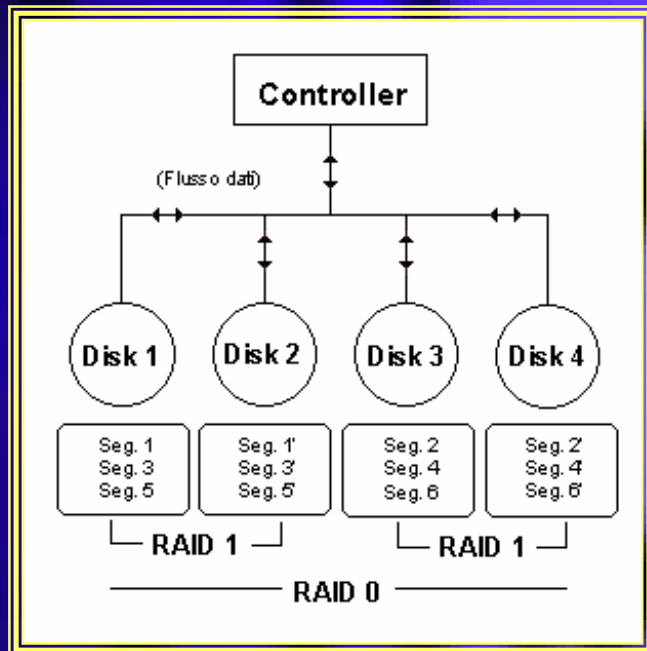
Ogni disco contiene parti di segmenti (dati) e parità necessarie a ricostruire uno dei tre dischi dell'array che può andare perso.

Come si può notare dall'esempio se si perde uno dei dischi la somma delle parità presenti sugli altri due mi consente di ricostruire il disco danneggiato.

## RAID 10 (0,1)

- Questo sistema implementa due diversi livelli di RAID:
  - RAID 1 e RAID 0.
- Un sistema di questo tipo necessita di un minimo di quattro dischi, due coppie in RAID 1 e l'insieme di questi raggruppati in RAID 0. Grazie all'accoppiata dei sistemi striping e mirror avremo un sistema sicuro e prestante. Questo sistema è in grado di rigenerare i dati anche in presenza di rotture in più dischi mantenendo buone prestazioni.
- Il RAID 10 ha migliori prestazioni del RAID 5, poiché non vengono gestite le informazioni di parità. Per contro è il sistema più costoso, in quanto la capacità complessiva dei dischi viene dimezzata e l'espansione del sistema va effettuata con coppie di dischi uguali.

## RAID 10 (0,1): Schema



Sopporta la rottura di due dischi dei differenti array in quanto le informazioni presenti sui due dischi di ogni singolo array vengono duplicate (mirrorate)

### VANTAGGI:

- Alte prestazioni.
- Alta protezione dati.
- Risulta affidabile anche in presenza di un blocco di più di un disco.

### SVANTAGGI:

- Dimezzamento della capacità dei dischi.
- Alto costo in rapporto alla capacità totale

## RAID 30 (0,3)

Questo sistema implementa due diversi livelli di RAID:

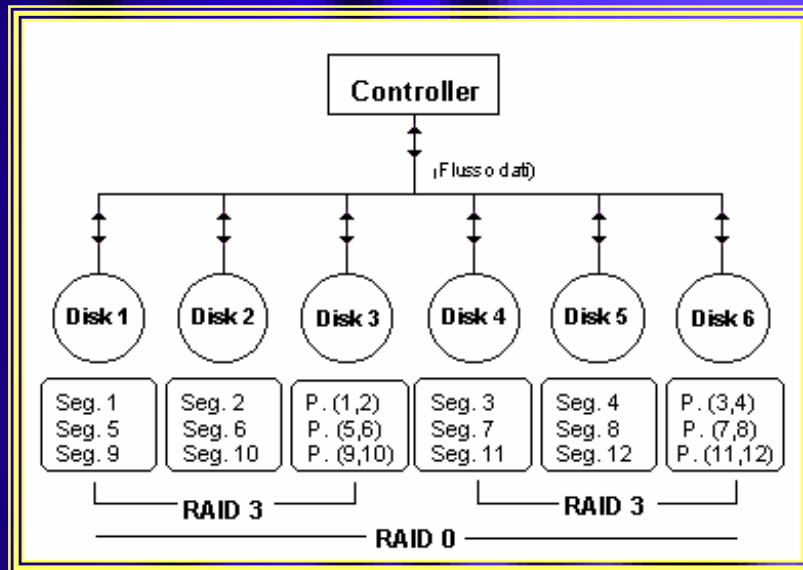
- RAID 3 e RAID 0.

Un sistema di questo tipo necessita di un minimo di sei dischi, due gruppi di tre dischi configurati in RAID 3 e l'insieme di questi raggruppati in RAID 0. Grazie all'accoppiata dei due sistemi avremo un sistema sicuro e prestante. Questo sistema è in grado di rigenerare i dati anche in presenza di rotture in più dischi mantenendo buone prestazioni. Il RAID 30 rispetto al RAID 10 risulta meno costoso (ma rimane meno prestante), in quanto a parità di capacità necessita di un disco in meno.

Come nel RAID 3 ha il suo collo di bottiglia nell'usare un disco dedicato per le informazioni di parità ogni sottogruppo del sistema.



## RAID 30 (0,3): Schema



Sopporta la rottura di due dischi dei differenti array in quanto le informazioni di parità presenti sui singoli dischi di parità (uno per array) consentono la ricostruzione dell'array globale.

### VANTAGGI:

- Alte prestazioni.
- Alta protezione dati.
- Risulta affidabile anche in presenza di un blocco di più di un disco.

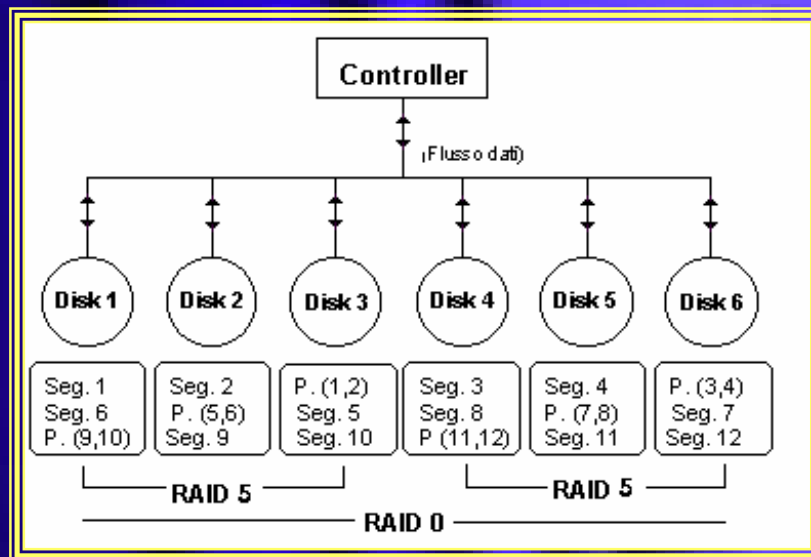
### SVANTAGGI:

- Difficile upgradabilità a causa dei dischi di parità.

## RAID 50 (0,5)

- Questo sistema unisce due diversi livelli di RAID:
  - RAID 5 e RAID 0.
- Un sistema di questo tipo necessita di un minimo di sei dischi, due gruppi di tre dischi configurati in RAID 5 e l'insieme di questi raggruppati in RAID 0. Grazie all'accoppiata dei due sistemi avremo un sistema sicuro e prestante. Questo sistema è in grado di rigenerare i dati anche in presenza di rotture in più dischi relative ai due array in RAID 5

## RAID 50 (0,5) : Schema



Sopporta la rottura di due dischi dei differenti array in quanto le informazioni di parità presenti sugli altri dischi consentono la ricostruzione dell'array.

Consente l'aggiunta di nuove unità di spare

### VANTAGGI:

- Alte prestazioni.
- Alta protezione dati.
- Risulta affidabile anche in presenza di un blocco di più di un disco.
- Facile upgradabilità.

### SVANTAGGI:

- Unico svantaggio la minor ridondanza rispetto al RAID 10.

## Evoluzione dello Storage

- Storage si riferisce all'”immagazzinamento” dei dati in dispositivo fisico.
- La nuova evoluzione dello storage passa per:
  - SAN (Storage Area Network)
  - LVM (Logical Volume Manager)

## Dallo SCSI alle SAN

- SCSI ha il limite della lunghezza del cavo (pochi mt) e dei dispositivi collegabili (qualche decina)
- Queste limitazioni producono l'effetto di legare rigidamente i dischi ai singoli Server e di rendere estremamente complicato l'impianto di cablaggio
- Difficile espandere la struttura

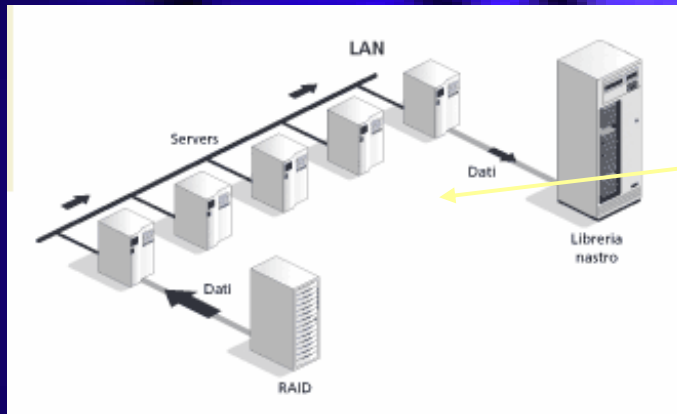
## Storage Area Network

- Utilizzano la tecnologia Fibre Channel rendono il cablaggio molto più semplice permettendo di superare i limiti delle distanze e della connettività. Inoltre, con i “fibre channel switch” si può ovviare ai problemi di Fault Tolerance e introdurre una gestione intelligente per la ripartizione dei Device.
- Maggiori prestazioni (velocità fino 200 Mb/sec) dei driver e della struttura di storage in generale.
- Supporto di un numero illimitato di periferiche, l'espandibilità è illimitata e garantita.

## SAN - Tecnologia Fibre Channel

- I sistemi Raid basati su piattaforma Fibre Channel possono essere posizionati ad una distanza fino a 30 metri dal server se connessi con cablaggio in copper (rame). Utilizzando cablaggio in fibra ottica, si può raggiungere una distanza di circa 10 KM.
- Fibre Channel è l'unica tecnologia che permette l'utilizzo di HUB o SWITCH (FC), ciò permette di "migrare" verso una architettura *Storage Area Network*. Con questa architettura si possono configurare situazioni tipo: Clustering, Disaster Recovery, centralizzazione dello storage su più server, condivisioni dei dati su file system omogenei o eterogenei, connettività di più loop, situazioni di remote Mirroring tra diverse unità raid.

# SAN - Esempi



Con l'architettura RAID-SCSI i sistemi devono essere collegati lungo un BUS.

Con l'architettura SAN l'interconnessione avviene con cavi in fibra ottica ad alta velocità.





## Clustering

- Un cluster (grappolo) è un insieme di computer in grado di eseguire insieme una certa serie di compiti.
- Un gruppo di sistemi che lavorano come se fossero uno solo.

## Perché usare un cluster

- Consente di creare architetture scalabili
- Consente di ottimizzare le strutture disponibili
- Alte prestazioni con costi contenuti (buon rapporto qualità servizio/prezzo)

## Tipologie di Cluster

- Può essere utilizzato per:
  - **calcolo distribuito:** usati per rispondere a grandi capacità di calcolo (crittografia)
  - **calcolo distribuito:** usati per distribuire grandi carichi di rete
  - **alta disponibilità (High Availability) :** usato dove è necessario un alto livello di disponibilità e di Fault Tolerance (Servizio Ininterrotto).

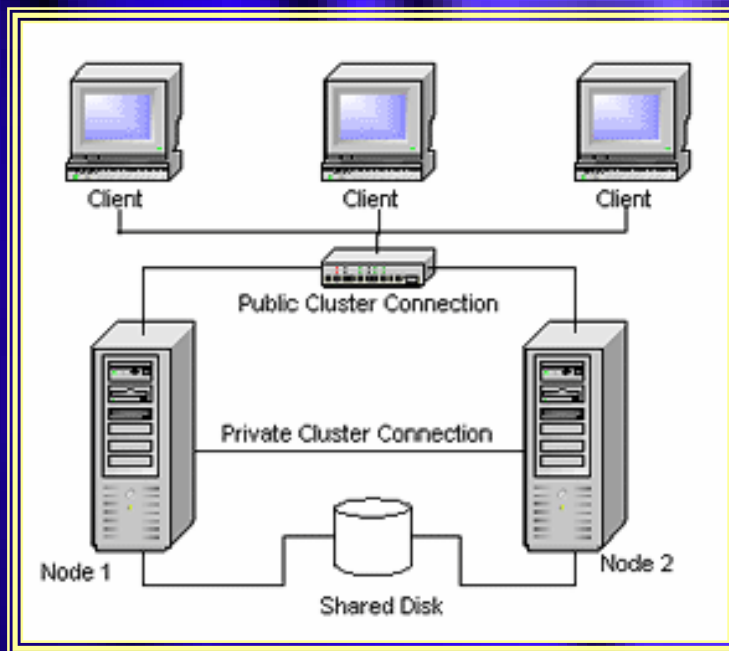
## Come funziona un cluster

- Computer farm o cluster di workstation
- Suddivide il lavoro per tutti i membri del cluster
- Sfrutta le potenzialità di calcolo di ogni singolo host, sommandole
- Necessità di un software apposito per gestire il cluster ed il dialogo tra i diversi elementi del cluster
- Utilizzo di programmi dedicati per il calcolo parallelo (PMI e PVI)
- Trasferisce il suo compito ad un altro membro del cluster in caso di guasto.

## Linux Clustering

- Coda (Filesystem Distribuito)
- Kernel LVM
- Beowulf
- Linux HA
- Linux OpenMosix

## Schema di un Cluster



### Elementi essenziali:

- Area di Storage condivisa
- Connessione fast in rete
- Connessione diretta (privata) tra i membri del cluster
- Ridondanza totale del sistema

## Risorse

- <http://www.networkingitalia.it/windows/faqraid.asp>
- <http://www.networkingitalia.it/server/raid.asp>
- <http://www.ecobyte.it/soluzioni/ita>