LOCALITATEA DE INTERCONECTARE

Cristian Lupu
Centrul pentru Noi Arhitecturi Electronice al Academiei Române

Rezumat: În articolul de față, se definește și se abordează un principiu nou în construcția și evaluarea rețelelor directe: *localitatea de interconectare*. Deși principiul localității este folosit aproape la orice nivel al unui proiect de calculator, el nu este suficient definit și întrebuințat în proiectarea și evaluarea rețelelor de interconectare directe. În articol, se dau unele definiții ale localității în sens clasic: localitatea referințelor și localitatea de comunicație. Apoi se introduce localitatea de interconectare, care este definită mai întâi structural (topologic) și apoi funcțional. Ca atare, localitatea unei rețele de interconectare va fi definită prin două localități: o localitate structurală și o localitate funcțională. Localitățile structurală și funcțională se măsoară prin vecinătăți. După ce se definesc o serie de vecinătăți, se introduc rezerva Moore și rezerva de vecinătate și se dau câteva exemple de evaluare a localităților structurale și funcționale.

Cuvinte cheie: Localitate, localitatea referințelor, localitate de comunicație, distribuție de mesaje, localitate de interconectare, localitate structurală, localitate funcțională, vecinătăți, diametru, distanță medie funcțională, rezerva Moore, rezerva de vecinătate, hipercub, hipercub binar, hipertor, hipergrilă.

1. Introducere: Localitatea referințelor și localitatea de comunicație

Se cunoaște că proprietatea generală a programelor este *localitatea referințelor* (locality of reference)[6]. *Codul* ascultă de o regulă empirică: *un program execută 90% din timpul de execuție 10% din cod.* Datele au o localitate mai slabă. Cele trei componente ale localității referințelor sunt: *temporală*, *spațială* și *secvențială*. Astfel, în *localitatea temporală* (temporal locality) există tendința ca, în viitorul apropiat, apelul să fie făcut la resursele la care *s-a apelat în trecutul recent*. O astfel de localitate prezintă construcțiile software ca buclele, variabilele temporare sau stivele. În *localitatea spațială* (spatial locality) tendința este ca referința următoare să se facă la resursele din *vecinătatea fizică, spațială* a ultimei referințe. Exemplul imediat este cel al referințelor la același bloc sau pagină de memorie, pe care se bazează construcția memoriilor *cache*. *Localitatea secvențială* se referă la tendința ca referința următoare să se facă la resursa următoare secvențial resursei la care s-a făcut referirea precedentă. Localitatea temporală reflectă o anumită *inerție temporală* a referințelor, cea spațială o *inerție spațială* a lor, iar cea secvențială o anumită, să zicem, *inerție funcțională, algoritmică*. *Localitatea referințelor* este unul dintre cele mai importante criterii cantitative de proiectare a calculatoarelor după Hennessy și Patterson, creatorii primelor calculatoare cu set redus de instrucțiuni. Aceste calculatoare au o arhitectură foarte simplă de tip *load/store*, instrucțiuni cu format fix pe 32 de biți și funcționare eficientă în *pipeline*, o memorie ierarhică, edificată pe baza principiului localității referințelor.

Deși principiul localității se vede, implicit, aproape la orice nivel al proiectului unui calculator, el nu este suficient definit și utilizat în proiectarea și evaluarea rețelelor de interconectare directe. În structuri de interconectare, abordarea care se cunoaște se bazează pe principiul *localității de comunicație* (communication locality) [4, 14]. În acest mod de a defini localitatea este necesară, mai întâi, definirea unor *modele de lucru* re~~prez~~entative pentru rețeaua în studiu. Cu toate acestea, rețeaua se poate comporta diferit de la o arhitectură la alta sau de ~~la~~ o aplicație la alta. De exemplu, unele aplicații rulate pe multicalculatoare generează mesaje foarte lungi, în timp ce, aceleași aplicații rulate pe multiprocesoare cu memorie distribuită și *cache*-uri coerente generează mesaje foarte scurte.

Modelul de lucru este definit prin trei parametri de bază: *distribuția destinațiilor, rata de injecție* și *lungimea mesajului*. Distribuția destinațiilor indică destinația pentru următorul mesaj în fiecare nod. În studiul rețelelor de interconectare, cea mai frecvent utilizată distribuție este cea *uniformă*. În această distribuție, probabilitatea ca un nod i să-i trimită un mesaj nodului j este aceeași pentru toți i și j, $i \neq j$. Distribuția uniformă nu face nici-o presupunere despre ce tip de calcul se generează prin mesajele trimise în rețea.

Localitatea de comunicație se împarte în două: *spațială* și *temporală*. O *aplicație* manifestă localitate spațială când distanța medie între noduri este mai mică decât cea corespunzătoare distribuției uniforme. Ca rezultat, fiecare mesaj consumă mai puține resurse, reducând disputele la aceeași resursă. Pe de altă parte, o aplicație are o localitate temporală când manifestă o afinitate de comunicație pe o submulțime de noduri. Probabilitatea să trimită mesaje nodurilor recent utilizate ca destinație pentru alte mesaje este mai mare decât pentru alte noduri. Trebuie notat că nodurile, manifestând o afinitate de comunicație, nu trebuie să fie vecine în rețea.

Când traficul de rețea nu este uniform, ne așteptăm ca orice corespondență rezonabilă a unui calcul paralel să plaseze *task*-urile sale, care schimbă mesaje cu frecvență mare, în locații fizic apropiate. Pentru a modela localitatea spațială, se folosesc două distribuții simple: *sfera localității* și *distribuția probabilității descrescătoare*. În prima, un nod trimite mesaje nodurilor în interiorul unei sfere centrată pe nodul sursă cu o probabilitate de obicei înaltă, ϕ, și nodurilor din afara sferei cu probabilitatea $1-\phi$. Această distribuție modelează localitatea de comunicație a unor programe care rezolvă probleme structurate. Un exemplu este localitatea de comunicaţie a unor programe care rezolvă probleme structurate. Un exemplu este comunicaţia cu vecinii cei mai apropiaţi, tipică programelor care rezolvă ecuaţii diferenţiale parţiale iterative, cuplată cu comunicaţia globală pentru verificarea convergenţei.

În distribuţia probabilităţii descrescătoare, probabilitatea de trimitere a unui mesaj la un nod descreşte cu cât distanţa dintre sursă şi destinaţie creşte. În [14] Reed şi Grunwald au propus o funcţie de distribuţie $\Phi(d) = Decay(l, dmax) \times l^d$, $0 < l < 1$, unde d este distanţa dintre nodurile sursă şi destinaţie, $dmax$ este diametrul reţelei şi l este un parametru de localitate. $Decay(l, dmax)$ este o constantă de normalizare a probabilităţii Φ, aleasă astfel încât suma probabilităţilor să fie egală cu unu. Valorile mici ale parametrului de localitate l înseamnă un grad mare de localitate, în timp ce valorile mari înseamnă că mesajele pot parcurge distanţe mari. Cu cât l se apropie de unu, funcţia de distribuţie Φ se apropie de distribuţia uniformă. Cu cât l se apropie de zero, Φ se apropie de o comunicaţie cu vecinii cei mai apropiaţi.

Distribuţiile descrise mai sus au diferite grade de localităţi spaţiale, dar nu au localitate temporală. Recent, la evaluarea performanţelor reţelelor de interconectare, au fost utilizate diferite reguli de comunicaţie ale perechilor de noduri sursă şi destinaţie. Aceste reguli de comunicaţie iau în considerare permutările folosite în mod uzual în algoritmii numerici paraleli. În aceste reguli de comunicaţie, nodul-destinaţie pentru mesajele generate de un anumit nod este totdeauna acelaşi. De aceea, utilizarea legăturilor reţelei este neuniformă şi aceste distribuţii ajung la un maxim de localitate temporală.

În sfârşit, pentru a modela localitatea temporală, a fost propusă o distribuţie bazată pe un model de stivă "cel mai puţin recent utilizat". În acest model, fiecare nod are o stivă a sa, conţinând cele m noduri care i-au trimis cel mai recent mesaje. Pentru fiecare poziţie în stivă, există o probabilitate de a trimite un mesaj nodului în poziţia respectivă. Suma probabilităţilor pentru nodurile din stivă este mai mică decât 1. Un nod care nu este în stivă poate fi ales ca destinaţie pentru următoarea transmisie şi, în acest caz, după transmisia mesajului, nodul destinaţie va fi inclus în stivă, înlocuind cea mai puţin recent utilizată destinaţie.

În ultimul timp, deşi articolele care privesc definiţia explicită a localităţii sunt rare, se vede o mişcare spre o caracterizare mai sofisticată a localităţii şi spre metode mai precise de măsurare şi de vizualizare a ei, în special la proiectarea şi evaluarea ierarhiilor de memorie [3, 5]. O altă problemă a localităţii este şi plasarea resurselor: problema distanţei-t, problema adiacenţei-j şi problema plasării generalizate [2]. Preocupările privind localitatea fac parte dintr-o problemă mult mai vastă, *calculul poziţional* sau *calculul bazat pe localizare* (location-aware computing, location-based computing). În acest tip de calcul, sunt posibile aplicaţii în care entităţile care iau parte la calcul au capacitatea de a-şi sesiza poziţia şi de a-şi modifica parametrii, interfeţele şi funcţiile [1, 15].

2. Localitatea de interconectare. Vecinătăţi

Am văzut în paragraful introductiv că, în structurile de interconectare, definiţia localităţii se bazează mai mult pe comunicaţii şi pe calcul decât pe proprietăţile de localitate intrinseci ale topologiilor. Termenul "localitate de interconectare" a mai fost folosit de Hillis atunci când prezintă problemele de alocare a memoriei la Connection Machine [7]. El generalizează caracteristica de localitate fizică a memoriei, ascunsă programatorului în calculatoarele convenţionale von Neumann, sau caracteristica de localitate fizică bidimensională a tehnologiei de implementare a circuitelor integrate. Noi privim *localitatea de interconectare* mai întâi *structural*, topologic şi apoi *funcţional*. Ca atare, localitatea unei reţele de interconectare va fi definită prin două localităţi: o *localitate structurală* şi o *localitate funcţională*.

Evaluarea localităţilor structurale într-o reţea de interconectare începe, de obicei, prin calcule de distribuţii de noduri cu distanţa logică faţă de o *origine aleasă arbitrar*. Localităţile structurale se apreciază, se *măsoară*, prin *vecinătăţi*. Vecinătăţile se împart în *vecinătăţi de suprafaţă* (radiale) şi *vecinătăţi de volum* (sferice).

Definiţia 2. 1: *Vecinătatea de suprafaţă* este numărul de noduri ale unei reţele la o distanţă d faţă de o origine aleasă arbitrar:

$SN_d(O) = N_d(O)$
unde O este originea.

Definiţia 2. 2: *Vecinătatea de volum* este numărul de noduri ale unei reţele la o distanţă d faţă de o origine aleasă arbitrar O:

Prin vecinătăţi, localitatea structurală se poate măsura analitic. O altă măsură, sintetică, a localităţii structurale este *diametrul* reţelei: la acelaşi număr de noduri, cu cât diametrul este mai mic, cu atât localitatea structurală a reţelei este mai mare.

O problemă pe care o vom menţiona doar pe scurt aici este aceea că vecinătăţile şi diametrele depind de *poziţia originii*. La reţelele regulate omogene, ca hipercubul generalizat sau hipertorul generalizat, poziţia originii nu contează, pe când la reţele neregulate, ca hipergrila generalizată, şi/sau reţele neomogene ca hiperstructurile generalizate[12], poziţia originii contează. Vom analiza cum depind vecinătăţile, diametrele şi alte elemente de măsură topologice sau funcţionale în funcţie de poziţia originii într-un articol viitor.

Localitatea structurală este o informaţie *invariabilă*, dependentă de topologia reţelei. Un punct de vedere funcţional asupra localităţii de interconectare trebuie să ia în considerare, în afară de topologie, distribuţiile de rutare a mesajelor, $\Phi_O(d)$, unde O este originea şi d este distanţa faţă de origine. Ca şi localitatea structurală, localitatea funcţională se apreciază prin vecinătăţi, o *vecinătate funcţională de suprafaţă* şi o *vecinătate funcţională de volum*.

Definiţia 2. 3: *Vecinătatea funcţională de suprafaţă* la distanţa d faţa de o origine O într-o reţea este dată de relaţia:

$$FSN_d(O) = \Phi_O(d) \times N_d(O),$$

unde $\Phi_O(d)$ este distribuţia de rutare a mesajelor la distanţa d. $\Phi_O(d)$ şi $N_d(O)$ depind de originea aleasă.

Definiţia 2. 4: *Vecinătatea funcţională de volum* la distanţa d faţa de o origine O într-o reţea este dată de relaţia:

$$FVN_d(O) = \sum_{i=1}^{d} \Phi_O(i) \times N_i(O),$$

unde $\Phi_O(i)$ este distribuţia de rutare a mesajelor la distanţa i. $\Phi_O(i)$ şi $N_i(O)$ depind de originea aleasă.

Ca şi diametrul, există o măsură sintetică pentru localitatea de interconectare funcţională: *distanţa medie funcţională* sau numărul mediu de legături între noduri.

Definiţia 2. 5: *Distanţa medie funcţională* a unei reţele de interconectare cu originea O este dată de relaţia:

$$\overline{d_F}(O) = \sum_{d=1}^{D(O)} d \times \Phi_O(d),$$

unde d este distanţa, $D(O)$ este diametrul care depinde de poziţia originii şi $\Phi_O(d)$ este distribuţia de rutare a mesajelor la distanţa d şi care depinde de asemenea de poziţia originii.

Vecinătăţile de suprafaţă şi de volum, precum şi diametrul sau gradul, sunt mijloace de evaluare analitică şi, respectiv, sintetică, a capacităţii de intercomunicare a unei reţele directe măsurând *localitatea structurală* a reţelei. Prin vecinătăţile funcţionale şi, indirect, prin distanţa medie funcţională, se exprimă, analitic şi sintetic, ce parte din localitatea structurală a fost folosită în procesul de comunicaţie implementat pe reţea. Cu alte cuvinte, vecinătăţile funcţionale şi distanţele medii funcţionale exprimă *localitatea funcţională* a reţelei.

3. Evaluarea localităţii de interconectare structurale prin vecinătăţi şi rezerve Moore

Să luăm ca exemplu o comparaţie din punct de vedere al localităţii *structurale* a patru tipuri de reţele directe omogene: un hipercub, un hipercub binar, un hipertor şi o hipergrilă.

Ca să le putem compara, pentru acelaşi sau aproape acelaşi număr de noduri, să introducem *vecinătăţile normalizate*, care depind *distanţa normalizată* $\delta = d/D$, unde D este diametrul reţelei. Deci, o să avem $SN_\delta(O) = N_\delta(O)$, $VN_\delta(O) = \sum_{i=1}^{d/D} N_i(O)$, $FSN_\delta(O) = \Phi_O(\delta) \times N_\delta(O)$, $FVN_\delta(O) = \sum_{i=1}^{d/D} \Phi_O(i) \times N_i(O)$.

Hipercubul (HC), hipertorul (HT) şi hipergrila (HG) din exemplul nostru, sunt în trei dimensiuni, câte 10 noduri pe dimensiune, totalizând $N_{HC} = N_{HT} = N_{HG} = 1000$ noduri. Hipercubul binar (HCB) este construit în 10 dimensiuni totalizând $N_{HCB} = 1024$ noduri. În continuare, dăm în tabelul 1, vecinătăţile de suprafaţă normalizate şi vecinătăţile de volum normalizate pentru cele patru tipuri de reţele. Menţionăm că poziţia originii este numai 0 ... 0, în sistemul de numeraţie în bază multiplă, în care se reprezintă reţeaua.

Tabelul 1: Vecinătățile de suprafață normalizate SN_δ și vecinătățile de volum normalizate VN_δ

Structuri δ	HC SN_δ	HC VN_δ	HT SN_δ	HT VN_δ	HG SN_δ	HG VN_δ	HCB SN_δ	HCB VN_δ
0.1	27	27	18	24	10	19	10	10
0.2	27	27	38	62	28	83	45	55
0.3	27	27	99	227	55	219	120	175
0.4	243	270	128	355	69	351	210	385
0.5	243	270	144	643	75	574	252	637
0.6	243	270	128	771	63	779	210	847
0.7	729	999	66	936	45	879	120	967
0.8	729	999	38	974	21	964	45	1012
0.9	729	999	6	998	6	995	10	1022
1.0	729	999	1	999	1	999	1	1023

Întâi de toate să observăm în figura 1 că distribuțiile hipertorului, hipergrilei și hipercubului binar sunt diferite de cea a hipercubului. Primele trei sunt sub formă de clopot iar cea din urmă este crescătoare. Mai departe, distribuția hipergrilei este sub formă de clopot *asimetric*, cele ale hipertorului și hipercubului binar fiind *simetrice* (vezi și tabelul 1). Distribuția hipercubului ar începe și ea să semene cu un clopot dacă mărim foarte mult numărul dimensiunilor în care este construit. Vecinătățile de suprafață ale hipertorului, hipergrilei și hipercubului binar seamănă cu distribuțiile a doi arbori reuniți la jumătatea diametrului. Distribuția hipercubului este cea mai apropiată de cea a unui arbore.

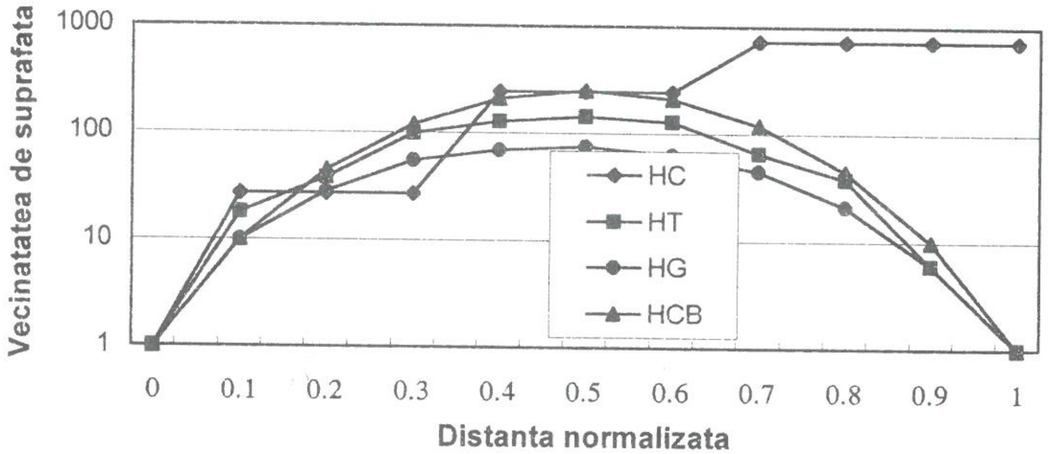


Figura 1. Vecinătățile de suprafață normalizate pentru HC, HT, HG și HCB

Pe de altă parte, diametrele rețelelor în discuție sunt pentru hipercub 3, pentru hipertor 15, pentru hipergrilă 27, iar pentru hipercubul binar 10. Localitatea structurală, la același număr de noduri, scade pe măsură ce diametrul crește. Ordinea rețelelor, în sensul micșorării localității structurale, este hipercubul, hipercubul binar, hipertorul și hipergrila. Este o localitate măsurată sintetic, la *exteriorul* rețelei. În *interiorul* rețelei, pot fi diverse localități (măsurate prin vecinătăți de suprafață normalizate): de la cea mai mică (1, d=0 pentru toate rețelele și d=D pentru HT, HG, HCB) până la cea mai mare (729, d=D pentru HC; 144, d=D/2 pentru HT; 75, d=D/2 pentru HG; 252, d=D/2 pentru HCB). Asta înseamnă că, de exemplu, la δ=0,3 HCB, HT și HG au o localitate de suprafață mai mare decât a HC, deși localitatea la exteriorul rețelei, *diametrală*, este inversă. Deci, deși localitatea la exteriorul rețelei este dată de diametru, în interiorul rețelei pot fi localități de suprafață care se abat de la ordinea fixată de diametru.

Să vedem ce se petrece și cu vecinătățile de volum normalizate, VN_δ, figura 2. Să observăm că vecinătatea de volum este crescătoare și, în mare, respectă ordinea dată de diametru. Dar, în curbele de vecinătate de volum pentru HT, HG și HCB există unele zone de inflexiune. Totodată, există și zone de intersecție. Curbele pleacă cu HG şi HT mai mari, apoi în zona de inflexiune se produce rearanjarea în ordinea dată de diametru: HCB, HT şi HG. HC "decolează" greu şi intersectează curbele HCB, HT şi HG undeva între δ=0,6 şi δ=0,7.

Ceea ce se reţine din exemplele pe care le-am dat este că diametrul măsoară localitatea structurală exterioară şi vecinătăţile pe cea interioară. Localitatea exterioară este diferită de cea interioară.

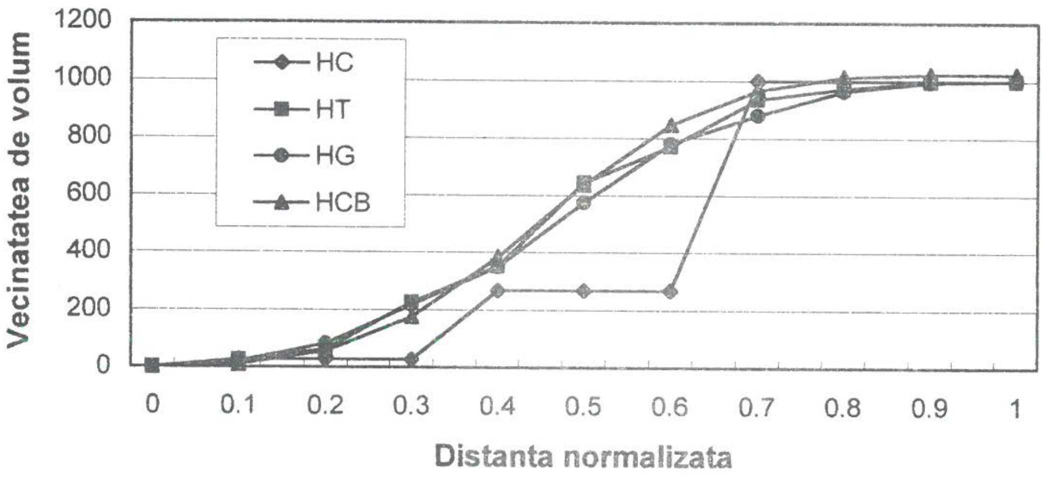


Figura 2. Vecinătăţile de volum normalizate pentru HC, HT, HG şi HCB

Pentru măsurarea absolută a localităţii structurale propunem *rezerva Moore* [9, 10] bazată pe limita Moore [13]. Limita Moore desemnează numărul maxim de noduri al unei reţele pentru care sunt date grad l şi diametrul D. Această limită se deduce pe baza ideii că, fiind dat gradul maxim l, un nod are cel mult l vecini la distanţa 1, $l(l-1)$ la distanţa 2, ..., $l(l-1)^{D-1}$ la distanţa D:

$$N_{Moore} = = 1 + l\frac{(l-1)^D - 1}{l-2}.$$

Limita îndeplinită pentru un l-arbore cu diametrul D, este o *limită absolută* a localităţii de interconectare structurale. În general, o reţea cu N noduri, grad l şi diametru D va satisface relaţia: $N \leq N_{Moore}$. Diferenţa este, în principiu, *rezerva Moore*. Cu cât ne îndepărtăm de această limită (rezerva Moore este mai mare), cu atât localitatea structurală este mai scăzută. Vom da definiţiile pentru rezerva Moore care este de trei feluri:

Definiţia 3. 1: *Rezerva Moore de suprafaţă* a unei reţele la o distanţă d de origine, SMR_d, este diferenţa între numărul de noduri al l-arborelui Moore cu gradul corespunzător reţelei la distanţa d, $l(l-1)^{d-1}$, şi vecinătatea de suprafaţă în reţeaua considerată:

$$SMR_d = l(l-1)^{d-1} - N_d.$$

Definiţia 3. 2: *Rezerva Moore* a unei reţele la distanţa d de origine, MR_d, este diferenţa între limita Moore la distanţa d în l-arborele Moore corespunzător reţelei, $N_{Moore}(d) = l\frac{(l-1)^d - 1}{l-2}$, şi vecinătatea de volum la aceeaşi distanţă d:

$$MR_d = l\frac{(l-1)^d - 1}{l-2} - \sum_{i=1}^{d} N_i.$$

Definiţia 3. 3: *Rezerva Moore diametrală* este $MR_D = N_{Moore}(D) - \sum_{i=1}^{D} N_i$.

Rezervele Moore se calculează exact pentru reţele *regulate*, cu toate nodurile având acelaşi grad. Pentru reţele iregulate, având grad variabil, putem să aproximăm luând gradul maxim.

În figura 2, se dau rezervele Moore normalizate pentru cinci rețele cu aproximativ același număr de noduri: un hipercub generalizat, GHC1, trei hipercuburi, HC1÷HC3 (HC1 este un hipercub binar) și două hipertoruri, HT1 și HT2. După cum ne așteptam, rețelele manifestă o localitate structurală în creștere cu gradul, adică cu produsele $\sum_{i=1}^{r}(m_i - 1)$, $r(m-1)$ și, respectiv, $2r$. Ordinea în care rezerva Moore normalizată scade și, ca atare, crește localitatea structurală, este HT2, HC1 (HT1), HC2, HC3 și GHC1. Creșterea gradului este 6, 10, 15, 27 și 68.

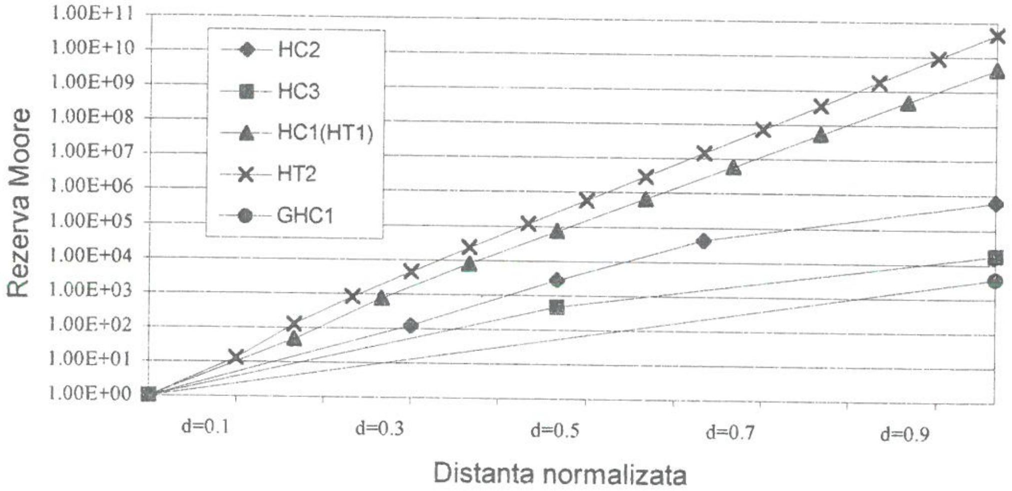


Figura 3. Rezerva Moore normalizată pentru un hipercub generalizat GHC1=20x50, trei hipercuburi HC1=2^{10}, HC2=4^5, HC3=10^3, și două hipertoruri HT1=4^5, HT2=10^3

Gradul, ca și diametrul, este o măsură sintetică. În timp ce gradul este o măsură mai abstractă, rezerva Moore este o măsură analitică, mai intuitivă decât gradul. Ea ne ajută să evaluăm localitatea structurală, în special la începutul curbelor, la distanțe mici, unde nu contează chiar așa de mult limita Moore. De exemplu, la rețelele cu grade 6 (HT2) și 10 (HC1, HT1) avem rezerva Moore la distanța $\delta=0.2$ de 124 de noduri și, respectiv, 45 de noduri. La distanța $\delta=0.4$ rezerva Moore crește la HT2 la 23081 noduri în timp pentru HC1(HT1) crește doar la 7815 noduri [11].

4. Evaluarea localității de interconectare funcționale prin rezerva de vecinătate

Între vecinătatea de volum și vecinătatea funcțională de volum este evident că există, pentru orice rețea directă, relația: $VN_d \geq FVN_d$. Diferența între cele două vecinătăți reprezintă ceea am denumit *rezervă de vecinătate* [8].

Definiția 4. 1: *Rezerva de vecinătate* pentru o rețea directă la distanța d, NR_d, este dată de:

NR_d [în noduri] = $VN_d - FVN_d$,
sau, în procente, de relația:
NR_d [%] = 1 – (FVN_d / VN_d)

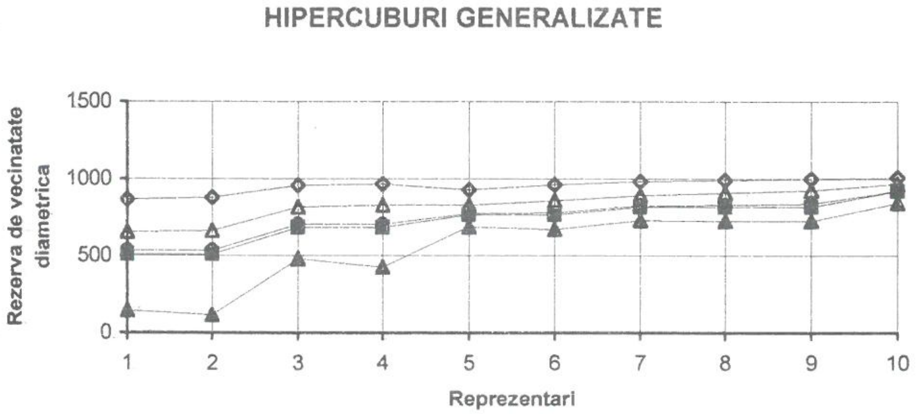
Prin vecinătăți de volum se evaluează capacitatea de intercomunicare a unei rețele directe, măsurând localitatea structurală a rețelei. Prin vecinătăți funcționale se exprimă localitatea funcțională. Utilizând rezerva de vecinătate putem să introducem un criteriu de proiectare sau de evaluare a unei topologii enunțând următoarea conjectură:

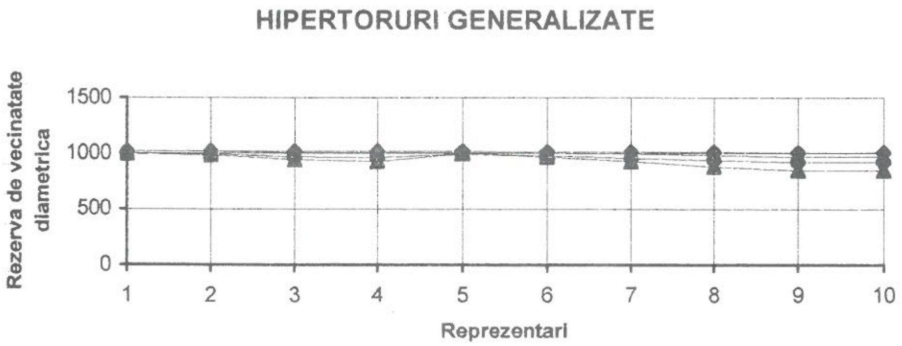
Conjectura 4. 1: Potențialul de intercomunicare structural al unei rețele directe este optim utilizat într-un proces de comunicație caracterizat printr-o distribuție de rutare Φ dacă rezerva de vecinătate este optimă (minimă pentru *NR* exprimat în noduri și maximă pentru *NR* exprimat în procente).

Acest criteriu poate fi utilizat pentru a căuta cele mai potrivite structuri la o distribuție de rutare dată sau, invers, pentru a sugera cea mai potrivită distribuție pentru o topologie dată. De asemenea, criteriul poate fi utilizat pentru a proiecta topologii noi, în acord cu distribuții de rutare specifice.

Pentru a ilustra cele spuse mai sus, să considerăm trei tipuri de distribuţii de rutare: *structurală*, $\Phi(d)=N_d/(N-1)$; *uniformă*, $\Phi(d)=\varphi=1/D$; *exponenţială*, $\Phi(d)=K\cdot\lambda^d$ [8, 14]. Pentru distribuţia exponenţială se aleg trei valori ale parametrului de localitate λ: 0.9, 0.5 şi 0.1.

În figura 4, se dau rezervele de vecinătate diametrale NR_D pentru trei tipuri de structuri. Structurile sunt hipercuburi generalizate(GHC), figura 4a, hipertoruri generalizate(GHT), figura 4b, şi hipergrile generalizate(GHG), figura 4c [12]. Structurile au acelaşi număr de noduri (N=1024), dar 10 reprezentări în sistemul de numeraţie în bază multiplă: (1)N=16x64, (2)N=32x32, (3)N=4x16x16, (4)N=8x8x16, (5)N=2x2x4x64, (6)N=2x4x4x32, (7)N=2x2x2x8x16, (8)N=2x2x4x8x8, (9)$N=4^5$, (10)$N=2^{10}$. NR_D oferă posibilitatea unei evaluări globale asupra reţelelor şi asupra distribuţiilor. Putem vedea că hipercuburile generalizate sunt cele mai sensibile structuri care sunt foarte dependente şi de distribuţia de rutare şi numărul de dimensiuni. Cele mai insensibile sunt hipergrilele. O altă observaţie este aceea că potenţialul de comunicare, măsurat de NR_D, este descrescător cu numărul de dimensiuni pentru GHC şi crescător pentru GHT şi GHG. În aceste condiţii, un BHC este cel mai prost GHC şi cel mai bun GHT sau GHG! Observăm, de asemenea, efectul parametrului de localitate λ care este acelaşi pentru toate structurile: cu cât este mai locală comunicaţia (λ mic), cu atât este mai proastă utilizarea reţelei (NR_D mai mare). Acest *efect* λ este mai vizibil pe structuri cu mai multe dimensiuni şi pe GHC, decât pe GHT sau GHG. În general, exceptând hipercuburile generalizate, putem aprecia că aceste structuri au caracteristici de localitate intrinseci slabe şi, ca atare, un potenţial mai prost de comunicare.





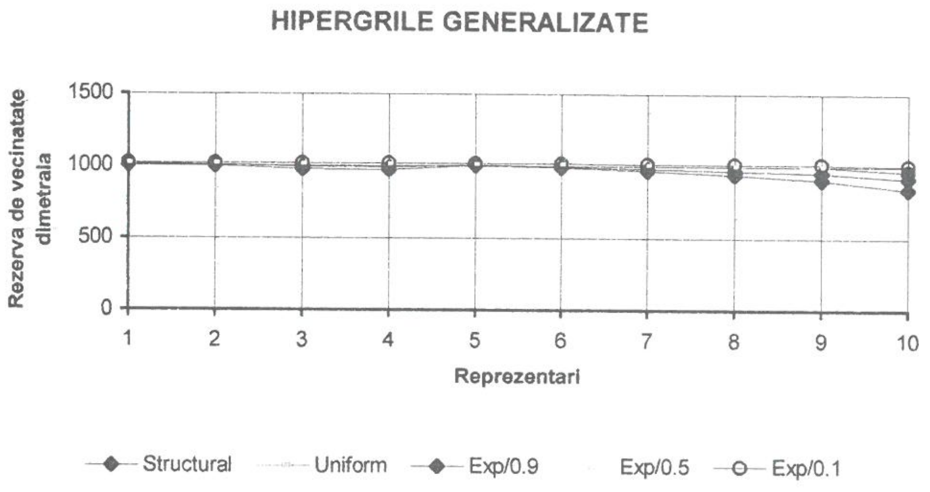


Figura 4. Rezerva de vecinătate diametrală NR_D pentru zece GHC, zece GHT şi zece GHG

Concluzii

În articolul de faţă, se dă o definiţie unitară pe baza vecinătăţilor *localităţii de interconectare*. Localitatea unei reţele de interconectare va fi definită prin două localităţi: o *localitate structurală* şi o *localitate funcţională*. Vecinătăţile de suprafaţă şi de volum, precum şi diametrul sau gradul, sunt mijloace de evaluare analitică şi, respectiv, sintetică, a capacităţii de intercomunicare a unei reţele directe măsurând *localitatea structurală* a reţelei. Prin vecinătăţile funcţionale şi, indirect, prin distanţa medie funcţională se exprimă, analitic şi sintetic, ce parte din localitatea structurală a fost folosită în procesul de comunicaţie implementat pe reţea. Cu alte cuvinte, vecinătăţile funcţionale şi distanţele medii funcţionale exprimă *localitatea funcţională* a reţelei. Localitatea structurală este o informaţie *invariabilă*, dependentă de topologia reţelei. Localitatea funcţională este o informaţie variabilă *dependentă* de topologia reţelei şi de distribuţia de rutare implementată pe reţea. Prin vecinătăţi structurale şi funcţionale se asigură o definire a localităţii de interconectare, bazată mai mult pe proprietăţile intrinseci ale topologiilor decât, indirect, pe comunicaţii şi pe calcul.

Vecinătăţile de suprafaţă şi de volum sunt mijloace de evaluare analitică a capacităţii de intercomunicare a unei reţele directe, măsurând localitatea structurală a reţelei. Diametrul măsoară localitatea structurală exterioară şi vecinătăţile pe cea interioară. Localitatea exterioară este diferită, în general, de cea interioară. Prin vecinătăţile funcţionale se exprimă, tot analitic, ce parte din localitatea structurală a fost folosită în procesul de comunicaţie implementat pe reţea. Cu alte cuvinte, vecinătăţile funcţionale exprimă localitatea funcţională a reţelei. La sfârşitul capitolului despre evaluarea structurală am introdus rezerva Moore ca o măsură absolută a localităţii de interconectare structurale. Rezerva Moore este necesară în proiectarea/evaluarea topologică, bazată pe localitate a reţelelor de interconectare directe, ca o alternativă la grad, mai intuitivă şi mai precisă.

Cu ajutorul rezervei de vecinătate, care este diferenţa între vecinătatea de volum şi vecinătatea funcţională, am introdus un criteriu de proiectare/evaluare al reţelelor directe. Prin utilizarea acestui criteriu, putem să apreciem cât de mult utilizăm o structură de interconectare având un proces de comunicaţie implementat pe ea şi invers. Criteriul poate fi utilizat pentru a căuta cele mai potrivite structuri la o distribuţie de rutare dată sau, invers, pentru a sugera cea mai potrivită distribuţie pentru o topologie dată. De asemenea, criteriul poate fi utilizat pentru a proiecta topologii noi, în acord cu distribuţii de rutare specifice. Utilizând rezerva de vecinătate diametrală am comparat trei reţele cunoscute: hipercubul generalizat, hipertorul generalizat şi hipergrila generalizată. Reţeaua cea mai dependentă structural de distribuţia de rutare este hipercubul generalizat, iar efectul λ la o distribuţie de rutare exponenţială este mai vizibil pe structuri cu mai multe dimensiuni. Exceptând hipercubul generalizat, celelalte au caracteristici intrinseci de localitate mai proaste şi un potenţial de comunicare mai slab.

Vecinătăţile, rezerva Moore, rezerva de vecinătate şi criteriul rezervei de vecinătate sunt încercări de a măsura mai precis proprietăţile de localitate ale reţelelor de interconectare directe, fiind rezultate incipiente în încercarea de a utiliza localitatea ca principiu cantitativ principal în proiectarea topologică.

Bibliografie

1. **ADDLESEE, M., R. CURWEN, J. HODGES, J. NEWMAN, P. STEGGLES, A. HOPPER**: Implementing a Sentient Computing System. În: Computer, vol. 34, no. 8, August 2001, pp. 50-56.
2. **BAE, M. M., B. BOSE**: Resource Placement in Torus-Based Networks. În: IEEE TC, vol. 46, no. 10, October 1997, pp. 1083-1092.
3. **DANDAMUDI, S. P., D. L. EAGER**: Hot-Spot Contention in Binary Hypercube Networks. În: IEEE TC, vol. 41, no. 2, February 1992, pp. 239-244.
4. **DUATO, J., S. YALAMANCHILI, L. NI**: Interconnection Networks. An Engineering Approach. În: IEEE Computer Society Press, Los Alamitos, California, 1997.
5. **GRIMSRUD, K., J. ARCHIBALD, R. FROST, B. NELSON**: Locality as a Visualization Tool. În: IEEE TC, vol. 45, no. 11, November 1996, pp. 1319-1326.
6. **HENNESSY J., D. A. PATTERSON**: Computer Architecture. A Quantitative Approach, Morgan Kaufmann Pub. Inc, San Mateo, California, 1990.
7. **HILLIS, W. D.**:The Connection Machine, *The* MIT Press, Cambridge, Massachusets, 1985.
8. **LUPU, C., A. NICOLESCU**: Neighbourhood Reserve -- A Measure Of Locality of the Direct Interconnection Networks. În: Proc. of the 9-th Mediterranean Electrotechnical Conferance–MELECON'98, vol. II, May 1998, Israel, Tel Aviv, pp. 1380-1384.
9. **LUPU, C.**: Moore Reserve – An Absolute Measure of Interconnection Structural Locality. În: Proc. of the 12-th International Conference on Control Systems and Computer Science, CSCS12, Vol. II, May 1999, pp. 221-225, Bucharest, Romania.
10. **LUPU, C., A. HAGIESCU**: Moore reserve for orthogonal networks. A case study. În: Proc. of the 10th Mediterranean Electrotechnical Conferance–MELECON'2000, vol. I, May 2000, pp. 97-100, Cyprus.
11. **LUPU, C., A. HAGIESCU**: Programe de evaluare topologică a rețelelor directe ortogonale omogene şi neomogene ", *Raport CNEA*, 2000.
12. **LUPU, C.**: Rețele ortogonale. În: Revista Română de Informatică şi Automatică, vol. 11, nr. 2, 2001, pp. 28-36.
13. **PRIOR, D. M. N., M. G. NORMAN, N. J. RADCLIFFE, L. J. CLARKE**: What price regularity. În: Concurrenncy: Practice and Experience, vol. 2, no. 1, March 1990, pp. 55-78.
14. **REED, D. A., D. C. GRUNWALD**: The Performance of Multicomputer Interconnection Network. În: Computer, vol. 20, no. 6, June 1987, pp. 63-73.
15. **WANT, R., B. SCHILIT**: Expanding the Horizons of Location-Aware Computing. În: Computer, vol. 34, no. 8, August 2001, pp. 31-34.