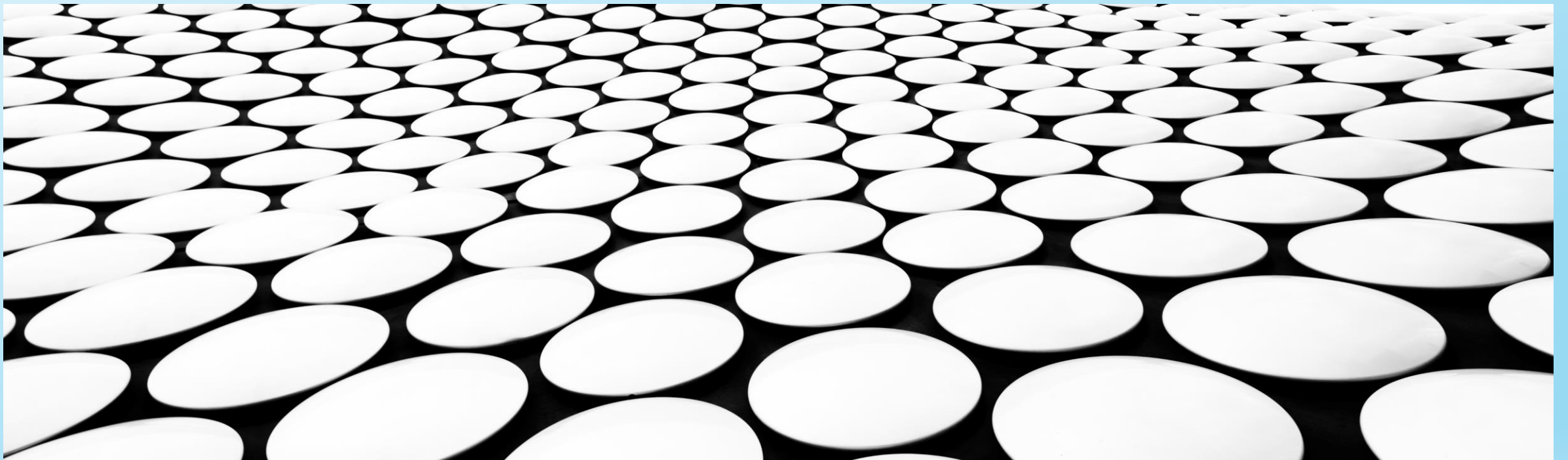

ARHITECTURA SISTEMELOR DE CALCUL

UB, FMI, CTI, ANUL III, 2022-2023



Arhitecturi RAID

Ce înseamnă RAID?

- RAID (**R**edundant **A**rray of Independent **D**isks sau **R**edundant **A**rray of Inexpensive **D**isks) reprezintă un sistem de HDD(SSD)-uri care are ca scop:
 - creșterea vitezelor scriere / citire,
 - protecția datelor
 - combinația lor.
- Există mai multe **tipuri** de RAID. Cele mai folosite și cunoscute sunt RAID 0, RAID 1, RAID 5 și RAID 10.
- De obicei se utilizează mai multe discuri identice pentru o implementare RAID

RAID 0

Tipul de configurație **RAID 0** este utilizat în scopul îmbunătățirii **vitezei** de scriere și citire a datelor.

- Exemplu:

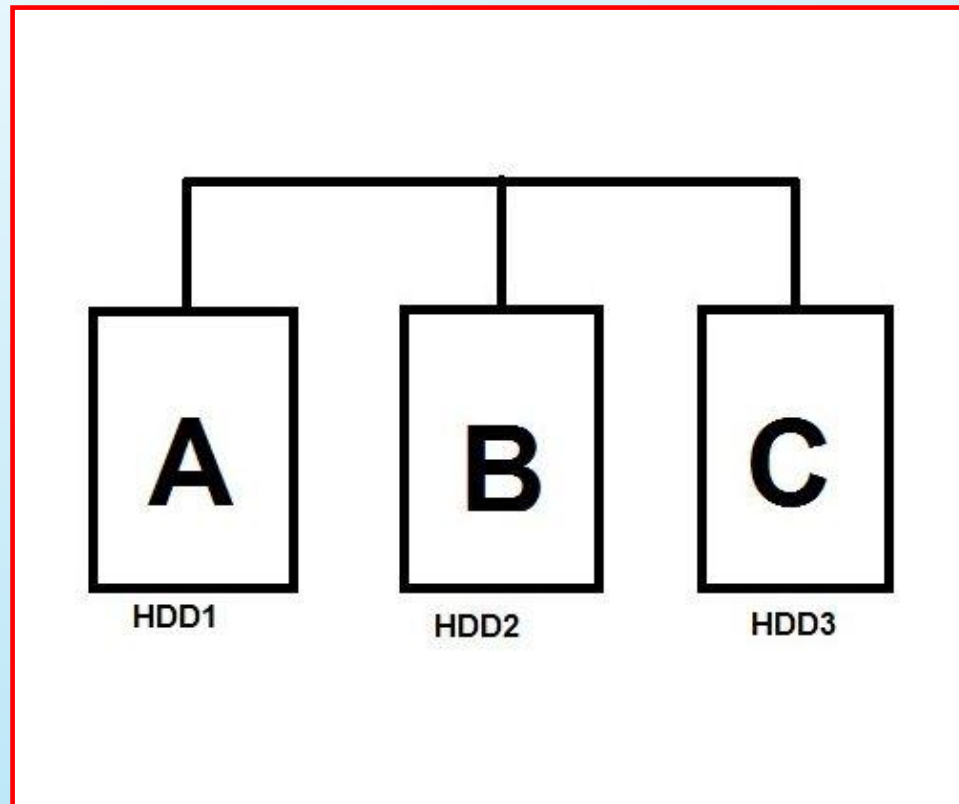
O configurație RAID 0 cu 3 HDD-uri.

Salvăm un fișier care este alcătuit din 3 părți (A,B și C).

Salvarea se efectuează în felul următor:

partea A a fișierului este salvată pe primul HDD,
partea B pe al doilea HDD, iar
partea C pe al treilea HDD.

File=A+B+C



Scrierea/citirea se face **simultan** pe mai multe HDD-uri, mărindu-se performanța totală a sistemului.

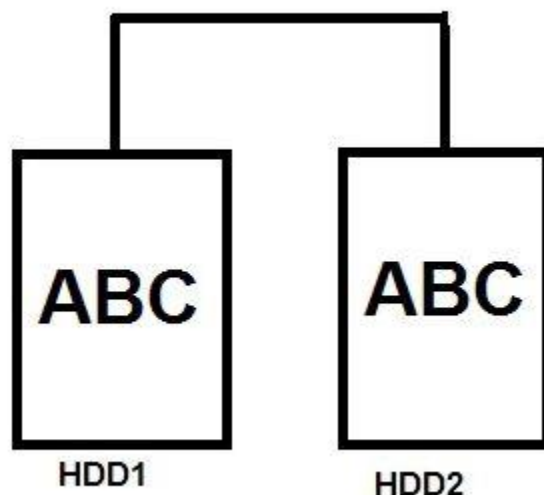
- RAID 0 este cea mai **rapidă** metodă de copiere, stocare și citire a datelor însă
- defectarea unuia dintre cele 3 HDD-uri atrage după sine **pierdere totală** a informațiilor salvate.

RAID 1

- RAID 1 se folosește mai ales când dorim o **securitate sporită** a datelor importante.
- Dezavantajul acestei configurații este că spațiul de stocare este de **jumătate** din suma capacităților HDD-urilor.



FILE=A+B+C



Exemplu: Se folosesc 2 HDD-uri pentru o configurație de tip **RAID 1**.

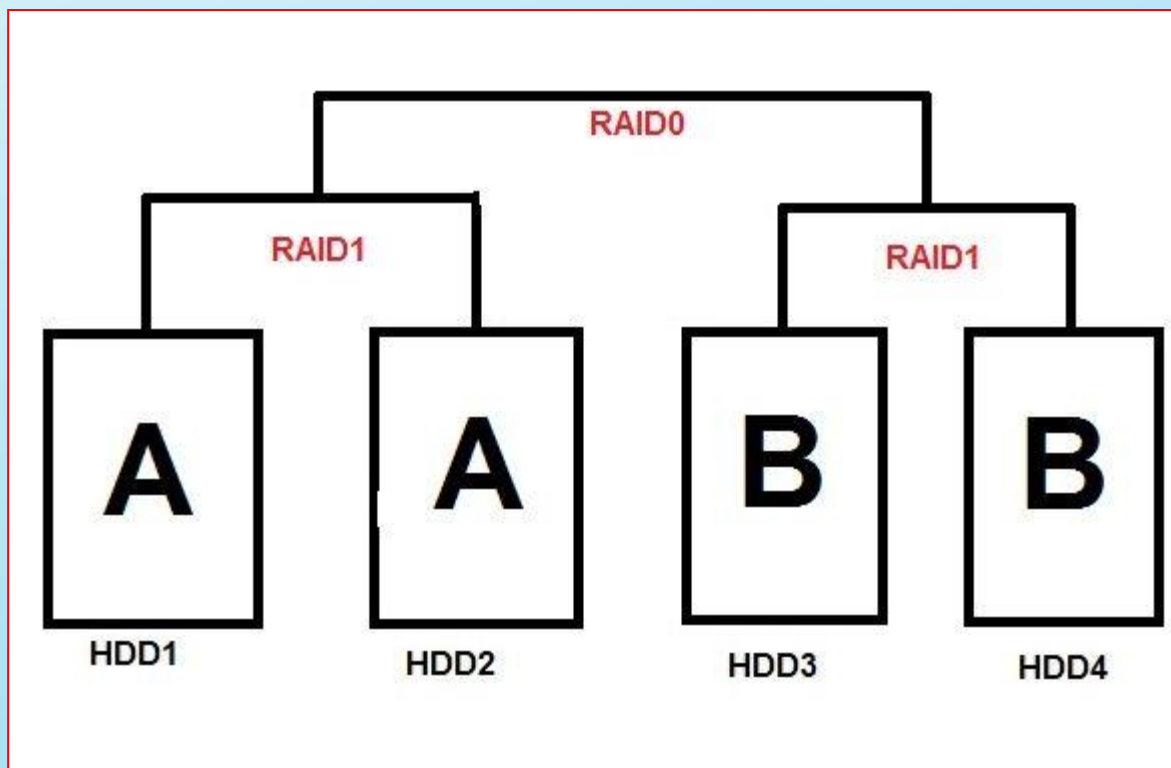
- Salvam același fișier cu 3 părți (A,B și C) ca și în exemplul anterior. Salvarea se face în felul următor:
Părțile A, B și C ale fișierului sunt salvate pe fiecare HDD în parte.
- Cu alte cuvinte fișierul este duplicat, îl avem pe ambele HDD-uri în aceeași formă. Dacă unul dintre cele 2 HDD-uri va ceda, atunci integritatea datelor (în exemplul nostru fișierul alcătuit din cele 3 părți: A, B și C) rămâne în siguranță deoarece fișierul există pe ambele HDD-uri.

RAID 10

- Această configurație este o combinație între RAID0 și RAID1.
- Este nevoie de **minim 4 HDD-uri**.
- RAID 10 oferă citirea datelor după modelul RAID 0, iar stocarea se face ca la RAID 1. În acest fel beneficiem simultan de **viteza RAID0** și **siguranța RAID1**.
- Dezavantajul este că putem folosi numai jumătate din capacitatea celor 4 HDD-uri.



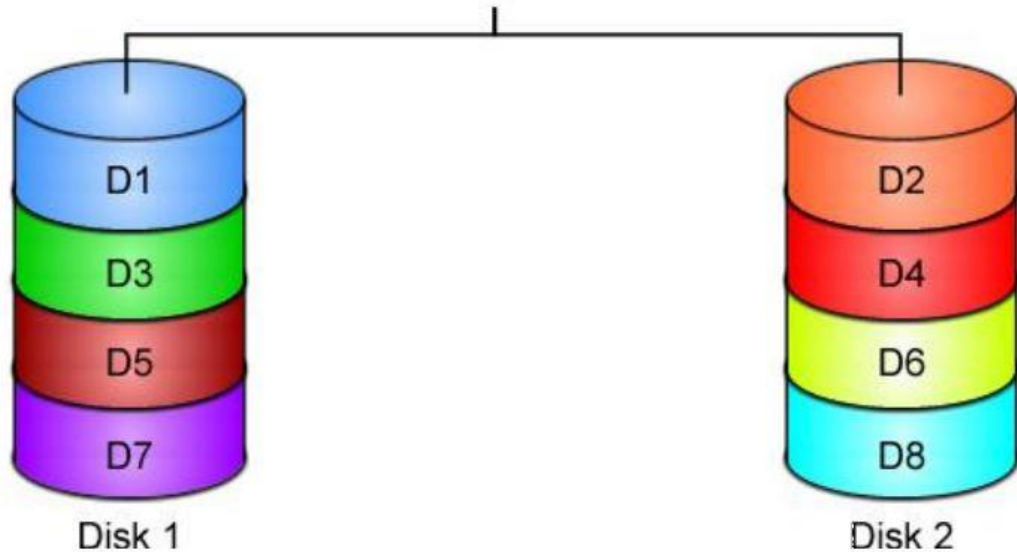
FILE=A+B



RAID 0

- RAID 0 nu este o arhitectură RAID în adevăratul sens al cuvântului deoarece nu asigură nicio redundanță a datelor.
- RAID 0 este folosit strict pentru maximizarea performanțelor de viteză în lucrul cu harddisk-urile.
- Datele sunt întrețesute (stripping) în mod secvențial pe mai multe discuri și sunt tratate ca un singur disc (sau **volum**) virtual.
- De obicei 4 discuri formează un volum.
- Implementările RAID 0 divizează volumele în **blocuri** și scriu datele în blocuri consecutive, localizate fizic pe discuri diferite din cadrul ariei de discuri.
 - de exemplu un bloc conține 512 sectoare, iar un sector 512 octeți

RAID 0



Felii: D1,D1D3,D4,D5,D6,D7,D8...

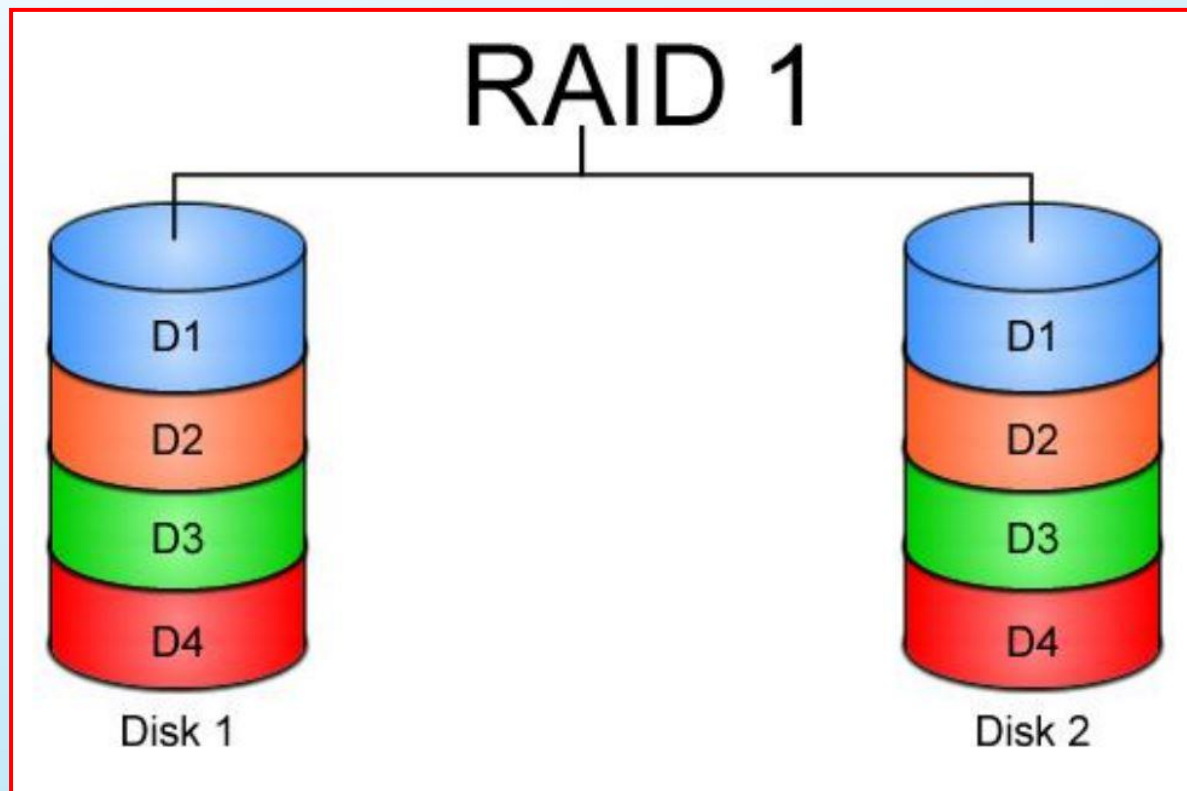
- Operațiile de scriere și citire au loc în paralel fiind executate SIMULTAN pe toate discurile aflate în volum.
- Mărimea blocurilor este definită de către utilizator și poate fi de 512 sectoare, spre exemplu.

- Utilizând această tehnică de întrețesere se obține o rată de transfer a datelor foarte ridicată, îmbunătățindu-se în felul acesta și performanța efectivă a sistemului.
- În eventualitatea defectării chiar și a unui singur disc, RAID 0 nu oferă redundanță datelor.
- În lipsa posibilității de regenerare a discului, datele ce se găsesc pe acesta se vor pierde. Din acest motiv, RAID 0 nu este folosit în aplicații de înaltă disponibilitate.



RAID 1

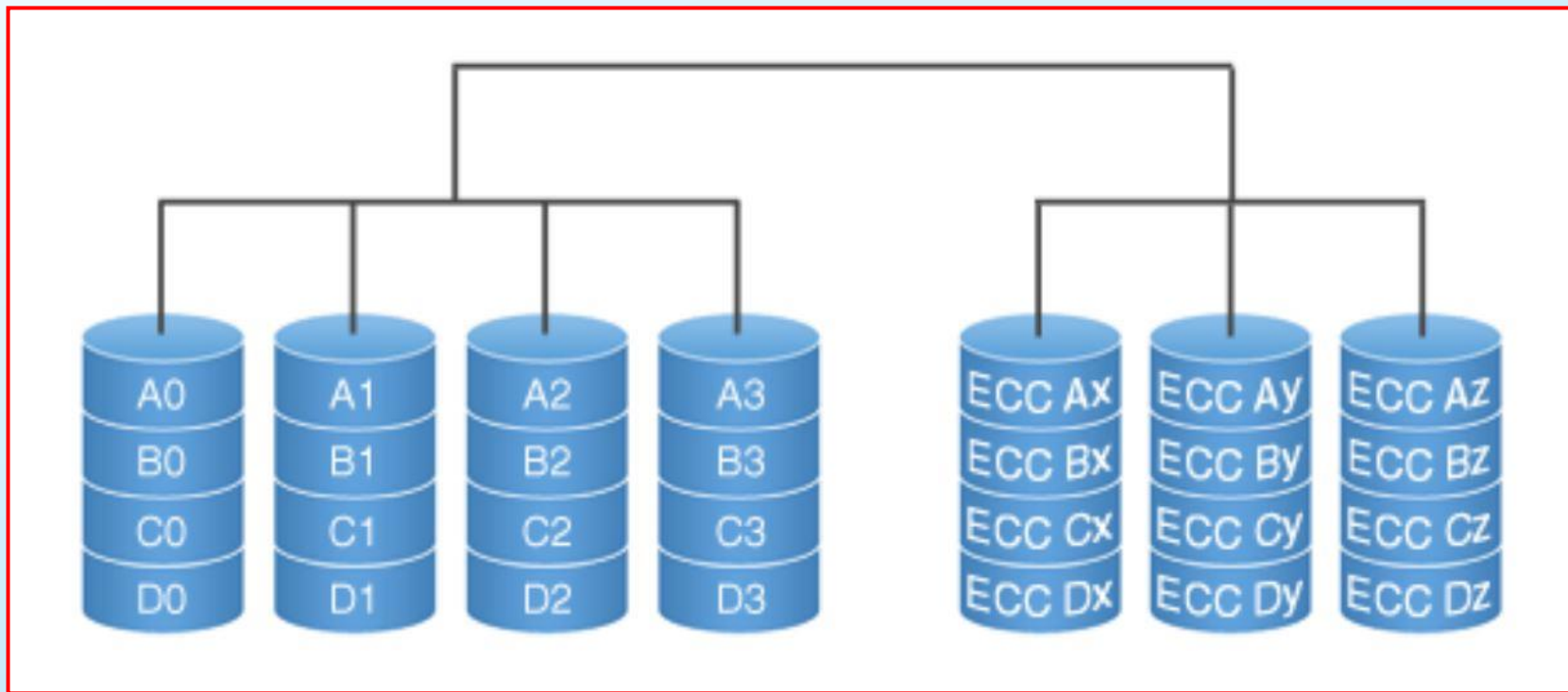
- RAID 1 asigură redundanța datelor prin oglindirea acestora (mirroring).
- Metoda aceasta creează două sau mai multe copii ale aceluiași volum de date și le plasează pe discuri separate. Fiecare copie este o imagine oglindită a celeilalte.
- RAID 1 poate scădea performanța generală a sistemului deoarece trebuie să efectueze tot atâtea scrieri câte copii sunt.
- Performanța citirilor este însă mult îmbunătățită deoarece **cererile sunt trimise simultan** controlerelor de discuri ale tuturor copiilor. Discul care răspunde primul cererii de citire va returna datele sistemului.



- RAID 1 este cea mai scumpă implementare deoarece asigură **redundanța** datelor în proporție de 100%
- Poate suporta **mai multe defecte** simultan.
În caz de defectare, datele nu trebuie reconstruite ci pur și simplu **copiate**.
- asigură **cea mai mare disponibilitate** a datelor prin utilizarea celui mai mic număr de discuri din configurație.

RAID 2

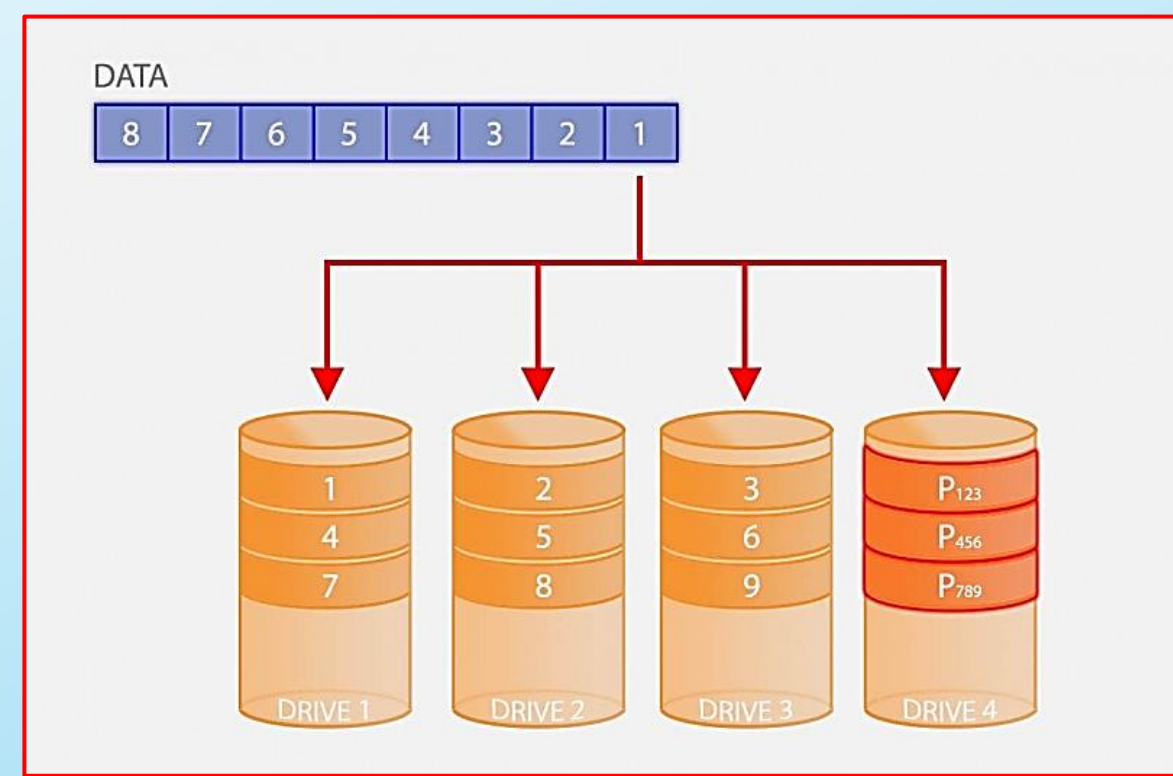
- Acest model presupune că se distribuie datele pe mai multe discuri în paralel, efectuându-se o **întrețesere la nivel de bit**,
 - În cadrul ariei de discuri sunt utilizate și discuri de verificare.
 - Citirea și scrierea datelor se face utilizând **tehnici de codificare** bazate pe coduri corectoare de erori de tip **Hamming** pentru a asigura detecția și corecția erorilor.



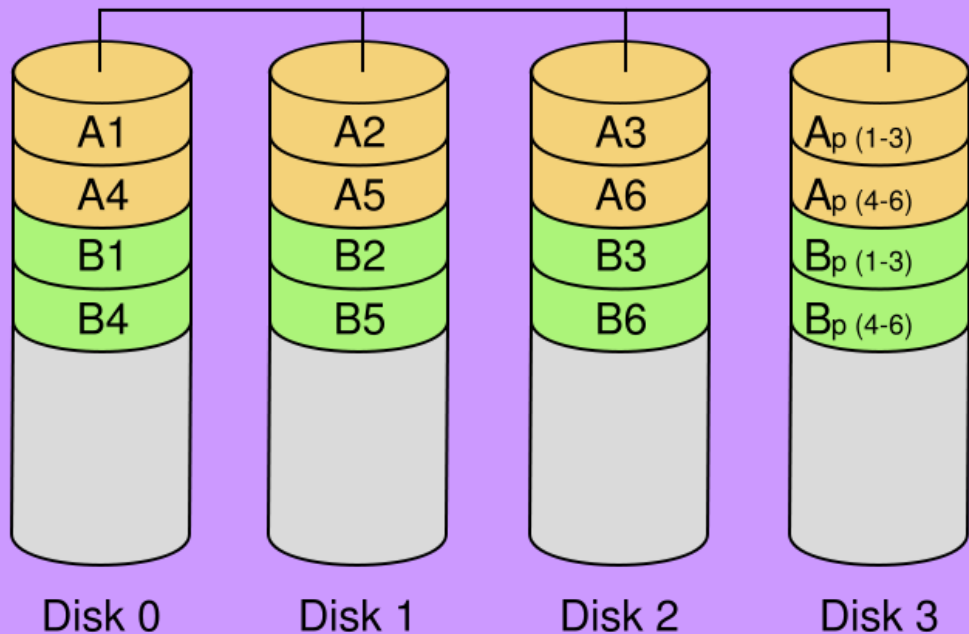
- RAID 2 are mai multe dezavantaje decât avantaje.
- Printre acestea trebuie amintit faptul că:
 - utilizează tehnici de codare bazate pe coduri corectoare de erori de tip Hamming
 - necesită utilizarea unor grupuri mari de discuri pentru a asigura consistența
(ex: 4 discuri de verificare pentru 10 discuri de date, 5 discuri de verificare pentru 25 de discuri de date).
 - rata de transfer este cel mult egală cu cea a unui disc
- Din cauză că această tehnică de codificare este foarte complexă și scump de implementat, RAID 2 nu prezintă un real interes pentru mediul comercial.

RAID 3

- RAID 3 folosește doar **un singur disc** dedicat pentru memorarea **informațiilor de paritate**, celelalte discuri din cadrul ariei de discuri fiind folosite pentru memorarea informației utile ce este împrăștiată pe toate discurile ariei. (exact ca și în cazul RAID 2).
- Acesta efectuează un **SAU EXCLUSIV** asupra datelor și nu un cod corector de erori.
- RAID 3 lucrează la nivel de **bit**



RAID 3



- Cantitatea minimă de date care este scrisă sau citită într-o operație de intrare sau de ieșire este egală cu numărul de discuri din arie înmulțit cu numărul de octeți ai sectorului discului.
- Aceasta este cunoscută ca **unitate de transfer**.

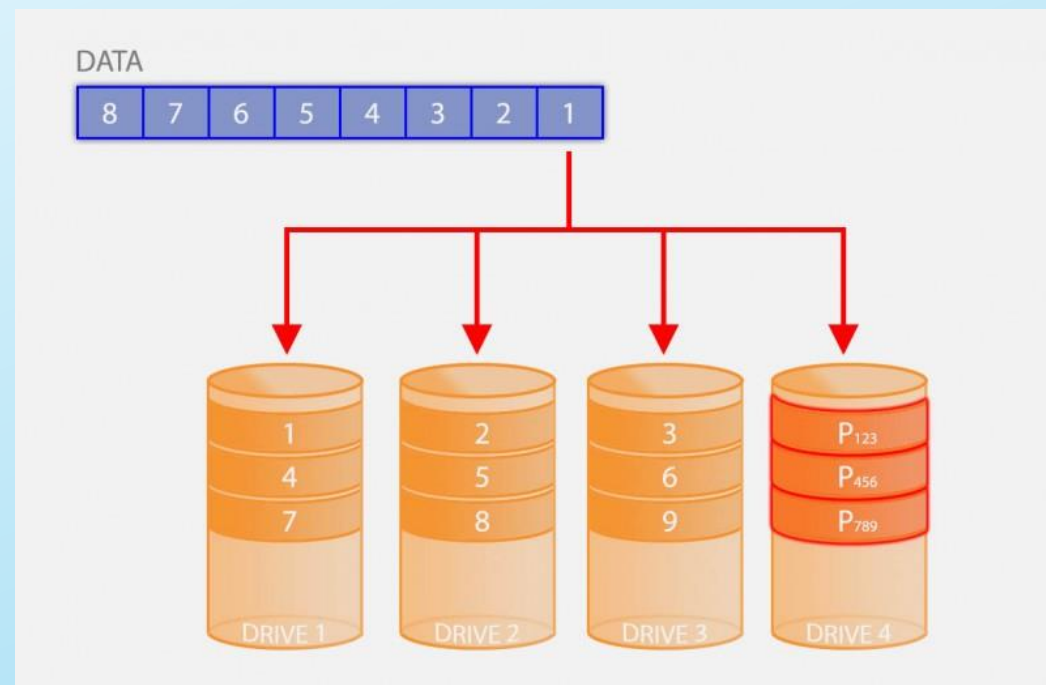
✖ RAID 3 oferă și posibilitatea **înlocuirii la cald** a unui disc și recuperarea datelor.

- La un moment de timp dat nu poate fi activă decât o singură cerere de intrare sau de ieșire deoarece capetele de acces se mișcă în paralel în aria de discuri.
- Acest lucru ne asigură o **rată bună** de transfer a datelor atunci când datele accesate **sunt dispuse secvențial pe disc**, dar face ca RAID 3 să nu fie suficient de practic pentru aplicații cu date dispuse aleator în aria de discuri.
- O rată de transfer foarte bună se obține atunci când unitatea de transfer are aceeași mărime cu dimensiunea blocurilor de date care se scriu sau se citesc. Cu cât blocurile transferate sunt de dimensiuni mai mici, cu atât rata de transfer scade.



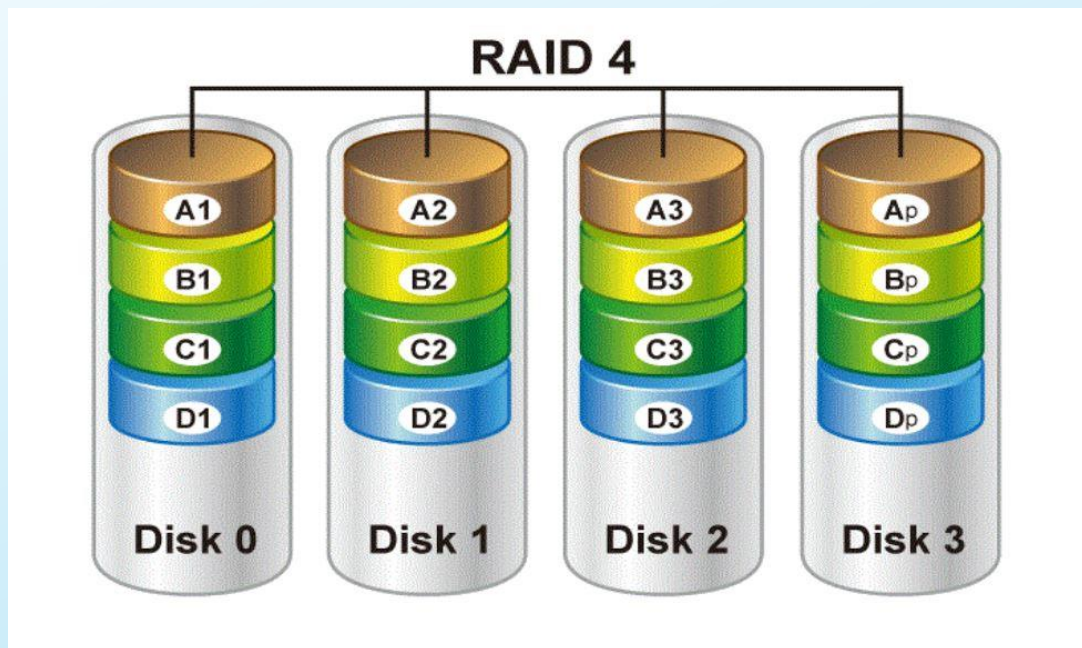
RAID 4

- În cazul RAID 4 datele sunt distribuite la fel ca și în cazul RAID 3 pe mai multe discuri de date și pe un disc de paritate.
- Diferența constă în faptul că **un bloc** de date se află **pe un singur disc**, astfel încât citirea se face accesând un singur disc și nu mai multe.
- Aria de discuri poate prelucra mai multe cereri deodată întrucât discurile nu se mai utilizează în paralel.
- RAID 4 lucrează la nivel de **bloc**



Date scrise astfel: blocul 1 merge la unitatea 1, blocul 2 la unitatea 2, blocul 3 la unitatea 3 și paritatea calculată la unitatea 4.

Dacă o unitate eșuează, datele „lipsă” pot fi recalculate pe baza datelor rămase.



Discul de paritate devine un element de scădere a performanței deoarece este implicat în toate operațiile de scriere.

Acest lucru se întâmplă deoarece paritatea, pentru fiecare operație de scriere, se calculează având în vedere și informația de pe celelalte discuri, în sectoarele de date asociate.

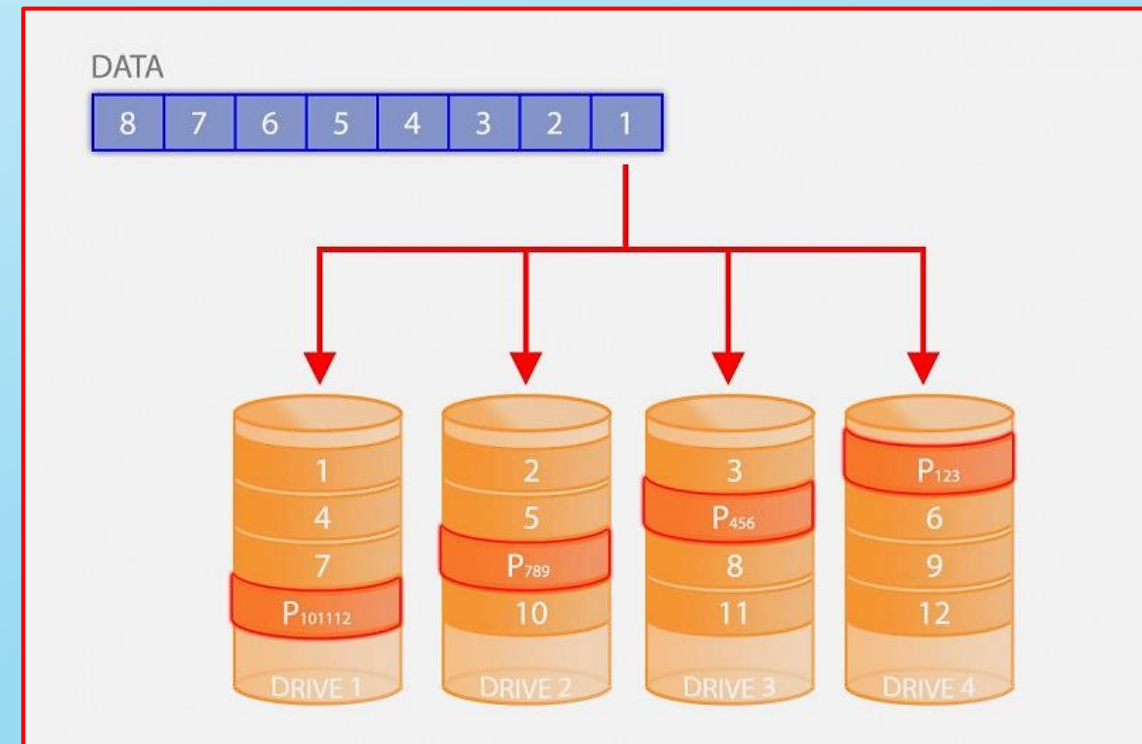
Această procedură face ca RAID 4 să fie considerat nepractic.

RAID 5

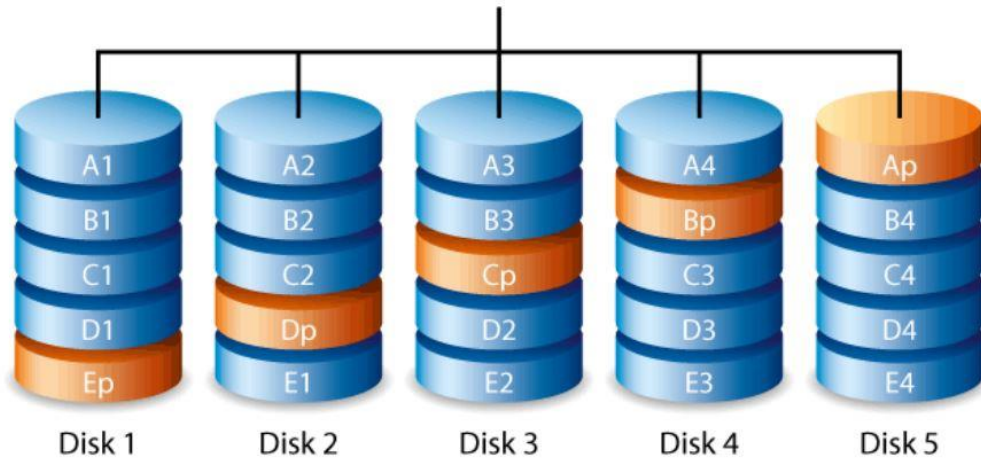
- În cazul RAID 5 datele sunt distribuite pe mai multe discuri de date și pe un disc de paritate.
- RAID 5 **nu are un disc dedicat** special numai pentru informația de **paritate**, ci distribuie informațiile de tip date și informațiile de tip paritate pe toate discurile ariei.
- RAID 5 permite **mai multe operațiuni de acces concurente** la dispozitivele din cadrul ariei de discuri, fiind satisfăcute astfel multiple cereri concurente de intrare sau ieșire.

Datele se recuperează ca și în cazul RAID 4

Aria suporta un singur disc defect.



RAID 5



Procesorul efectuează toate calculele pentru paritate, iar în cazul unei defecțiuni, operațiunile de intrare sau de ieșire sunt de până la de 3 ori mai consumatoare de resurse.

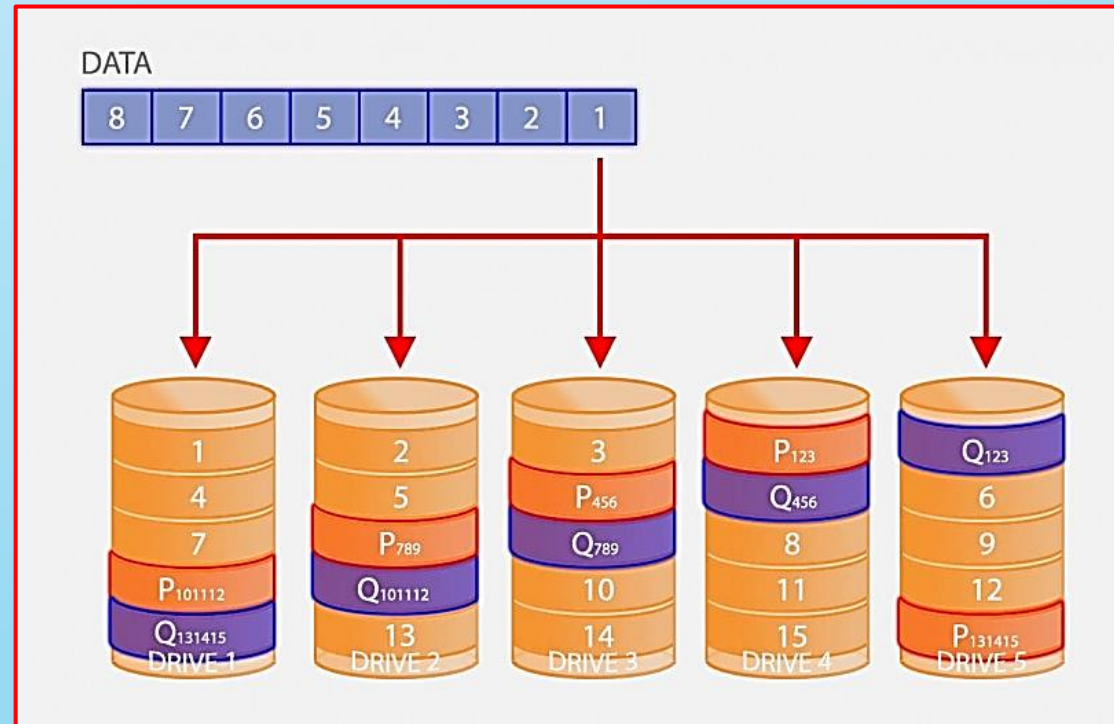
Având un cost relativ scăzut per GB utili și o performanță medie foarte bună, RAID 5 este utilizat cu succes pentru toate tipurile de transfer (servere pentru baze de date raționale, servere de fișiere)

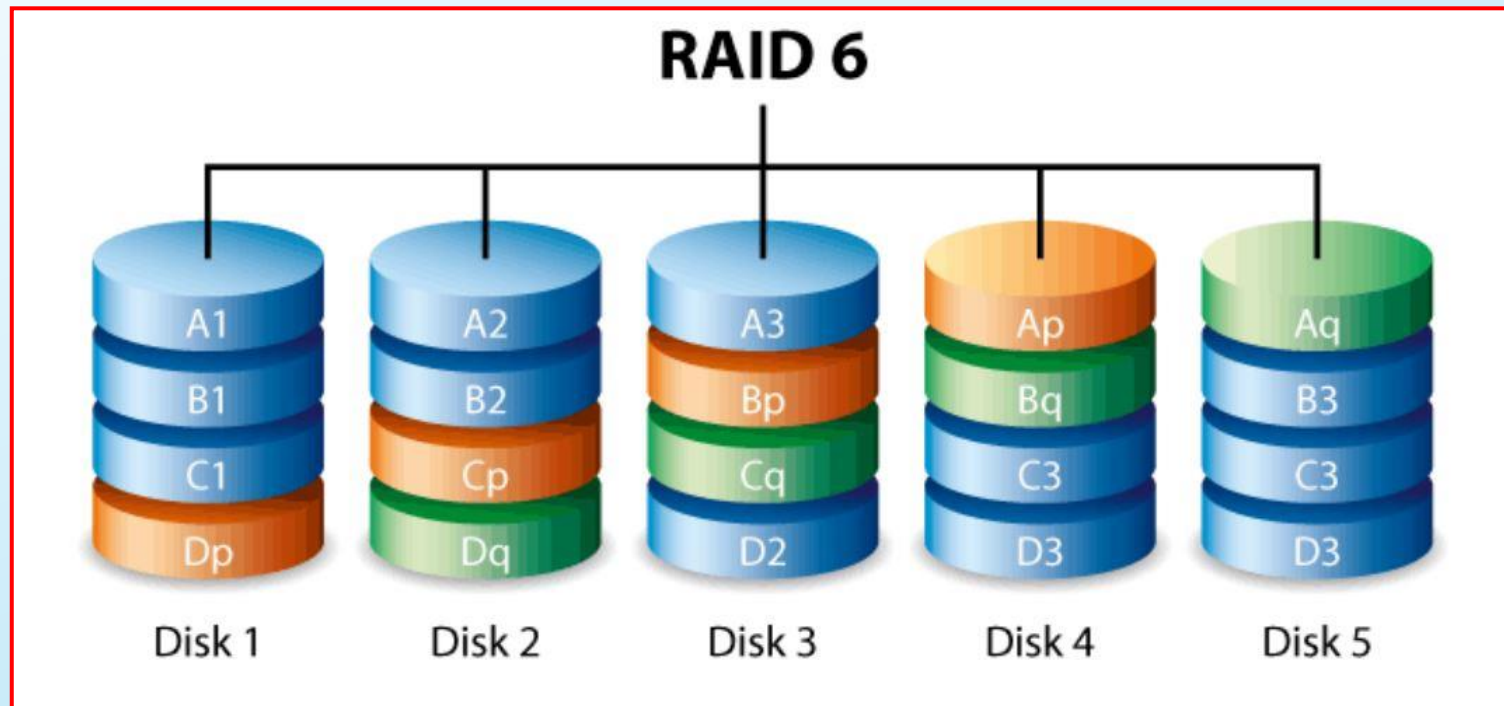
- Astfel, **performanța** transferurilor pentru aceste arii de discuri crește semnificativ, acestea devenind cele mai potrivite și utile în cazul operării cu blocuri de dimensiune mică, sau cu **datele dispuse în mod aleator** pe discuri. Ciclul de scriere se realizează în 3 pași:
 - Este citită paritatea existentă și data ce urmează a fi schimbată. Vechea dată din valoarea parității existente este scăzută și apoi se memorează paritatea corectă într-un buffer.
 - Este calculată noua paritate pe baza valorii din buffer și a valorii noii date de memorat și se scriu pe discurile corespunzătoare noua dată și paritatea actualizată.
- Implementările software ale RAID 5 sunt suportate de către multe sisteme moderne de operare (Linux, Windows Server etc.).



RAID 6

- Ceea ce diferențiază nivelul RAID 6 de nivelul RAID 5 este faptul că se calculează și se stochează paritate dublă pentru date.
- Sunt necesare **minim 4 discuri**, performanța scăzând datorită calculului dublu de paritate.
- Redundanța crește, aria suportând **două discuri defecte**.

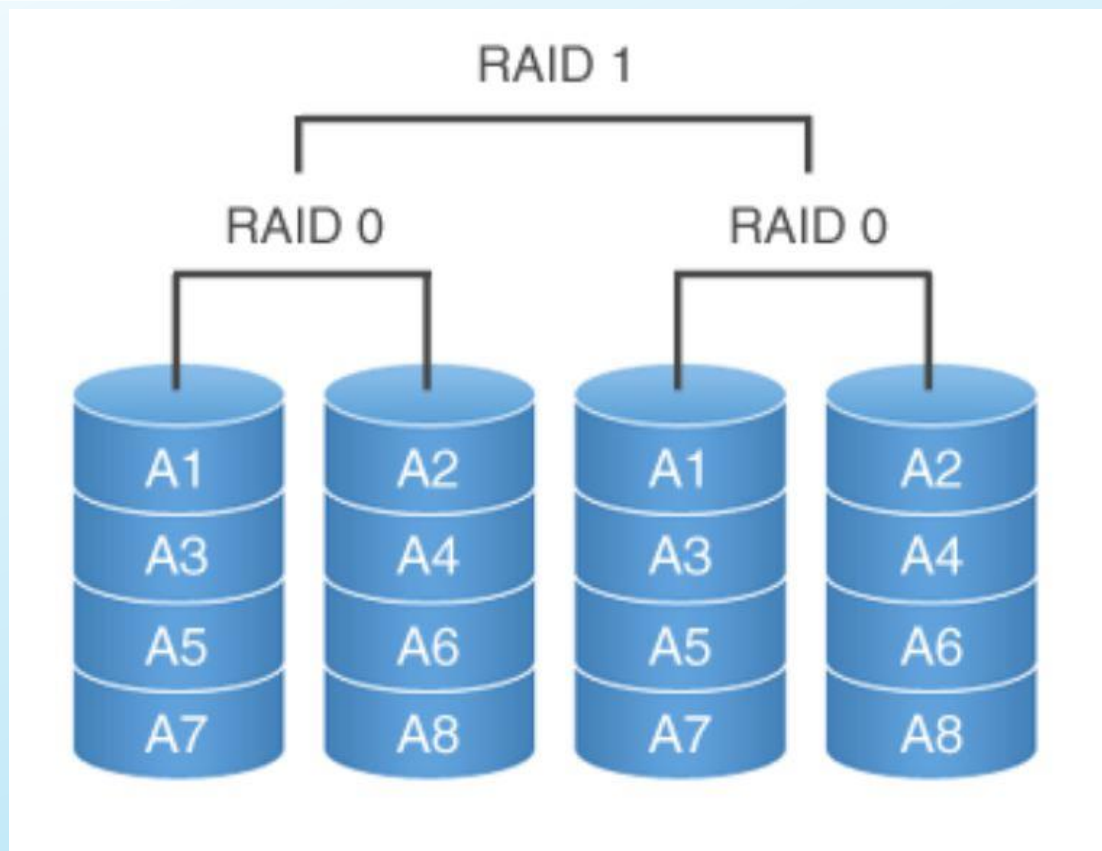




- Reconstrucția datelor poate fi un mare consumator de resurse, performanța fiind puternic afectată, dar redundanța dublă se compensează prin posibilitatea de amânare a procesului de rebuilding pentru un plus de viteză.
- Pentru a crea nivele etajate de RAID, oricare nivel poate fi combinat cu alt nivel, dar cel mai des sunt folosite nivelele 0 și 1.

RAID 01

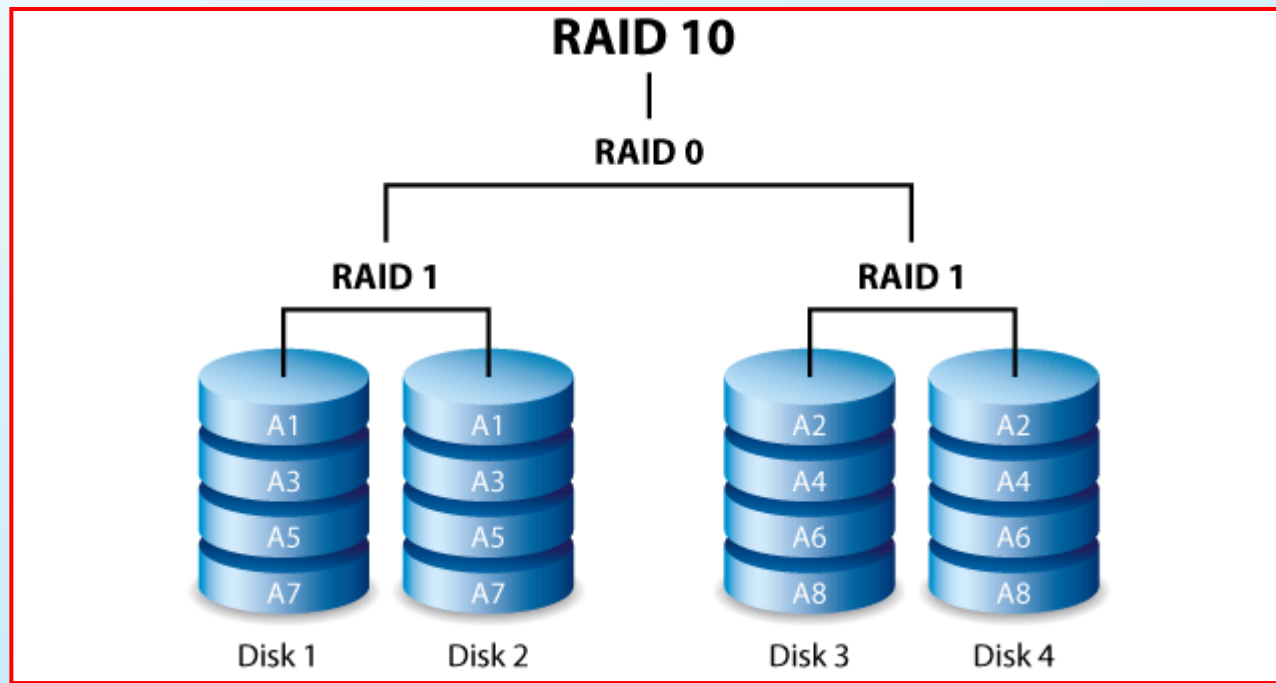
- Acest nivel dispune **minim 4 discuri** grupate în fâșii (oglindite),
- Înlătură neajunsurile și îmbunătățește performanțele, însă crește complexitatea.
- RAID 0+1 creează un al doilea set de fâșii oglindite cu un set primar de fâșii de discuri.



- Matricea continuă să funcționeze cu unul sau mai multe unități de stocare afectate în aceeași oglindă stabilită, dar dacă driverul cedează de pe ambele părți ale oglinzii, datele de pe sistemul RAID sunt pierdute

RAID 10

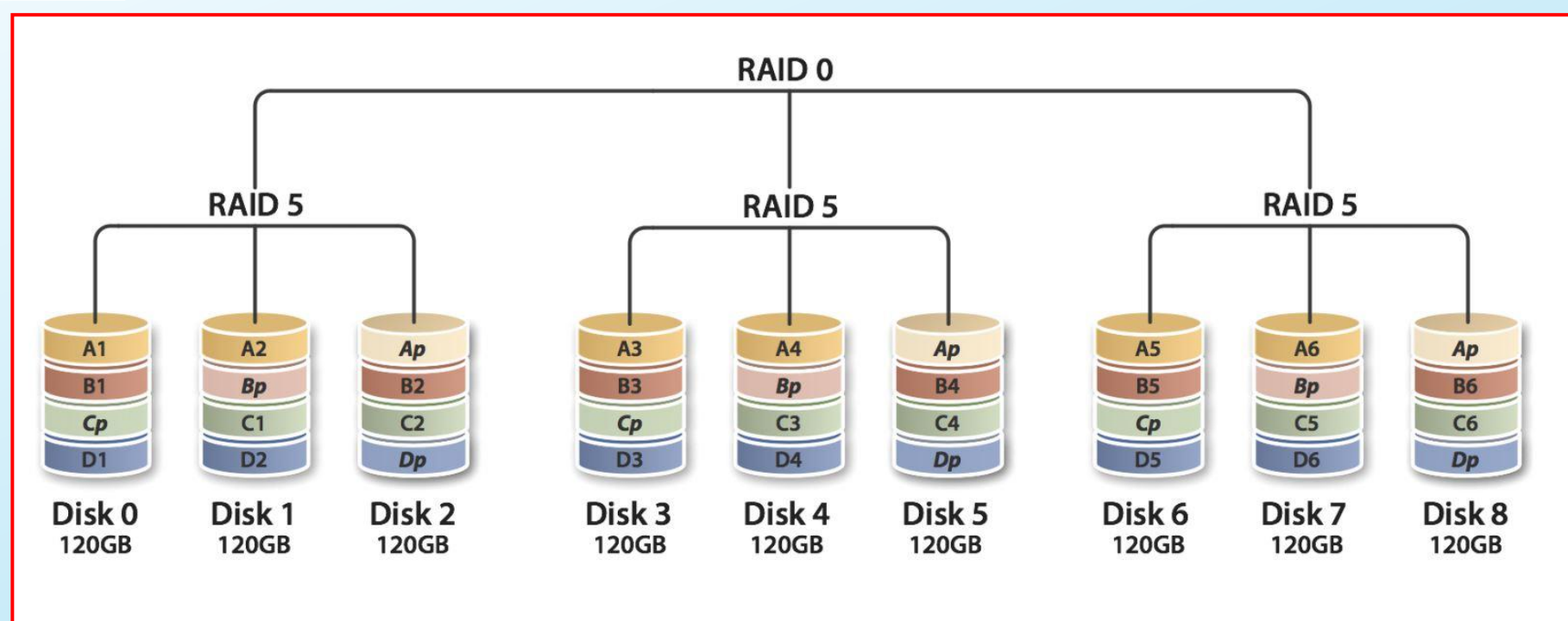
- Acest nivel dispune **minim 4 discuri** grupate în fâșii (oglindite), înlătură neajunsurile și îmbunătățește performanțele, însă crește complexitatea.
- Diferența față de RAID 0+1 este că RAID 1+0 creează discuri cu date întrețesute de la o serie de unități de stocare în oglindă.



- Într-o situație de eșuare, RAID 1+0 funcționează mai bine, deoarece toate celelalte discuri continuă să fie utilizate.
- Matricea poate susține mai multe discuri pierdute, atât timp cât oglinda nu pierde toate unitățile sale

RAID 50

- RAID50 combină feliile de blocuri de la nivelul RAID0 cu paritatea distribuită de RAID5
- Aceasta este de fapt o matrice RAID 0 întretesută de-a lungul elementelor RAID 5.
- Pentru a realiza un nivel hibrid RAID 50 avem nevoie de **cel puțin 6 discuri**.
- Avantajul acestei configurații este că ca pot eșua câte un drive din fiecare set, fără să se piardă date.

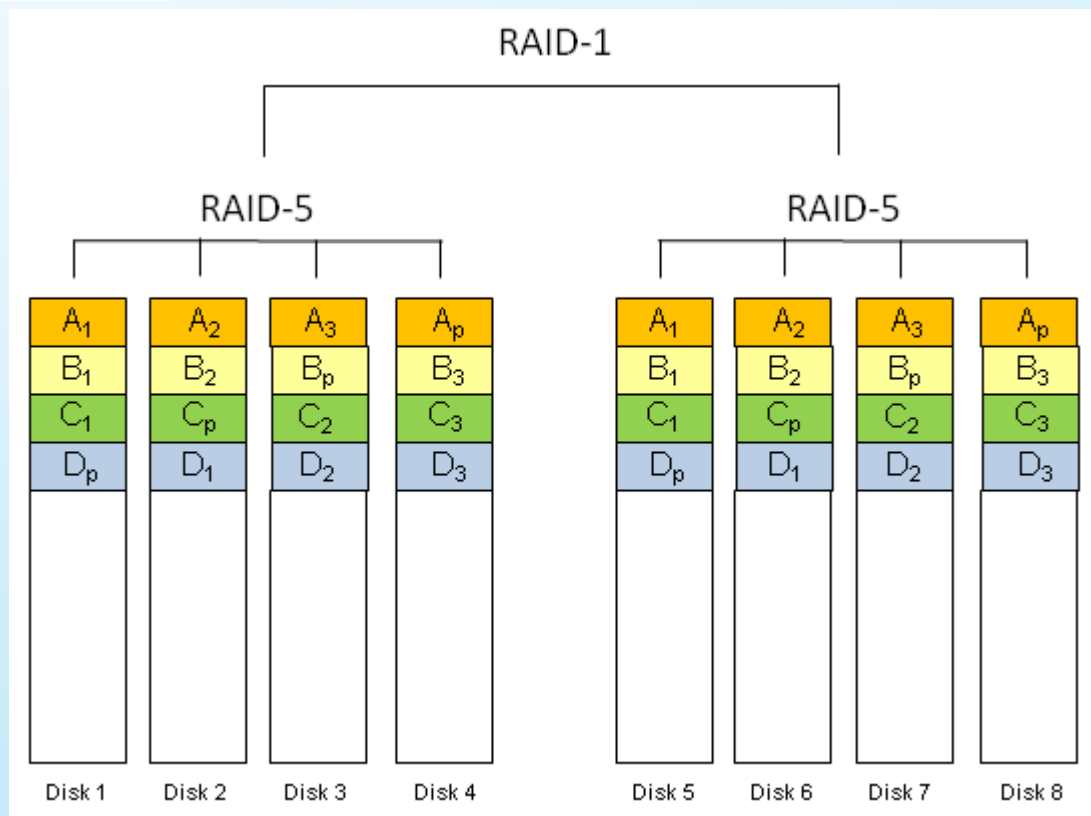


- Unitățile rămase în acel set devin un punct slab pentru întreaga matrice până când discul defect este înlocuit.
- Este suficient să se mai defecteze un singur disc (pe lângă cel inițial) și întreg sistemul cedează, datele stocate în întreaga matrice fiind pierdute.
- Timpul petrecut în procesul de rebuilding reprezintă o perioadă de vulnerabilitate a setului RAID.



RAID 51

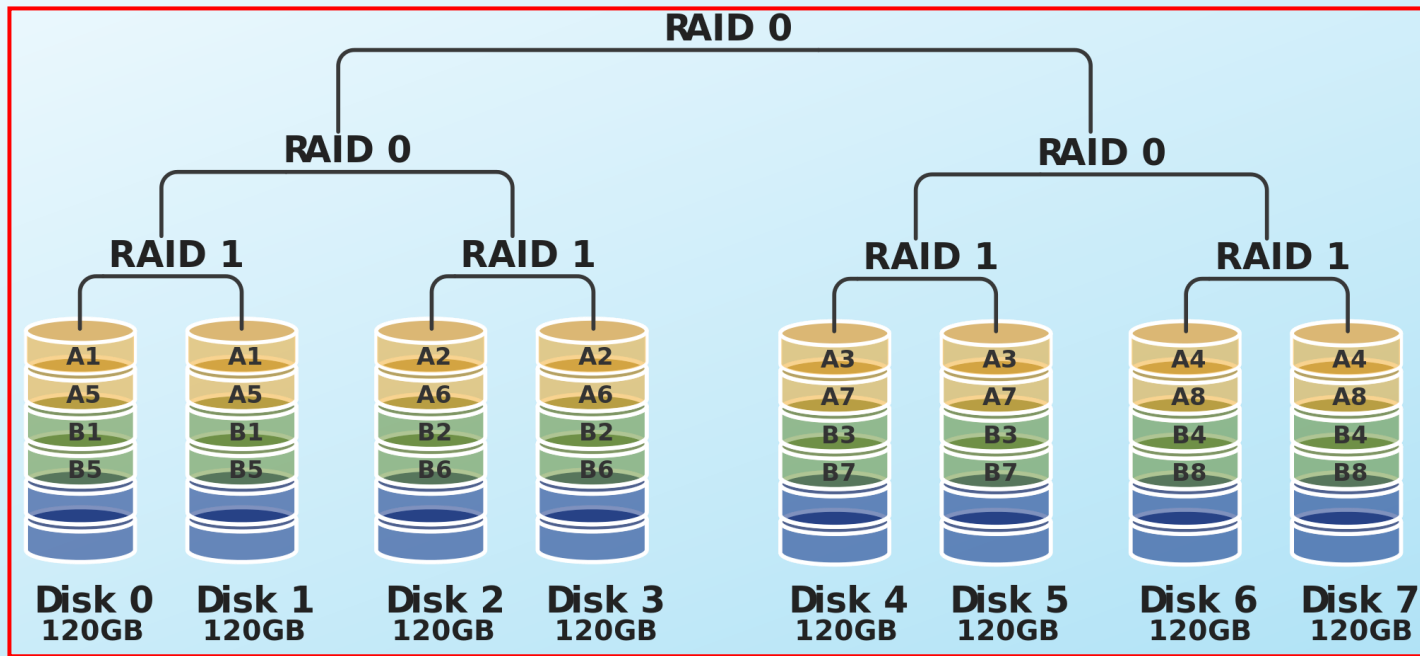
- Acest nivel hibrid este o matrice formată din două matrice RAID 5, care sunt oglindite una față de cealaltă.
- În general această configurație este folosită astfel încât fiecare set de RAID 5 se află pe un controller separat.
- Scrierile și citirile sunt echilibrate în ambele matrice RAID 5.



- Unele controllere suportă RAID 51 pe mai multe canale pentru a ține sincronizate părțile diferite.
- Această configurație poate susține eșecul oricărui disc din orice matrice, plus încă un disc suplimentar din cealaltă matrice înainte să existe pierderi de date.
- Spațiul maxim ce se poate găsi pe un RAID 51 este egal cu dimensiunea unui set individual RAID 5

RAID 100

- Un raid 100, numit uneori și RAID 10+0, este o felie de tipuri RAID 10.
- Din punct de vedere logic el este de fapt o matrice RAID 10 implementată cu ajutorul software-ului RAID 0 peste hardware-ul de RAID 10.



- Principalul avantaj al RAID 100 față de un singur nivel RAID este răspândirea încărcării pe mai multe controllere, obținându-se astfel performanță mai bună la citiri aleatoare și atenuarea riscurilor pe matrice.
- RAID 100 reprezintă cea mai bună alegere când vine vorba de baze de date foarte mari, unde controllerele hardware de RAID limitează numărul de discuri fizice permise în fiecare matrice standard

