

RAID

Il termine RAID significa Redundant Array of Independent Disks: sono tecniche di organizzazione dei dischi che permettono di migliorare performance e affidabilità.

Al sistema operativo un RAID appare come un singolo disco logico; il RAID controller si occupa della traduzione delle richieste da logiche a fisiche.

Il controller RAID può essere implementato tramite hardware o software. Sono inoltre usati dischi **hot-spare** per sostituire dischi fallati automaticamente. In un sistema RAID, la performance è migliorata tramite parallelismo e l'affidabilità tramite ridondanza.

Parallelismo

Per ottenere parallelismo, viene utilizzata una tecnica detta "**data striping**": i dischi logici sono divisi in **strip** e strip consecutive su dischi diversi sono dette **stripe**. Il data striping può avvenire a livello di bit, byte o blocco di byte.

Per trovare disco e offset uno strip in un array di n dischi senza ridondanza si possono usare le seguenti equazioni (IL = indirizzo logico):

- Disco = $IL \bmod n$
- Offset = IL / n

Richieste di strip differenti su dischi differenti possono avvenire in parallelo, ma richieste che coinvolgono strip sullo stesso disco devono avvenire sequenzialmente. Lo striping è vantaggioso nel caso vi siano richieste I/O di considerevoli dimensioni.

Ridondanza

Un array di dischi è meno affidabile di un solo disco (il MTTF dell'array è equivalente al minore tra i dischi): si sfrutta la ridondanza.

Con la tecnica del mirroring ogni disco è duplicato: se un disco fallisce, si usa il suo clone. I dati sono perduti se il disco clone fallisce prima di essere rimpiazzato. Ogni scrittura su un disco è effettuata anche sul suo clone: avvenendo in parallelo, il transfer rate è lo stesso.

RAID0

RAID0 non offre alcun tipo di ridondanza (quindi i disk failure non sono tollerati), i blocchi di dati sono distribuiti tra i dischi (striping) seguendo una politica round-robin.

Le performance sono elevate, data la possibilità di eseguire richieste in parallelo, così come la capacità, ma non è consigliato per sistemi che richiedono pochi dati alla volta.

RAID1

RAID1 usa mirroring per fornire ridondanza: sono tollerati tanti disk fault quanti sono i dischi presenti, ma serve il doppio dei dischi per implementarlo.

RAID2

RAID2 usa striping a livello di bit e ECC (error correction codes): i dati sono divisi in gruppi di bit e per ogni gruppo è calcolato un ECC. I bit di ECC e di dati sono memorizzati su dischi diversi: ci sono quindi n dischi di dati e m dischi per ECC. Le richieste di lettura coinvolgono solo i dischi di dati, mentre quelle di scrittura entrambi, dato che gli ECC vanno ogni volta ricalcolati.

Nonostante l'alta affidabilità e la buona performance in lettura, RAID2 non è molto utilizzato, data la sua complessità e costo.

RAID3

RAID3 usa striping a livello di bit più un bit di parità: sono necessari n dischi di dati e 1 per parità.

Il bit di parità è un bit posto alla fine di una stringa binaria che indica se il numero di bit a 1 della stringa è pari o dispari. Due varianti:

- even parity bit: bit di parità settato a 1 se il numero di bit a 1 è dispari
- odd parity bit: bit di parità settato a 1 se il numero di bit a 1 è pari

I bit di parità in RAID3 permettono di correggere errori: sappiamo quale disco è fallito, quindi basta sostituire il bit appartenente al disco in questione per ripristinare la parità. Come RAID2, anche RAID3 è poco utilizzato.

RAID4

RAID4 usa striping a livello di blocchi e un blocco di parità: sono necessari n dischi di dati e 1 per parità.

Il blocco di parità è calcolato mettendo in XOR (0 se i bit sono uguali) due blocchi della stessa stripe. Ci sono due approcci per la computazione del blocco di parità:

- parità additiva: tutte le strip sono lette e messe in XOR (utile per operazioni che coinvolgono molti dischi)
- parità sottrattiva: solo la strip da sostituire, la nuova strip e il blocco di parità sono messi in XOR (utile per operazioni che coinvolgono pochi dischi)

Il cutoff per l'uso di parità sottrattiva è $\frac{n}{2} - 1$.

RAID4 offre buona performance su grandi richieste, ma non può gestire scritture di pochi dati su stripe diverse, dato che devono scrivere sullo stesso disco di parità.

RAID5

RAID5 usa striping a livello di blocchi e blocco di parità distribuito: interlacciando blocchi di dati e parità si risolve il problema del collo di bottiglia che si presenta in RAID4. Tuttavia, il problema delle scritture di pochi dati persiste.

RAID6

RAID6 usa striping a livello di blocchi e blocchi di parità distribuiti: ogni stripe ha due blocchi di parità contenuti in dischi diversi. Molto affidabile, ma la performance in scrittura è peggiore di quella di RAID5.

RAID ibridi

RAID 0+1 (mirror of stripes) consiste nello striping e mirroring di n dischi.

RAID 1+0 (stripe of mirrors) consiste nel mirroring e poi striping di n dischi.