

# Sistemi Operativi 1

AA 2018/2019

Sistemi RAID

# Sommario

- Motivazioni
- Concetti di base
- Livelli RAID



# Motivazione

- L'evoluzione tecnologica ha permesso di avere dischi sempre più piccoli e meno costosi
  - E' facile equipaggiare un sistema con molti
- Dischi
- La presenza di più dischi consente:
  - maggiori prestazioni: letture/scritture in parallelo
  - maggior affidabilità: tramite ridondanza

# RAID – definizione e obiettivi

- A Case for Redundant Arrays of Inexpensive Disks (RAID) by D. A. Patterson, G. Gibson, and R. H Katz, articolo pubblicato nel 1988
  - (<http://wla.berkeley.edu/~cs61c/resources/Patterson88.pdf>)
- RAID: Redundant Array of Independent (Inexpensive) Disks.
- Obiettivi:
  - Migliorare l'affidabilità
  - Incrementare le prestazioni
- Opposto all'approccio SLED; Single Large Expensive Disk



# Struttura dei dispositivi RAID

- Struttura SW
  - Più dischi indipendenti collegati al bus
  - Funzionalità RAID implementata dal sistema operativo
- Struttura HW
  - Controllore intelligente gestisce diversi dischi collegati alla macchina
- Batteria RAID
  - Unità a se stante composta da controllore, cache, e dischi autonomi collegati a una macchina

# Concetti di base

- Le strutture RAID si basano su:
  - Copiatura speculare dei dati (mirroring)
  - Sezionamento dei dati (data striping)
- per implementare una sorta di parallelismo che garantisce aumento di:
  - affidabilità
  - prestazioni

# Affidabilità

- Un guasto a un disco comporta la perdita di dati
- Per migliorare l'affidabilità si deve ricorrere alla *ridondanza*:
  - Memorizzare informazioni non strettamente necessarie, ma utili per ricostruire le informazioni perse in caso di guasto

# Copiatura speculare

- Il modo più semplice per implementare la ridondanza è il *mirroring* (o *shadowing*)
  - Un disco logico corrisponde a due dischi fisici
  - Ogni scrittura avviene su entrambi i dischi
  - I dati si perdono solo se si guastano entrambi i dischi
- Tempo medio di perdita dei dati in caso di mirroring dipende da:
  - Tempo medio di guasto di ogni singolo disco
  - Tempo medio di riparazione





# Esempio

- Supponendo che:
  - Possibili guasti sui dischi siano indipendenti
  - Tempo medio di guasto sul singolo disco = 100.000 ore
  - Tempo di riparazione = 10 ore
- Si ha il seguente tempo medio di perdita dei dati su un sistema con mirroring:
  - $100.000^2 / (2 * 10) = 57.000$  anni!

# Guasti indipendenti

- Non sempre i guasti su dischi sono indipendenti
- Ad esempio in caso di:
  - disastro
  - calo di tensione
  - difetti di fabbricazione in una partita di dischi
  - ...
- i guasti potrebbero verificarsi contemporaneamente

# Prestazioni

- Con mirroring:
  - la frequenza di gestione delle letture raddoppia perché si può leggere da uno qualunque dei due dischi
  - il tempo di trasferimento rimane inalterato

# Sezionamento dei dati

- Usando più dischi è possibile migliorare la *capacità di trasferimento* distribuendo i dati in sezioni su più dischi
- Tale tecnica si chiama *data striping*

# Sezionamento dei dati

- Sezionamento a livello di bit
  - Distribuzione dei bit di ciascun byte su più dischi
  - Es. con 8 dischi, il bit  $i$ -esimo di un byte è memorizzato nel disco  $i$ -esimo
- Sezionamento a livello di blocco
  - Distribuzione dei blocchi di un file su più dischi
  - Es. con  $n$  dischi, il blocco  $i$ -esimo di un file è memorizzato nel disco  $(i \bmod n)$ -esimo

# Effetti del parallelismo

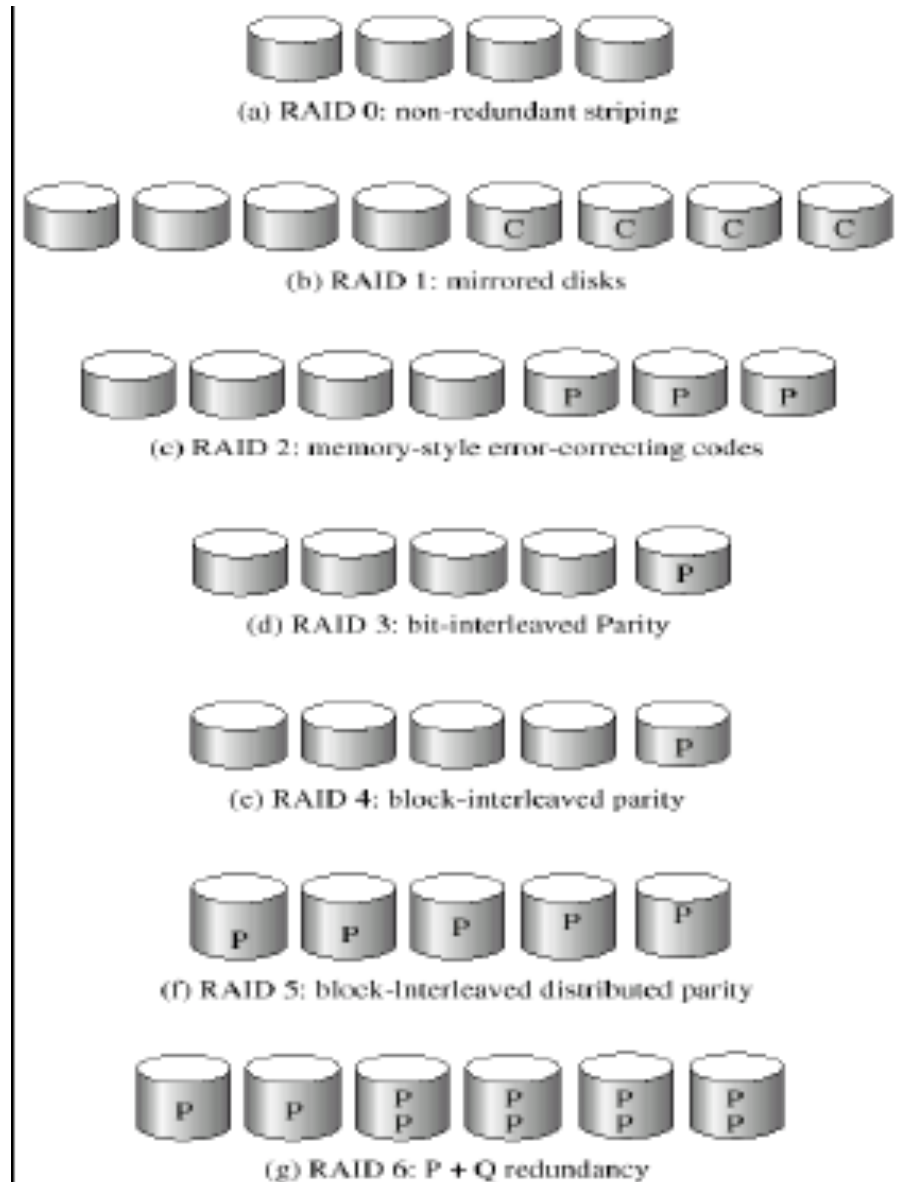
- Aumento, tramite *bilanciamento del carico*, della **produttività** per *accessi multipli* a piccole porzioni di dati
- *Riduzione del tempo di risposta* relativo agli accessi a grandi quantità di dati

# Livelli di RAID

- Mirroring:
  - alta affidabilità ma costoso
- Data striping:
  - alta capacità di trasferimento dati, ma non migliora l'affidabilità
- Spesso si usano tecniche basate sui bit di parità
- Utilizzo combinato di queste tecniche è stato schematizzato in 6 livelli RAID

# Livelli di RAID

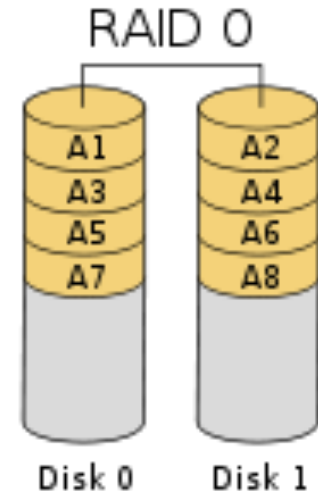
- C = mirroring
- P = bit di correzione degli errori
- Es.: 4 dischi dati, gli altri servono per ridondanza





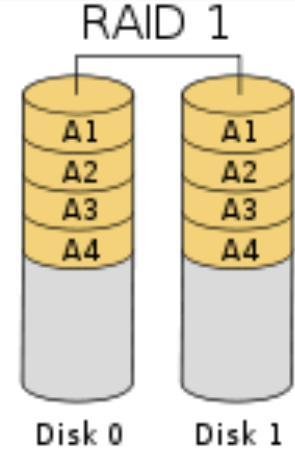
# Livello RAID 0

- Caratteristiche:
  - Sezionamento a livello di blocco
  - Senza ridondanza
- Vantaggi:
  - Economico
  - Alte prestazioni grazie al parallelismo delle operazioni di lettura/scrittura
- Svantaggi:
  - Non ha ridondanza
  - L'affidabilità cala all'aumentare del numero di dischi impiegati (es.: l'affidabilità di un sistema con 2 dischi è la metà rispetto a quella con un disco solo)



# Livello RAID 1

- Caratteristiche:
  - Mirroring senza sezionamento di blocco
- Vantaggi:
  - Affidabilità aumenta linearmente con il numero di copie (dischi)
  - Aumento delle prestazioni in lettura
    - Se un disco è occupato posso leggere dall'altro
- Svantaggi:
  - Alto costo
  - Bassa scalabilità



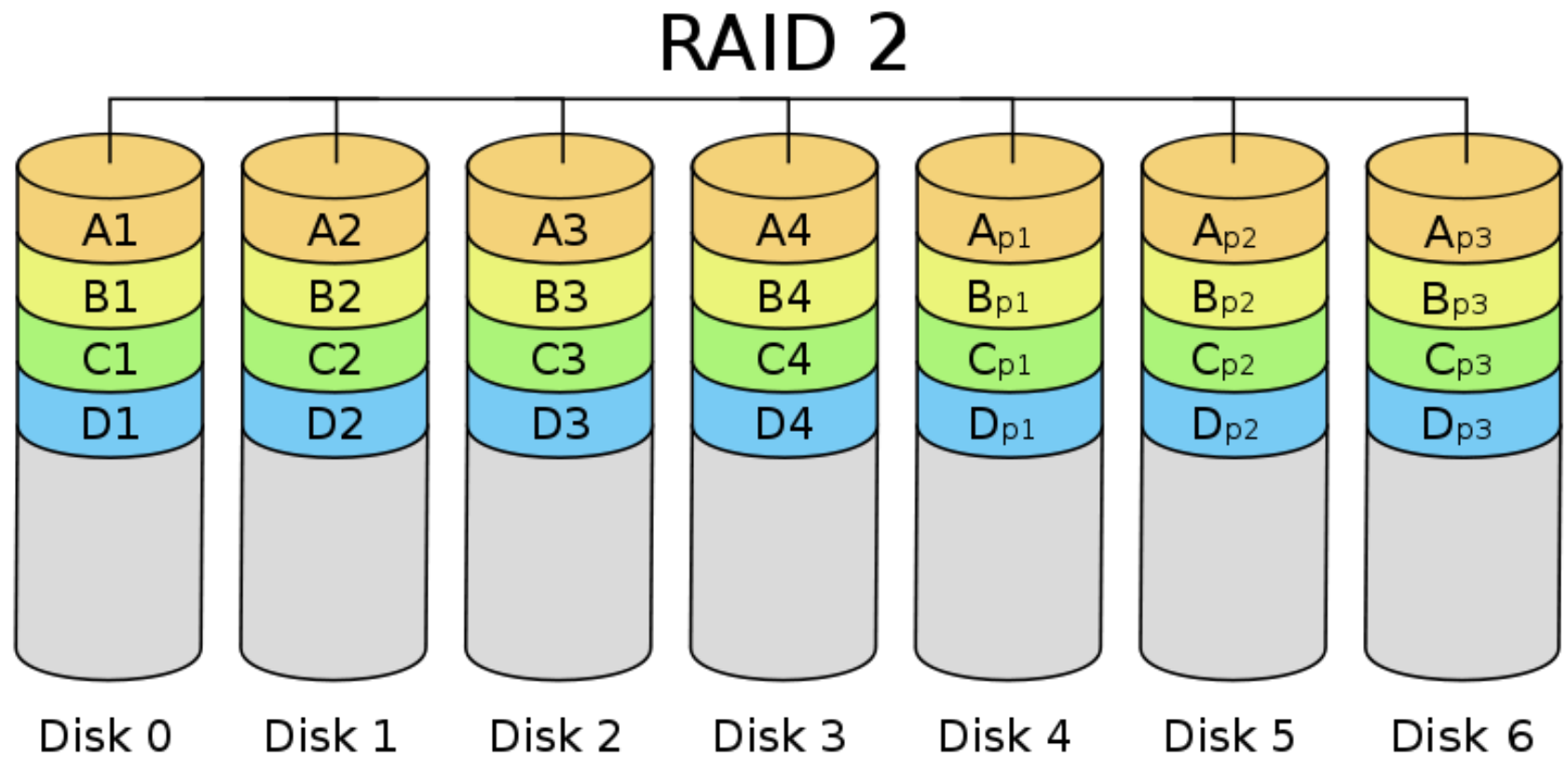
# Codici per la correzione degli errori

- Bit di parità
  - Ad ogni byte è associato un bit che indica se gli 1 presenti nel byte sono in numero pari (parità 0) o dispari (parità 1)
  - identificano tutti gli errori su un singolo bit
- Usando più bit supplementari si riescono a individuare/correggere un maggior numero di bit

# Livello RAID 2

- Caratteristiche:
  - Sezionamento a livello di bit
  - Utilizza i codici per la correzione degli errori (ECC)
    - Primo bit di ogni byte memorizzato nel disco 1, il secondo nel disco 2, ecc.
    - I bit di correzione sono memorizzati singolarmente in dischi separati differenti rispetto a quelli usati per i dati
  - Se un disco si guasta, i bit rimanenti del byte dati e i bit di correzione associati vengono usati per ricostruire il dato danneggiato
  - Il RAID 2 richiede solo 3 dischi in più per 4 dischi dati, contro i 4 richiesti dal RAID 1

# Livello RAID 2



# Codice Hamming(7,4)

- E' un codice di Hamming che codifica 4 bit di dati in 7 bit, aggiungendo 3 bit di parità.
- Può correggere ogni errore di singolo bit, oppure rivelare tutti gli errori di singolo bit e gli errori su due bit, ma senza poterli correggere.

$$\mathbf{x} = \mathbf{Gp} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

- Quindi trasmetteremo 0110011 prodotto dal messaggio originale 1011



# Codice Hamming(7,4)

- Controllo di parità'

$$\mathbf{z} = \mathbf{H}\mathbf{r} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{z} = \mathbf{H}\mathbf{r} = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

# Codice Hamming(7,4)

- Ricostruzione dati

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\mathbf{p}_r = \mathbf{R}\mathbf{r} \quad \mathbf{p}_r = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$



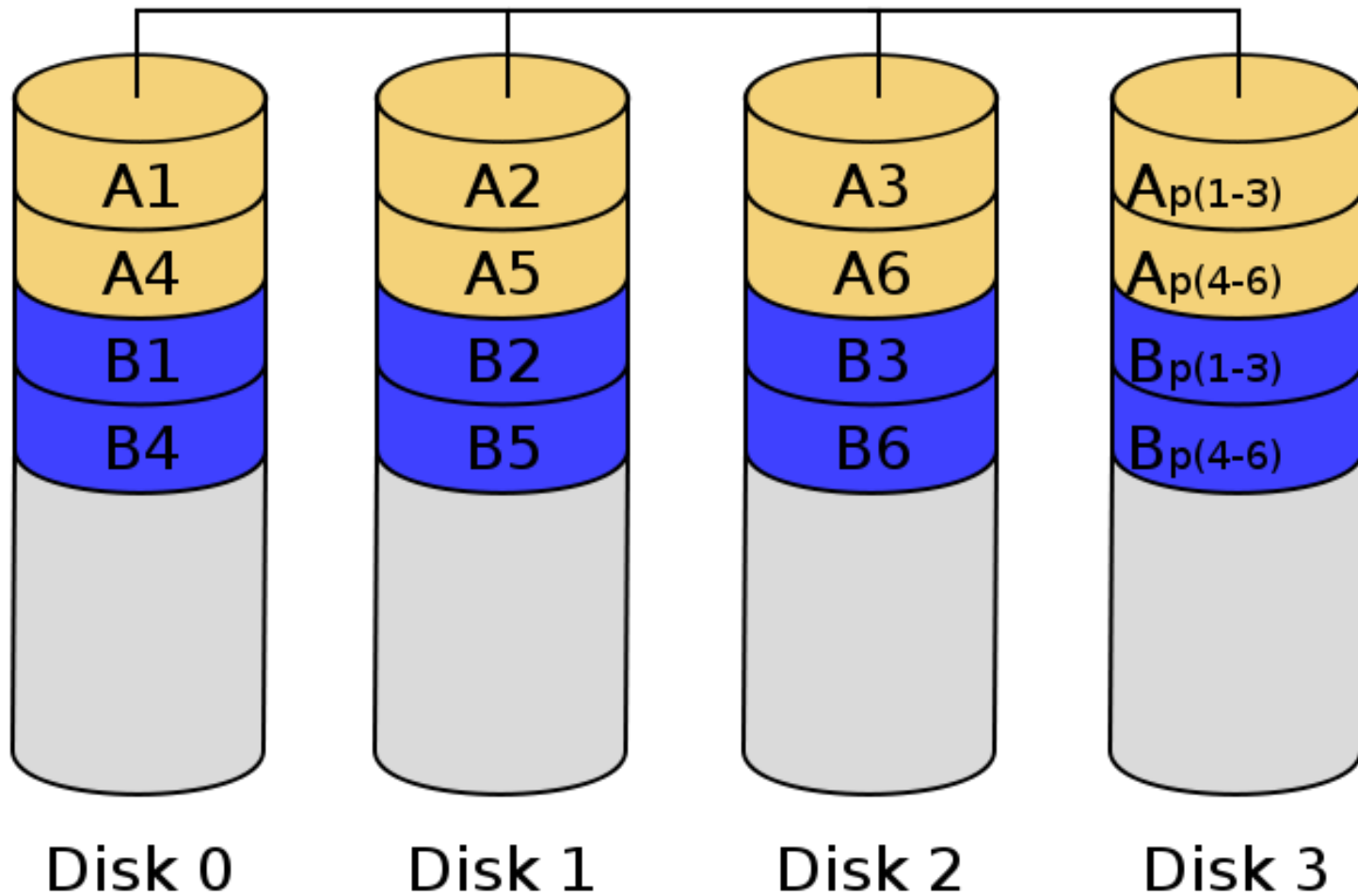
# Livello RAID 2

- Vantaggi
  - E' un RAID 0 con maggior affidabilità
- Svantaggi
  - Costoso

# Livello RAID 3

- Caratteristiche:
  - Sezionamento a livello di byte con un disco dedicato al bit di parità
    - Noto come organizzazione con bit di parità
  - I controllori dei dischi sono in grado di rilevare se un settore è stato letto correttamente
    - Se un settore è danneggiato, per ogni bit del settore è possibile determinare se deve valere 0 oppure 1 calcolando la parità dei bit corrispondenti dai settori degli altri dischi.
    - Se la parità dei rimanenti bit è uguale a quella memorizzata, allora il bit mancante è 0, altrimenti è 1

# Livello RAID 3



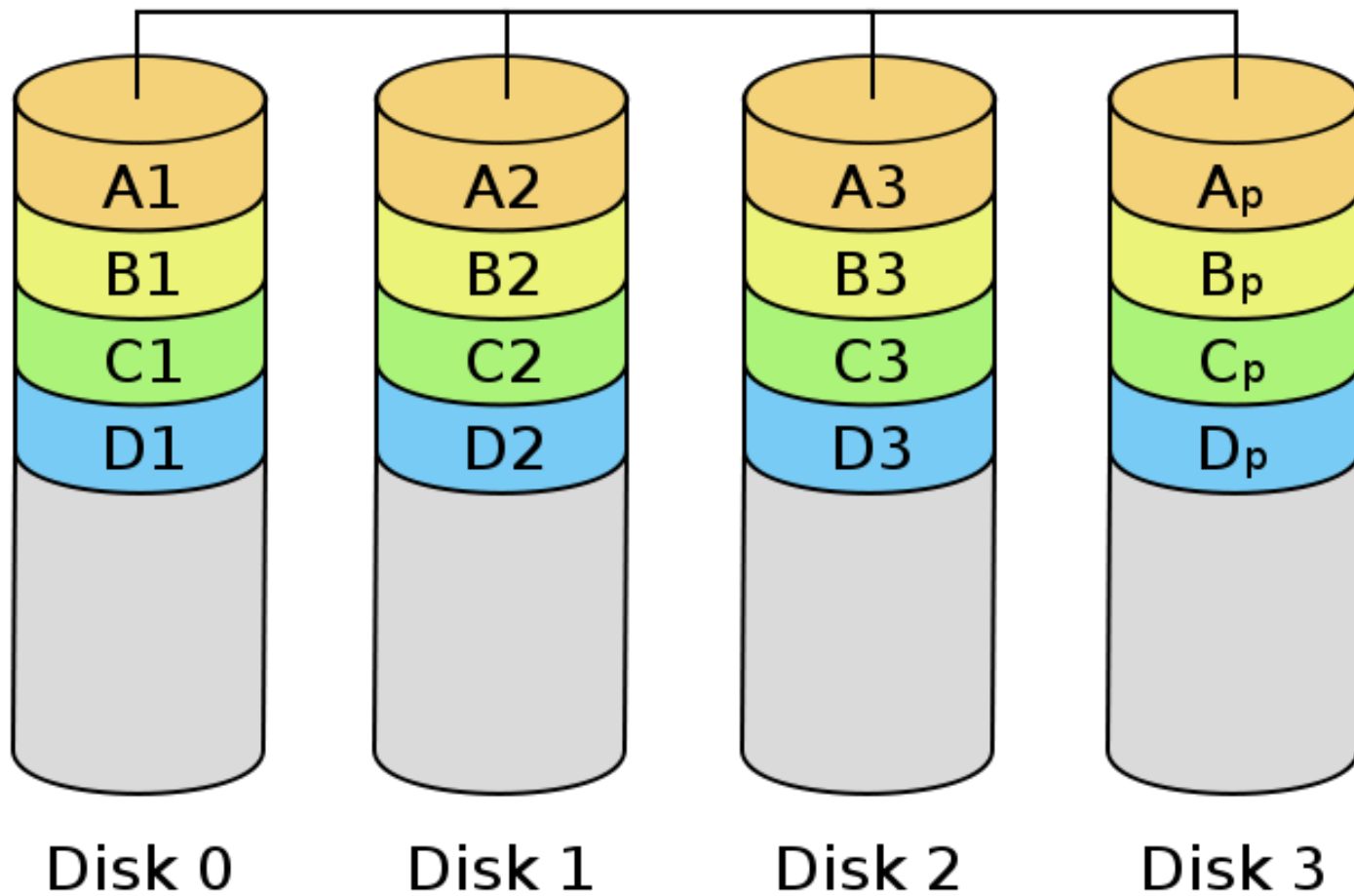
# Livello RAID 3

- Vantaggi:
  - Stessa efficienza del RAID 2, ma usa un solo disco per i bit di parità
  - Velocità di trasferimento pari a  $n$  volte quella del RAID 1, grazie al data striping
- Svantaggi (rispetto al RAID 1):
  - Meno operazioni di I/O al secondo perché ogni disco è coinvolto da tutte le richieste
  - Tempo più lungo per scritture perché è necessario calcolare il bit di parità
    - Soluzione: controllore RAID è capace di gestire il calcolo della parità, sollevando la CPU dall'incarico

# Livello RAID 4

- Caratteristiche:
  - Sezionamento a livello di blocco con disco dedicato alla parità
    - Noto come organizzazione con blocchi di parità intercalati
  - Come RAID 0, più un blocco di parità in un disco separato

# Livello RAID 4



# Livello RAID 4

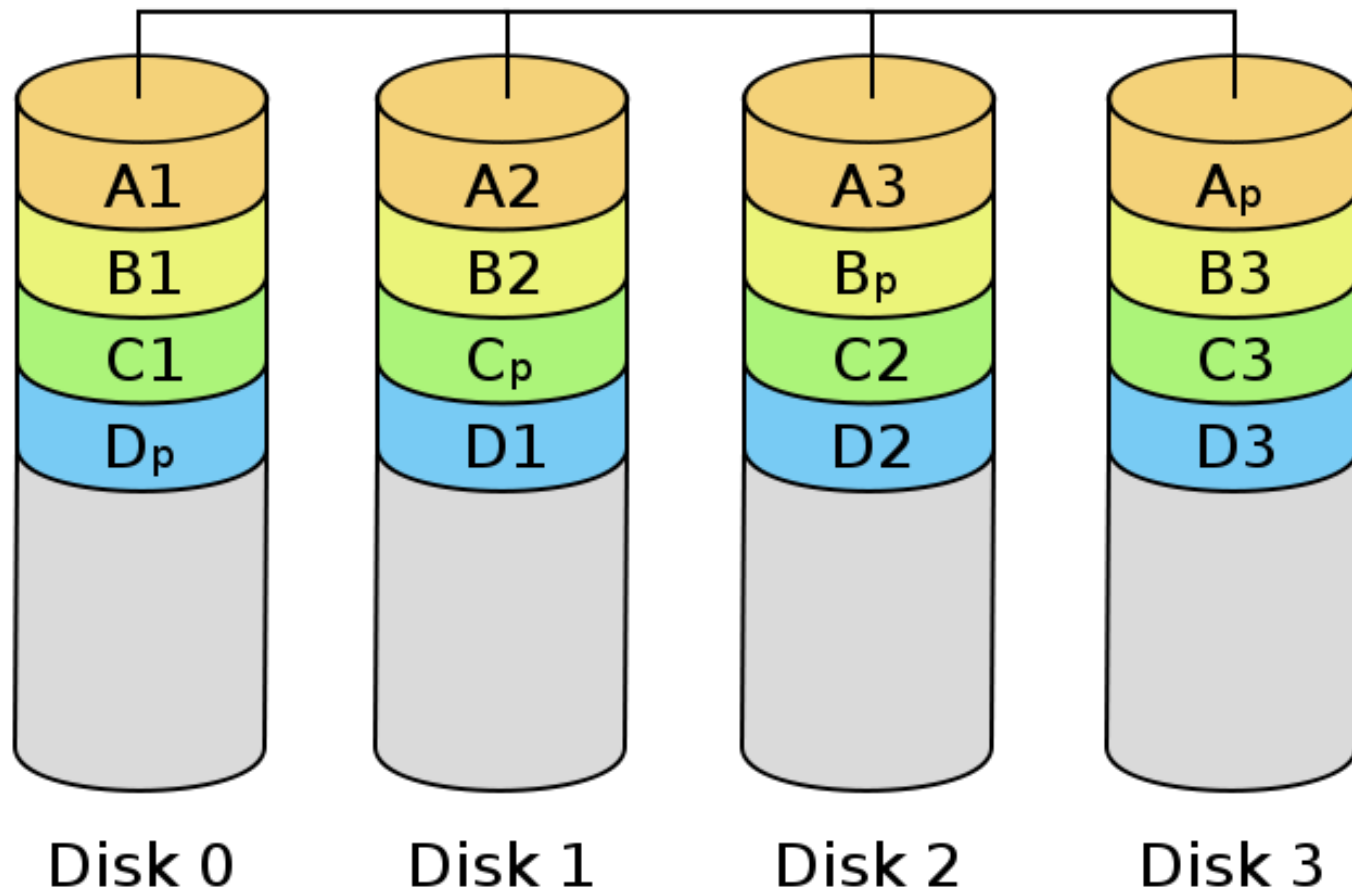
- Vantaggi:
  - Tolleranza ai guasti
  - Letture più veloci grazie al parallelismo
- Svantaggi:
  - Il disco usato per la parità può essere collo di bottiglia
  - Scritture lente a causa del calcolo della parità

# Livello RAID 5

- Caratteristiche:
  - Sezionamento a livello di blocco con bit di parità distribuiti tra tutti i dischi del RAID
    - Noto come organizzazione con blocchi intercalati a parità distribuita
  - Un blocco di parità non può contenere informazioni di parità per blocchi che risiedono nello stesso disco
    - Altrimenti un guasto al disco farebbe perdere i dati
  - E' un implementazione molto popolare



# Livello RAID 5



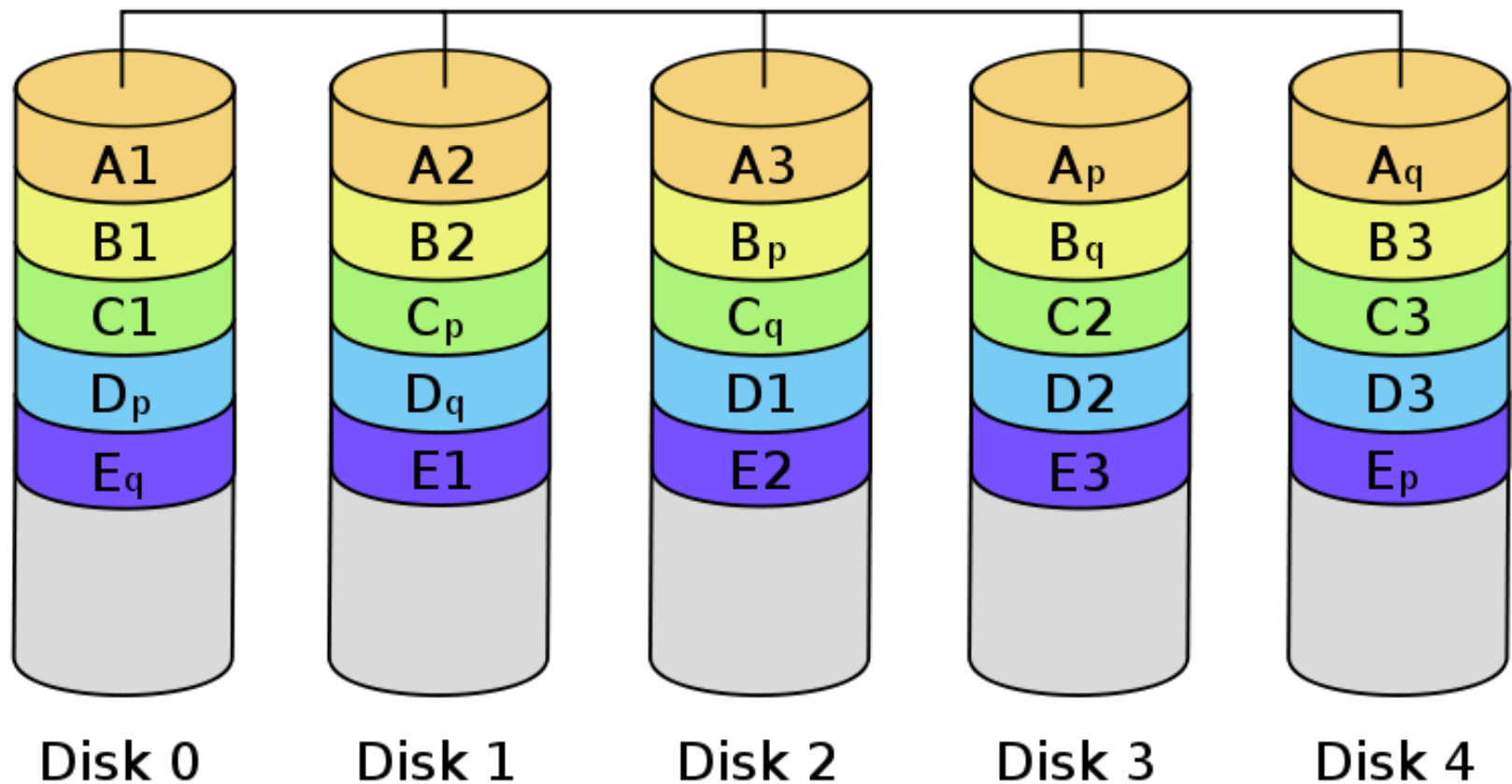
# Livello RAID 5

- Vantaggi
  - Come per il RAID 4 ma senza avere il collo di bottiglia del disco di parità
- Svantaggi
  - Scritture lente come per il RAID 4

# Livello RAID 6

- Caratteristiche:
  - Simile al RAID 5, ma con maggiori informazioni di ridondanza per gestire guasti contemporanei su più dischi.
  - Al posto della parità usa altri codici per la correzione dell'errore (Reed-Solomon)

# Livello RAID 6

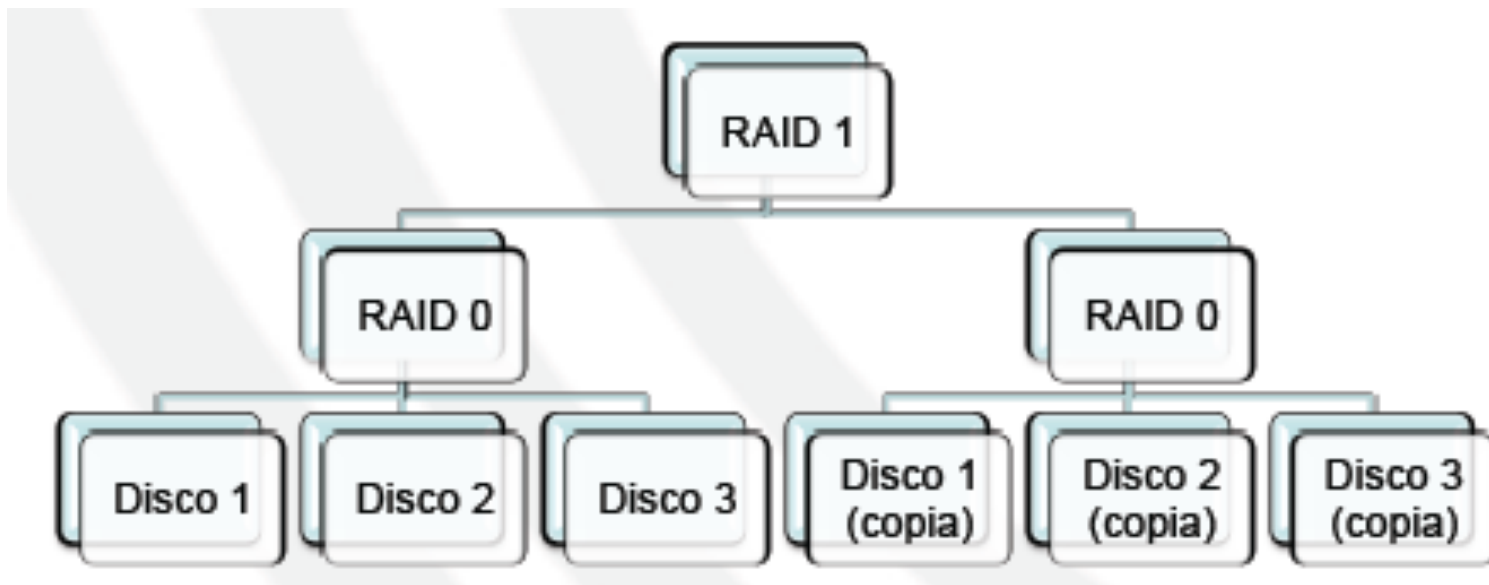


# Livello RAID 6

- Vantaggi:
  - Altissima ridondanza
- Svantaggi:
  - Molto costosi
  - Scritture molto lente per la gestione dei codici per la correzione degli errori

# RAID di livello 0+1

- Caratteristiche
  - Combinazione di 0 e 1 per fornire affidabilità e alte prestazioni

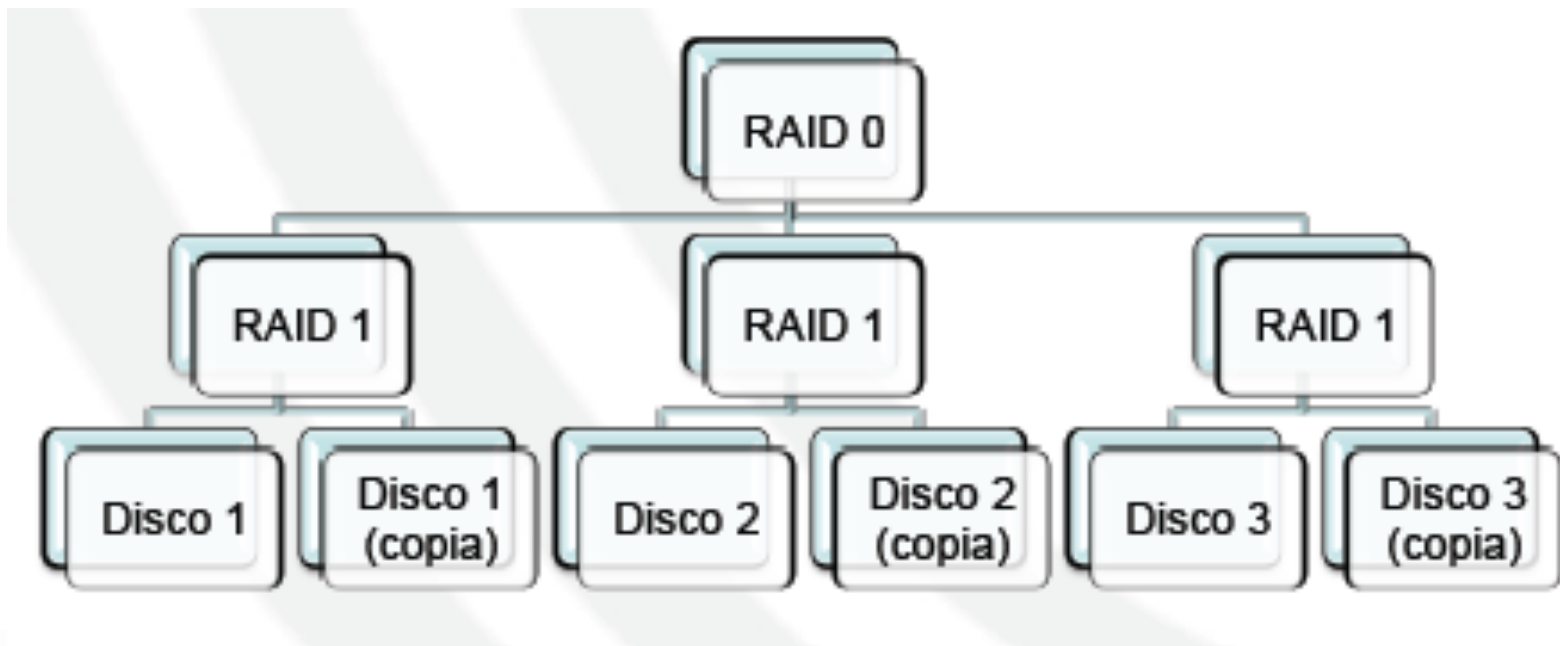


# RAID di livello 0+1

- Vantaggi:
  - Prestazioni migliori rispetto al RAID 5
  - Alta affidabilità
- Svantaggi:
  - Richiede il raddoppio del numero di dischi necessari per memorizzare i dati
  - Più costoso del RAID 5
  - Non supporta la rottura simultanea di 2 dischi se non appartengono allo stesso stripe

# RAID di livello 1+0

- Caratteristiche
  - Combinazione di 1 e 0 per fornire affidabilità e alte prestazioni





# RAID di livello 1+0

- Vantaggi:
  - Più robusto del RAID 0+1
    - Ogni disco di ogni stripe può guastarsi senza far perdere dati al sistema
- Svantaggi:
  - Costoso

# Riassumendo

- **RAID 0**: aumento delle prestazioni in lettura e scrittura
- **RAID 1**: tollera il guasto di un disco, con ricostruzione automatica
- **RAID 2**: prevede codici a correzione d'errore
- **RAID 3**: oltre ad avere codici a correzione d'errore, comporta minor sovraccarico
- **RAID 4**: consente l'esecuzione di letture contemporanee
- **RAID 5**: letture e scritture contemporanee, parità distribuita tra i dischi
- **RAID 6**: doppia parità, tollera il guasto di 2 dischi
- **RAID 0+1**: sfrutta velocità del livello 0, implementando la sicurezza come nel livello 1
- **RAID 1+0**: Tollera il guasto di piu' dischi se non sono sotto lo stesso raid

