

## Interconnection Networks

Asist. Felician ALECU

Catedra de Informatică Economică, A.S.E. București

*The design of interconnection system between processors and memory may be a bus oriented one using cache memories. In other computers, there are multiple memory modules connected to the processors using switches. There is always a possibility of memory contention problem, which may degrade the overall program performance.*

**Keywords:** *interconnection network, parallel computer, messages routing, contention.*

**P**entru conectarea procesoarelor și a blocurilor de memorie ce compun un sistem paralel se folosesc rețelele de interconectare. O rețea de interconectare reprezintă în esență un graf  $G = (V, M)$ , unde  $V$  este mulțimea vârfurilor (nodurilor) iar  $M$  mulțimea muchiilor (conexiunilor). O muchie între două noduri există numai dacă cele două noduri sunt conectate. Legăturile dintre nodurile grafului pot fi direcționate (orientate) sau nedirecționate (neorientate). În nodurile rețelei se pot afla procesoare sau blocuri de memorie. De asemenea pot exista și noduri complexe formate din procesoare și blocurile de memorie atașate. Topologia rețelei are o mare influență asupra performanțelor de ansamblu și asupra costurilor.

Rețelele de interconectare utilizate în cadrul calculatoarelor paralele pot fi:

- *rețele statice* – sunt formate din conexiuni fixe, punct la punct, între nodurile rețelei. Rețelele statice sunt în general utilizate în cazul calculatoarelor cu transfer de mesaje.

- *rețele dinamice* – sunt compuse din conexiuni variabile în timp. Nodurile sunt conectate între ele în mod dinamic cu ajutorul comutatoarelor. Rețele dinamice sunt utilizate cu preponderență pentru construirea calculatoarelor cu memorie partajată.

Pentru descrierea caracteristicilor rețelelor de interconectare se folosesc o serie de indicatori care vor fi detaliați în cele ce urmează:

- *dimensiunea rețelei* – reprezintă numărul de noduri care compun rețeaua de interconectare;

- *gradul unui nod* – semnifică numărul de muchii incidente în nodul respectiv. Dacă muchiile grafului sunt direcționate, gradul

unui nod reprezintă suma dintre gradul de intrare al acestuia (numărul muchiilor care sunt orientate astfel încât acestea intră în nodul ales) și gradul de ieșire (numărul de muchii care ies din nodul respectiv). Fiecare muchie incidentă atrage după sine necesitatea existenței unui port dedicat de intrare sau de ieșire. Pentru obținerea unor costuri de realizare cât mai reduse se impune ca gradul nodurilor să se mențină la valori cât mai mici;

- *diametrul rețelei* – se definește ca distanța maximă între două noduri oarecare ale rețelei unde prin distanța între două noduri înțelegem drumul cel mai scurt dintre acestea, măsurat în număr de muchii (conexiuni). Cu cât diametrul unei rețele este mai mic cu atât timpii de comunicație dintre noduri sunt mai mici ceea ce atrage după sine obținerea unor performanțe mai bune;

- *lățimea biseției* – semnifică numărul minim de conexiuni care, dacă sunt eliminate, conduc la obținerea a două subrețele egale (sau aproximativ egale);

- *costul rețelei* – este definit ca fiind numărul de conexiuni prezente în cadrul rețelei de interconectare.

Atunci când se proiectează o rețea de interconectare se dorește obținerea unor grade ale nodurilor cât mai mici, a unui diametru minim și a unei lățimi a biseției cât mai mari, toate acestea la un cost cât mai scăzut. Cu cât gradul nodurilor este mai mic cu atât este nevoie de mai puține porturi atașate procesoarelor ceea ce conduce către minimizarea costului de construcție a rețelei. Necesitatea minimizării diametrului rețelei apare din dorința ca nodurile să comunice pe distanțe cât mai mici. Un diametru mai mic al rețelei atrage

după sine obținerea unor performanțe mai bune. Lățimea bisecției trebuie să fie cât mai mare deoarece astfel se evită ștrangulările care pot afecta în sens negativ performanțele de ansamblu ale rețelei. Desigur că, din rațiuni economice, se dorește obținerea unor rețele de interconectare care să aibă un cost de construcție și de exploatare cât mai redus.

Topologia rețelei are o mare influență asupra performanțelor de ansamblu și asupra costurilor. În continuare vor fi prezentate cele mai frecvent întâlnite topologii ale rețelelor de interconectare statice:

- *rețeaua complet conectată* – fiecare procesor este direct legat cu toate celelalte noduri ale rețelei. Comunicația dintre oricare două noduri se face într-un singur pas datorită existenței unei conexiuni directe. Rețeaua complet conectată nu este o rețea blocantă deoarece conexiunea dintre oricare două noduri nu exclude o posibilă legătură între oricare alte două noduri. Deși poate părea a fi topologia ideală, datorită existenței unei legături între oricare două noduri, totuși, în practică, rețeaua complet conectată este utilizată destul de puțin datorită costului ridicat de construcție care provine din faptul că fiecare procesor trebuie să aibă porturi și canale de comunicație separate cu toate celelalte procesoare care fac parte din sistem.

- *rețeaua stea* – o rețea de tip stea cu  $n$  noduri este compusă dintr-un nod central la care sunt conectate toate celelalte  $n-1$  noduri. Comunicarea între oricare două noduri se face prin rutarea mesajelor prin nodul central. Acest lucru atrage după sine o importantă limitare a performanțelor acestui tip de rețea deoarece există posibilitatea ca nodul central, datorită traficului ridicat, să nu mai poată face față cererilor de transfer de mesaje. Astfel, nodul central poate reprezenta o strangulare în ceea ce privește comunicațiile din cadrul unei rețele de tip stea atunci când traficul este ridicat.

- *rețeaua liniară* – reprezintă cel mai simplu mod de a conecta un număr de  $n$  noduri. Fiecare nod este direct legat cu două noduri vecine, excepție făcând nodurile terminale. Comunicația între două noduri care nu sunt direct conectate se face prin înaintarea repe-

tată a mesajului prin nodurile intermediare dintre sursă și destinație. Direcția în care se transmite mesajul (stânga sau dreapta) se alege în funcție de modul în care sunt poziționate nodurile sursă și destinație în cadrul lanțului format.

- *rețeaua inel* – dacă nodurile extreme ale unei rețele liniare se leagă printr-o conexiune directă, atunci se obține o rețea de tip inel. Comunicația între două noduri care nu sunt direct conectate se face prin înaintarea repetată a mesajului prin nodurile intermediare dintre sursă și destinație. Direcția în care se transmite mesajul (stânga sau dreapta) se alege în funcție de drumul cel mai scurt dintre sursă și destinație.

- *rețeaua de tip grilă* – fiecare nod, cu excepția celor terminale, va fi conectat în mod direct cu patru noduri vecine, pe două direcții. Acest tip de rețea reprezintă o extensie la două dimensiuni a rețelei liniare. Transferul datelor se face prin înaintarea acestora mai întâi pe una dintre dimensiuni și apoi pe cealaltă până când se atinge nodul destinație. Dacă dimensiunile grilei sunt egale atunci grila poartă numele de grilă pătrată, altfel ea se numește grilă rectangulară.

- *rețeaua tor* – dacă se pleacă de la o rețea de tip grilă bidimensională și se conectează nodurile marginale pe fiecare linie și coloană, atunci se obține o rețea de tip tor. Acest tip de rețea reprezintă o extensie la două dimensiuni a rețelei de tip inel. Transferul datelor se face prin înaintarea acestora mai întâi pe una dintre dimensiuni și apoi pe cealaltă până când se atinge nodul destinație. Ca și în cazul grilelor, torurile pot fi pătrate sau rectangulare, respectiv bidimensionale sau multidimensionale.

- *rețeaua arbore* – fiecare nod, cu excepția celor terminale, este conecta la un număr de  $k$  noduri descendente. Aceste noduri poartă numele de noduri fii iar cele terminale se mai numesc noduri frunză. Nodurile terminale nu au descendenți. Singurul nod al arborelui care nu are ascendenți este numit nod rădăcină. Dacă  $k$  este egal cu 2 atunci arborele poartă numele de arbore binar: fiecare nod, cu excepția nodurilor frunză, are maxim doi descendenți. Un arbore binar în care fiecare nod

are exact doi descendenți se numește arbore binar complet. Transferul datelor se face prin înaintarea acestora prin arbore de la nodul sursă către cel destinație. Oricare ar fi nodurile sursă și destinație, între acestea va exista întotdeauna un singur drum în cadrul arborelui. Dacă se consideră subarborele care conține nodurile sursă și destinație în ramuri diferite, atunci toate mesajele vor trece prin nodul rădăcină al acestuia. Ținând cont de acest aspect putem concluziona că principalul dezavantaj al rețelelor de tip arbore este aglomerarea ce apare la nivelul nodurilor apropiate de rădăcină. Astfel, ori de câte ori va fi nevoie ca un procesor aflat în subarborele stâng să comunice cu unul aflat în subarborele drept, mesajul va trece prin nodul rădăcină. Pentru înlăturarea acestui dezavantaj a fost introdus conceptul de *arbore gras* care reprezintă un arbore binar modi-

ficat astfel încât cu cât ne apropiem de nodul rădăcină cu atât numărul conexiunilor este dublat. Dublarea conexiunilor se face prin duplicarea acestora sau prin dublarea capacității.

- *rețeaua hipercub* – o rețea hipercub cu  $N$  dimensiuni este constituită dintr-un număr de  $n = 2^N$  noduri etichetate de la 0 până la  $n - 1$ . Două noduri ale rețelei sunt conectate printr-o legătură directă numai dacă reprezentarea binară a etichetelor acestora diferă exact pe o singură poziție. Datorită acestui aspect putem afirma că fiecare nod al rețelei este direct conectat cu alte  $N$  noduri. Transferul datelor se face prin înaintarea acestora succesiv, pe fiecare din dimensiunile hipercubului, de la nodul sursă către cel destinație.

În continuare sunt prezentate comparativ caracteristicile rețelelor statice detaliate anterior (tabelul 1).

**Tabelul 1 – Caracteristicile rețelelor statice de interconectare**

| Tipul rețelei        | Dimensiunea rețelei | Gradul nodurilor | Diametrul rețelei                      | Lărgimea bisecției | Costul rețelei           |
|----------------------|---------------------|------------------|--|--------------------|--------------------------|
| Complet conectată    | $n - 1$             | $n - 1$          | 1                                      | $n^2 / 4$          | $n \cdot (n - 1) / 2$    |
| Stea                 | $1, n - 1$          | $1, n - 1$       | 2                                      | 1                  | $n - 1$                  |
| Liniară              | 1, 2                | 1, 2             | $n - 1$                                | 1                  | $n - 1$                  |
| Inel                 | 2                   | 2                | $\lfloor n/2 \rfloor$                  | 2                  | $n$                      |
| Grilă bidimensională | 2, 3, 4             | 2, 3, 4          | $2 \cdot (\sqrt{n} - 1)$               | $\sqrt{n}$         | $2 \cdot (n - \sqrt{n})$ |
| Tor bidimensional    | 4                   | 4                | $2 \cdot \lfloor \sqrt{n} / 2 \rfloor$ | $2 \cdot \sqrt{n}$ | $2 \cdot n$              |
| Arbore binar complet | 1, 2, 3             | 1, 2, 3          | $2 \cdot (\log(n + 1) - 1)$            | 1                  | $n - 1$                  |
| Hipercub             | $\log n$            | $\log n$         | $\log n$                               | $n / 2$            | $(n \cdot \log n) / 2$   |

Nici una dintre aceste topologii nu îndeplinește simultan toate criteriile de performanță enunțate anterior. Din acest motiv alegerea unui anumit tip de rețea se reduce de cele mai multe ori la realizarea unui compromis între costul sistemului și performanțele dorite.

Rețelele hibrid reprezintă o modalitate prin care se încearcă obținerea unor performanțe superioare prin combinarea mai multor topologii de interconectare: arbori grași, grile din arbori sau lanțuri, hipercuburi formate din grile, etc. Aceste rețele încearcă să elimine dezavantajele topologiilor individuale ce intră în componența lor în condițiile în care se conservă calitățile acestora.

O altă modalitate de îmbunătățire a performanțelor rețelelor de interconectare o reprezintă implementarea mai multor topologii între aceleași noduri. Fiecare astfel de topologie este dedicată unui anumit tip de operații.

Un alt aspect deosebit de important în ceea ce privește rețele statice îl reprezintă implantarea unei rețele în altă rețea de interconectare deoarece dezvoltarea unui algoritm presupune implantarea grafului programului în graful ce descrie sistemul pe care se va implementa. În plus, implantarea permite adaptarea la alte tipuri de rețele a algoritmilor de rutare dezvoltați pentru o anumită topologie de rețea. Implantarea unei rețele într-o altă rețea presupune să asociem fiecărui nod din

rețeaua sursă unul sau mai multe noduri din rețeaua destinație iar fiecărei muchii din rețeaua inițială una sau mai multe muchii din rețeaua în care se face implantarea.

Rețelele dinamice de interconectare sunt compuse din conexiuni variabile în timp realizate cu ajutorul comutatoarelor. Principalele topologii ale rețelelor dinamice de interconectare sunt:

- *rețeaua crossbar* – este utilizată pentru a conecta un număr de intrări ( $m$ ) cu o serie de ieșiri ( $n$ ) prin folosirea unei grile de  $m \times n$  comutatoare binare (cu două poziții: închis sau deschis) având o singură intrare și o singură ieșire. Fiecare comutator poate să faciliteze sau să întrerupă conexiunea dintre un modul de intrare și unul de ieșire. Rețeaua crossbar nu este o rețea blocantă deoarece o conexiune între două noduri nu afectează modul de conectare a oricărui altor două noduri. O rețea crossbar de dimensiune  $n \times n$  permite transferul simultan a  $n$  date. Dacă sunt admise numai conexiuni 1-la-1, atunci rețeaua devine o rețea de permutare deoarece se pot realiza  $n!$  conexiuni între elemente. Starea comutatoarelor la un anumit moment determină permutarea realizată, adică modul de conectare a elementelor de intrare cu cele de ieșire. Rețeaua crossbar este o rețea scalabilă din punct de vedere al performanțelor și nescalabilă din perspectiva costurilor pe care le implică.

- *rețeaua magistrală* – o magistrală este formată dintr-un număr de linii folosite pentru transferul datelor între procesoare, blocurile de memorie și dispozitivele de intrare/ieșire. La un moment dat este posibilă o singură legătură între un nod sursă și unul destinație. Dacă au loc simultan mai multe cereri pentru transfer de date atunci se va folosi un mecanism de arbitrar care va selecta elementul care va avea acces la magistrală. Performanțele rețelelor de tip magistrală sunt influențate de numărul de procesoare prezente în sistem. Creșterea numărului de procesoare atrage după sine scăderea performanțelor. Astfel, timpul de acces la memorie va fi din ce în ce mai mare deoarece magistrala este ocupată de alte procesoare. Performanțele

acestui tip de rețea sunt foarte bune pentru un număr mic de procesoare însă acestea scad drastic pe măsură ce numărul procesoarelor din sistem crește. Acest inconvenient poate fi remediat prin folosirea memoriilor cache locale. Astfel, datele solicitate de către procesor vor fi mai întâi căutate în memoria cache și numai în cazul în care nu au fost găsite se accesează magistrala pentru a aduce datele respective de la un bloc de memorie. În acest fel se reduce numărul de accese la memoria globală iar acest lucru are ca efect reducerea traficului prin magistrală și scăderea timpului de acces la memorie. Cu toate acestea folosirea memoriilor cache atrage după sine probleme legate de asigurarea coerenței conținutului acestora. Rețeaua magistrală este o rețea scalabilă din punct de vedere al costurilor și nescalabilă din perspectiva performanțelor.

- *rețeaua multinivel* – reprezintă o variantă intermediară între rețeaua de tip crossbar și cea de tip magistrală atât din perspectiva costurilor cât și din cea a performanțelor. Elementul care stă la baza acestui tip de rețea este comutatorul care stabilește în mod dinamic conexiuni între intrări și ieșiri. Cel mai simplu tip de comutator utilizat în practică este cel  $2 \times 2$  care realizează  $2!$  conexiuni posibile între intrări și ieșiri: conexiunea directă și cea inversă. În general, comutatoarele utilizate sunt de forma  $m \times n$ , unde  $m$  și  $n$  sunt puteri ale lui 2 egale ( $2 \times 2$ ,  $4 \times 4$ , etc.). O rețea multinivel este compusă dintr-un număr de  $k$  etaje formate din comutatoare de dimensiune  $m \times n$ . Nivelele sunt legate între ele prin intermediul unor conexiuni fixe care respectă a anumită topologie. Rețelele multinivel diferă între ele prin tipul comutatoarelor folosite și prin modalitatea în care sunt implementate conexiunile fixe existente între nivele. Aceste conexiuni realizează un anumit tip de permutare între ieșirile unui nivel și intrările următorului astfel încât să se poată realiza legătura între oricare intrare și ieșire prin poziționarea corespunzătoare a comutatoarelor.

Tabelul 2 prezintă comparativ caracteristicile rețelelor dinamice de interconectare detaliate anterior.

**Tabelul 2** – Caracteristicile rețelelor dinamice de interconectare

| Caracteristica                        | Rețeaua crossbar<br>$n \times n$ | Rețeaua magistrală<br>$n$ procesoare | Rețeaua multinivel<br>$n \times n$ |
|---------------------------------------|----------------------------------|--------------------------------------|------------------------------------|
| Timp de transfer                      | $\Theta(1)$                      | $\Theta(1)$                          | $\Theta(\log n)$                   |
| Număr de comutatoare                  | $\Theta(n^2)$                    | $\Theta(n)$                          | $\Theta(n \log n)$                 |
| Capacitate de transfer la fiecare pas | O singură conexiune              | O permutare dacă nu apare blocare    | O permutare                        |

Analizând informațiile din tabel reiese în mod clar faptul că rețeaua magistrală are costul asociat cel mai mic dar în condițiile în care lățimea de bandă scade drastic dacă numărul de procesoare crește semnificativ. La polul opus se află rețeaua crossbar care are cel mai mare cost asociat dar care prezintă și cea mai mare lățime de bandă.

Rețeaua multinivel se situează între rețeaua magistrală și cea crossbar atât din punct de vedere al costurilor cât și din cel al performanțelor. Modularitatea blocurilor constructive reprezintă un avantaj însă nu trebuie pierdut din vedere faptul că rețelele multinivel pot fi blocante pentru anumite tipuri de comutatoare și topologii de interconectare între niveluri.

### Bibliografie

- [Asm03] S. Asmussen, *Applied Probability and Queues*, Springer, 2003
- [Jos03] J. Joseph, C. Fellenstein, *Grid Computing*, Prentice Hall, 2003
- [Dod02] Gh. Dodescu, B. Oancea, M. Răceanu, *Procesare paralelă*, Editura Economica, București, 2002
- [Jor02] H. F. Jordan, H. E. Jordan, *Fundamentals of Parallel Computing*, Prentice Hall, 2002
- [Tan98] A. S. Tanenbaum, *Rețele de calculatoare*, Computer Press Agora, București, 1998
- [Tan96] A. S. Tanenbaum, *Computer Networks*, Prentice Hall, 1996
- [Sab95] G. W. Sabot, *High Performance Computing*, Addison-Wesley, 1995