

Problema unităților multiple care înlocuiesc o singură unitate este fiabilitatea. Presupunând că fiecare unitate, de capacitate mare sau mică, are aceeași rată de defectare, existența unităților multiple va crește în mod semnificativ rata de defectare față de cazul în care există o singură unitate de capacitate mare. Deci, utilizarea eficientă a matricilor de discuri implică două aspecte:

1. Distribuirea eficientă a datelor pe ansamblul unităților pentru a utiliza potențialul de acces permis de unitățile multiple.
2. Introducerea judicioasă a redundanței în cadrul matricii de discuri pentru a restabili fiabilitatea globală a ansamblului.

Dificultatea asigurării cerințelor precedente constă în faptul că accesul la matricea de discuri nu este uniform și nici staționar. Anumite modele de acces pot favoriza un anumit aranjament al unităților într-o matrice, dar atunci când ulterior datele sunt accesate într-un mod diferit, poate fi preferabilă o altă combinație de distribuție și sincronizare.

5.8.4. Redundanța în matricile de discuri: RAID

Ansamblurile de discuri pot asigura o îmbunătățire semnificativă a performanțelor în anumite aplicații particulare. Totuși, unitățile multiple implică puncte multiple de defectare. Indicatorul MTBF al unei matrici de discuri este:

$$\frac{MTBF \text{ pentru un singur disc}}{\text{Numarul de discuri din matrice}}$$

Aceasta presupune că defectele sunt independente. Dacă un singur disc are o rată de defectare de 300.000 de ore, o matrice de 100 de discuri ar avea o rată de defectare de 3.000 de ore, deci de 4 luni. Deoarece subsistemul de discuri reprezintă adesea singurul mediu complet nevolatil din sistem, integritatea acestui subsistem este foarte importantă. Se pot introduce discuri suplimentare conținând informații redundante în scopul păstrării integrității matricii de discuri.

Termenul RAID (*Redundant Array of Inexpensive Disks*) a fost introdus de R. Katz pentru a descrie mai multe tehnici, numite de acesta "nivele", în scopul îmbunătățirii integrității datelor stocate într-o matrice de discuri [9]. Metoda de bază descrisă constă în descompunerea matricii în grupuri, fiecare grup având un disc redundant conținând anumite informații de control. Atunci când un disc se defectează, sistemul detectează defectul, și prin utilizarea unei *rezerve calde* reface informația de pe discul defectat, utilizând discurile valide rămase din grup și discul de control. Informația refăcută care exista pe discul defect este scrisă pe discul de rezervă. Ulterior, în timpul unei proceduri de întreținere, discul defect este înlocuit cu un nou disc de rezervă.

R. Katz a descris șase nivele RAID, cu avantajele și dezavantajele fiecăruia atât în mediile de procesare a tranzacțiilor, cât și în cazul supercalculatoarelor. Mediul de procesare a tranzacțiilor constă din numeroase accesuri la fișiere de dimensiuni mici, în timp ce aplicația de acces la supercalculatoare constă dintr-un număr redus de accesuri la fișiere foarte mari.

Există mai multe implementări standard ale nivelelor RAID, ca și diferite versiuni nestandard. Nivelele RAID nu sunt ierarhice din punct de vedere calitativ, astfel că, de exemplu nivelul RAID 2 nu este în mod necesar superior nivelului RAID 1, ci este diferit de acesta. În continuare sunt descrise pe scurt nivelele RAID [9].

Nivelul RAID 1

Constă din discuri redundante (Figura 5.16). În acest caz, fiecare disc de date are un disc oglindă (*mirror disk*) pe care sunt duplicate toate informațiile de pe discul de date original. Toate operațiile de scriere pe discul de date se execută și pe discul redundant. Această metodă asigură o securitate ridicată a datelor, deoarece chiar și în cazul defectării totale a unui disc, nu există pierderi de date datorită duplicării. Prin această metodă se îmbunătățesc și performanțele la citire: deoarece există două discuri identice, datele cerute pot fi citite de pe discul care este disponibil sau de pe cel cu ansamblul de capete de citire/scriere care este mai apropiat de cilindrul care conține datele.

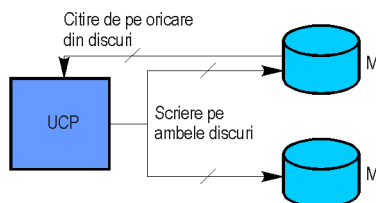


Figura 5.16. RAID 1: Fișierul *M* este duplicat pe două discuri.

Nivelul RAID 1 a fost promovat de *Tandem Computers*, care a utilizat această metodă și pentru controlere și magistrale de I/E. Dezavantajul nivelului RAID 1 este că se dublează costul subsistemului de discuri. De asemenea, se reduc performanțele la scriere, deoarece fiecare operație de scriere trebuie executată de două ori. La majoritatea serverelor, numărul operațiilor de citire depășește însă cu mult numărul operațiilor de scriere.

Nivelul RAID 2

Este o matrice întrepesută la nivel de bit, organizată în mod similar cu o memorie la care se utilizează coduri *Hamming* (Figura 5.17). Trebuie satisfăcut criteriul ca numărul biților de corecție, egal cu numărul discurilor de corecție k , să satisfacă relația:

$$2^k \geq k + m + 1$$

unde m este numărul discurilor de date. Dacă există 10 discuri de date, sunt necesare $k = 4$ discuri redundante.

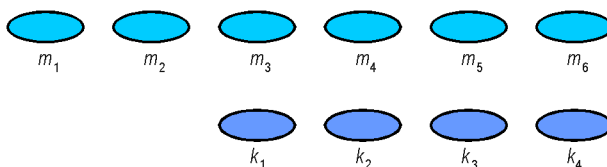


Figura 5.17. RAID 2: Fișierul $M = m_1, m_2, \dots, m_6$, este distribuit pe mai multe discuri. Pe fiecare disc sunt păstrați biți de corecție separați, $K = k_1, k_2, \dots, k_4$.

Cu toate că reduce costurile, RAID 2 are unele dificultăți. Acest nivel este potrivit numai matricilor de discuri sincronizate (*ds*). Fișierul minim trebuie să conțină 10 blocuri de date, deoarece acestea vor reprezenta unitatea minimă de scriere pe disc. Nivelul RAID 2 este util mai mult pentru aplicații de tip supercalculator, fiind utilizat de exemplu de *Thinking Machines* în mai multe sisteme ale acesteia.

Nivelul RAID 3

În cazul nivelului RAID 3, se introduce un disc de paritate (Figura 5.18), care creează un tip de cod de corecție. Datele sunt distribuite la nivel de bit pe mai multe discuri, ultimul disc conținând paritatea înregistrărilor de biți. O eroare este detectată la accesarea unui disc individual, iar discul de paritate poate fi utilizat pentru a reconstrui informația defectă. Utilizarea unui disc de paritate necesită accesarea acestuia la toate operațiile de scriere. Ca și RAID 2, nivelul RAID 3 este potrivit în general numai pentru matrici de discuri sincronizate.

Nivelul RAID 4

Este similar cu RAID 3, exceptând faptul că în locul distribuirii datelor pe discuri la nivel de bit, fiecărui disc i se alocă un bloc întreg al înregistrării (Figura 5.18). Astfel se îmbunătățește abilitatea sistemului de a accesa fișiere de dimensiuni mici (de exemplu, fișiere formate

dintr-un singur bloc), deoarece este necesară implicarea unui singur disc de date și a discului de paritate în procesul de acces.



Figura 5.18. RAID 3 și RAID 4: Fișierul $M = m_1, \dots, m_n$ este distribuit pe n discuri (la nivel de biți la RAID 3, la nivel de blocuri la RAID 4). Discul de paritate (k) păstrează paritatea înregistrărilor.

Nivelul RAID 5

Este ilustrat în Figura 5.19. În cazul nivelurilor RAID 3 și RAID 4, toate accesurile necesită un disc de paritate, care devine o zonă îngustă a sistemului. Se poate crea un sistem în care informația de paritate a sectoarelor este distribuită pe o serie de discuri. Aceasta îmbunătățește numărul de accesuri simultane care se pot efectua la matricea de discuri.

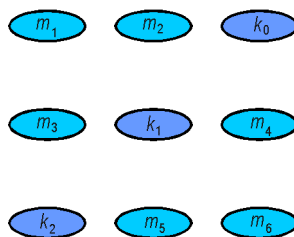


Figura 5.19. RAID 5: Fișierul $M = m_1, \dots, m_n$ este distribuit pe n discuri într-o matrice de $p \times p$, unde $(p-1) \times (p-1) = n$. Paritatea sectoarelor este distribuită pe p discuri. k_0 este paritatea între m_1 și m_2 , k_1 este paritatea pentru m_3 și m_4 etc.

Nivelul RAID 5 reprezintă un compromis adecvat între cost, performanțe și fiabilitate, fiind probabil cel mai popular nivel utilizat în cazul rețelelor locale de dimensiuni mici și medii. Acest nivel, ca și nivelele RAID 3 și RAID 4, nu oferă același grad de siguranță ca și nivelul RAID 1, deoarece nu există o redundanță completă, dar implementarea este mai ieftină, iar performanțele globale sunt teoretic mai bune atât la citire, cât și la scriere.

Nivelul RAID 6

Constă în avansarea nivelului RAID 5 la două sau mai multe dimensiuni de paritate. Se poate considera că există discuri pe mai multe dimensiuni în care se poate crea o paritate atât pe linie cât și pe coloană.

Problema diferitelor nivele superioare de redundanță este că, similar cu configurațiile matricilor de discuri, acestea nu sunt neapărat robuste față de diferitele distribuții de fișiere. Dacă un fișier este organizat în mod optim ca un fișier distribuit, iar un alt fișier este organizat în mod optim dacă este accesat printr-o matrice de discuri sincronizate, există dileme privind aranjarea atât a matricii de discuri, cât și a mecanismului de redundanță.

Nivelele RAID pot fi implementate fie prin hardware, fie prin software. Implementarea hardware se realizează fie printr-un controler RAID dedicat, fie printr-un adaptor SCSI. Adaptoarele SCSI diferă în mare măsură în privința nivelelor RAID pe care le implementează. Nivelele RAID pot fi implementate și prin software, de sistemul de operare. Unul din sistemele de operare care implementează nivelele RAID este *Windows NT*.

5.9. Interfețe pentru unitățile de discuri

Sarcina principală a unei interfețe sau a unui controler de disc este de a transmite și de a prelua date la și de la unitatea de disc. Diferitele tipuri de interfețe limitează viteza cu care sunt transmise datele între unitatea de disc și sistem, având performanțe în funcționare diferite. Tipul interfeței determină în mare măsură tipul unităților de disc care pot fi utilizate și viteza finală a subsistemului format din unitatea de disc, interfață și controler.

5.9.1. Interfața IDE/ATA

Cea mai răspândită interfață pentru unitățile de discuri fixe este numită atât IDE, cât și ATA. Denumirea IDE (*Integrated Drive Electronics*) este mai des utilizată, și se referă la unitățile de discuri care au un controler integrat în unitate. Această denumire este improprie, deoarece actualmente toate unitățile de discuri au controlerul integrat în unitate. Denumirea ATA (*AT Attachment*) reprezintă standardul ANSI care definește interfața de conectare la calculatoarele AT. Interfața IDE/ATA este de 16 biți, ca și magistralele de date și de I/E ale primelor calculatoare IBM PC/AT. Această dimensiune de 16 biți se menține și la variantele mai performante ale interfeței ATA, de exemplu ATA-2.

Ansamblul format din unitate și controler este conectat la unul din conectorii de pe placa de bază sau la o placă adaptoare pentru magistrală. Prin combinarea unității și a controlerului rezultă mai multe avantaje: