

## Homogeneous Sets of Servers Configuration

Ion BOLUN, Academia de Studii Economice, Chișinău, [bolun@ase.md](mailto:bolun@ase.md)  
 Anatol CIUMAC, Accent Electronic S.A., Chișinău, [aciumac@accent.md](mailto:aciumac@accent.md)  
 Andrei SVERINIUC, Academia de Studii Economice, Chișinău,  
[asveriniuc@mobiasbanca.md](mailto:asveriniuc@mobiasbanca.md)

*Results on the configuration of homogeneous set of servers for computer networks, aiming the minimization of its total cost are presented. The significant dependence of optimal solution with the number of server types (from which is done the selection) and, also, with the structure of total flow of users' requests is reveal.*

**Keywords:** computer network, server set, configuration, flow of requests, response time, optimal solution.

### Introducere

Edificarea societății informaționale necesită stocarea și procesarea unor volume mari de date. În acest scop se creează centre informatice interconectate în rețele. Informațiile de uz comun se procesează de către serverele rețelilor informatice. Creșterea volumului de date cere sporirea performanțelor seturilor de servere și, respectiv, a cheltuielilor cu crearea lor. De aceea este importantă optimizarea configurării seturilor de servere, urmărind minimizarea cheltuielilor necesare la satisfacerea cerințelor de procesare a fluxurilor de cereri ale utilizatorilor.

Diverse aspecte teoretice ale configurării seturilor de servere ale rețelelor de calculatoare sunt abordate în [1]. Tot aici este formulată problema de configurare a seturilor omogene de servere și propus un algoritm de soluționare a acesteia. În această lucrare sunt sistematizate unele rezultate ale calculelor în domeniu, obținute folosind aplicația informatică ConfigServ elaborată în acest scop.

### Concretizarea problemei de cercetare

Notății folosite:

$m$  – numărul de tipuri de servere, din care se selectează setul optim de servere;

$C_j, U_j$  – costul și, respectiv, productivitatea unui server de tip  $j = \overline{1, m}$ ;

$n$  – numărul de categorii de cereri ale utilizatorilor rețelei;

$u_i$  – laboriozitatea medie de deservire a unei cereri către serviciul  $i = \overline{1, n}$  (cereri de categoria  $i$ );

$u_i^{(2)}$  – momentul de ordinul doi al laboriozității de deservire a unei cereri de categoria  $i$ ;  
 $T_{di}$  – limita de sus pentru durata medie de răspuns la deservirea unei cereri de categoria  $i$ ;  
 $\Lambda_i$  – intensitatea sumară a fluxului de cereri de categoria  $i$  către setul de servere în cauză.

În linii mari, problema constă în următoarele. Fie sunt cunoscute valorile mărimilor:  $m; n; C_j, U_j, j = \overline{1, m}; u_i, u_i^{(2)}, T_{di}, \Lambda_i, i = \overline{1, n}$ . Se cere determinarea tipului  $j^*$  și a numărului optim  $M_{j^*}^*$  de servere de acest tip ce ar asigura costul minim  $C^*$  al setului de servere la deservirea de către acestea a tuturor cererilor fluxurilor  $\Lambda_i, i = \overline{1, n}$  și, totodată, nedepășirea duratelor  $T_{di}, i = \overline{1, n}$ , unde  $C^* = M_{j^*}^* \cdot C_{j^*}$ . (1)

Este evident că la creșterea numărului  $m$  de tipuri de servere se asigură posibilități mai largi de alegere a setului de servere și, eventual, de obținere a unei soluții mai bune. Dar pe cât de semnificativă ar putea fi îmbunătățirea soluției, este relativ greu de apreciat fără calcule speciale.

Pentru probleme reale, valorile mărimilor  $C_j, U_j, j = \overline{1, m}$  se specifică în funcție de caz în baza ofertei de servere pe piață ș.a. În scopuri de cercetare, pentru determinarea anumitor dependențe comune, este oportună generarea unor șiruri de valori pentru perechile  $\{C_j, U_j\}, j = \overline{1, m}$ , folosind următoarele repere:

$$U_{j+1} = hU_j, j = \overline{1, m-1}; \quad \frac{C_{j+1}}{C_j} = \left( \frac{U_{j+1}}{U_j} \right)^{\frac{p-1}{p}}. \quad (2), (3)$$

Relația (2) stabilește ca constantă valoarea raportului între productivitatea serverelor vecine în șirul  $j = \overline{1, m}$ ; fără a reduce dimensiunea problemei, se consideră că  $h > 1$ . În ce privește relația (3), în lucrare se consideră doar cazurile când  $p > 1$ , deoarece:

1) la  $p = 1$ , costul serverelor  $C_j$  nu ar depinde de productivitatea lor  $U_j$ , ceea ce în realitate nu are loc;

2) la  $0 < p < 1$ , odată cu creșterea productivității serverelor  $U_j$  ar descrește costul lor  $C_j$ , ceea ce în probleme reale nu ar trebui să fie. La întâlnirea unor asemenea cazuri, deja la

$$h = \sqrt[m-1]{\frac{U_m}{U_1}};$$

În cazul unor fluxuri multidimensionale de cereri ale utilizatorilor ( $n > 1$ ), structura fluxului sumar, determinată de ponderea laboriozității fluxurilor de cereri de diferite categorii, poate influența uneori soluția. Pentru a cerceta o asemenea dependență, este oportun

$$U_{\min} = \sum_{i=1}^n \Lambda_i u_i;$$

**Rezultate și recomandări practice**

Algoritmii de configurare a seturilor de servere propuși în [1] sunt realizați de aplicația ConfigServ elaborată în C++ Builder. Modulul executabil al aplicației este de circa 700 Ko.

Folosind aplicația ConfigServ, sunt efectuate un șir de calcule în scopul determinării dependenței costului optim  $C^*$  de numărul  $m$  de tipuri de servere dintre care se efectuează alegerea și, de asemenea, de structura fluxului sumar  $\Lambda$  de cereri ale utilizatorilor. Fluxul  $\Lambda$  se consideră constituit din cereri de două categorii ( $n = 2$ ): dialog ( $i = 1$ ) și cere-re-răspuns ( $i = 2$ ). Caracteristicile cererilor, la repartitia generală a duratei deservirii lor, sunt prezentate în tabelul 1.

În ce privește laboriozitatea sumară  $U_{\min}$  a fluxurilor de cereri ale utilizatorilor, sunt folosite, în funcție de caz, valorile: 2 Gflops;

etapa preliminară de formare a datelor inițiale, din mulțimea de tipuri de servere luate în considerație se exclud serverele cu caracteristici inoportune – cost mai înalt la productivitate egală sau mai joasă;

3) la  $p < 0$ , costul serverelor  $C_j$  ar crește mai rapid decât productivitatea lor  $U_j$  – caz care corespunde relației (11) din [1] și nu se cercetează în această lucrare.

De menționat că condiția (3) la  $p=2$  se reduce la legea lui Grosh [2]. Din relațiile (2) și (3) obținem:

$$p = \frac{\lg \frac{U_m}{U_1}}{\lg \frac{U_m C_1}{U_1 C_m}}. \quad (4), (5)$$

de fixat valoarea laboriozității  $U_{\min}$  a fluxului sumar  $\Lambda$  de cereri ale utilizatorilor și de modificat ponderea laboriozităților  $U_{\min i}, i = \overline{1, n}$  ale fluxurilor  $\Lambda_i, i = \overline{1, n}$  de cereri, unde:

$$U_{\min i} = \Lambda_i u_i, i = \overline{1, n}. \quad (6), (7)$$

20 Gflops; 50 Gflops, 250 Gflops și 1250 Gflops.

**Tabelul 1.** Caracteristicile cererilor

Caracteristicile cererilor \ $i$	1	2
$T_{di}, s$	0,3	60
$u_i, Gflops$	0,2	20
$u_i^{(2)}, Gflops^2$	0,4	300

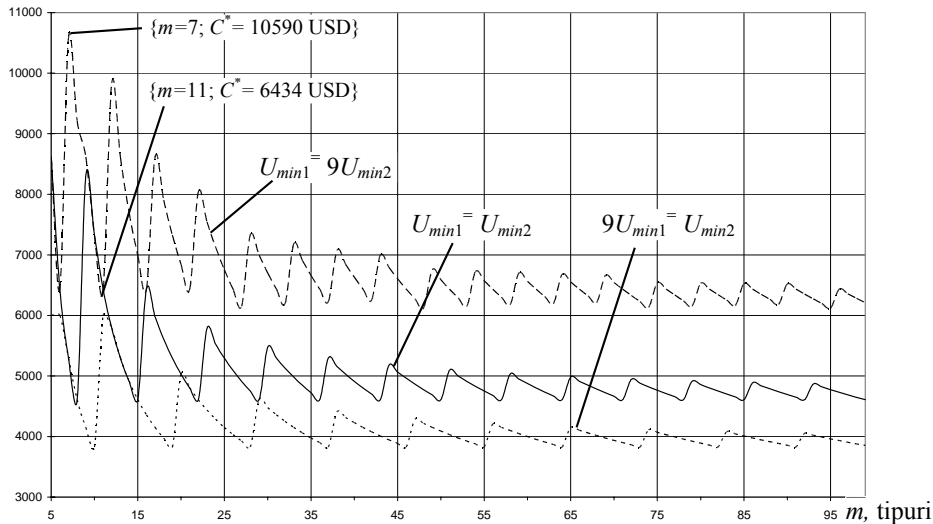
Ponderea laboriozității fiecăruia din fluxurile de cereri de categoriile 1 și 2 în laboriozitatea sumară  $U_{\min}$  este reprezentată în calcule de trei variante și anume:

- 1)  $U_{\min 1} = 9U_{\min 2}$ ;
- 2)  $U_{\min 1} = U_{\min 2}$ ;
- 3)  $U_{\min 2} = 9U_{\min 1}$ .

Valoarea  $p = 6,481$  folosită în calcule este determinată în baza informațiilor referitoare la calculatoare-servere concrete și anume:

1)  $U_1 = 2$  Gflops,  $C_1 = 2000$  USD – valori obținute ca media aritmetică a caracteristicilor respective pentru serverele IBM de tip xSeries 306 și serverele HP de tip ProLiant DL 320 G2;  
 2)  $U_m = 360$  Tflops,  $C_m = 100$  mln USD – valori caracteristice supercalculatorului Blue Gene/L Beta al companiei IBM.

$C^*$ , USD



**Fig. 1.** Dependența costului  $C^*$  al setului optim de servere de numărul  $m = \{5, 6, \dots, 99\}$  tipuri de servere și structura fluxului sumari de cereri ale utilizatorilor la  $U_{\min} = 2$  Gflops.

Din figurile 1-4 se poate observa o dependență puternică a costului  $C^*$  al setului optim de servere de numărul  $m$  de tipuri de servere din

$$1) \Delta C(7;11) = C^*|_{m=7} - C^*|_{m=11} = 4156 \text{ USD, iar } \delta C = \frac{\Delta C(7;11)}{C^*|_{m=7}} \cdot 100\% \approx 39\%;$$

$$2) \Delta C(7;999) = C^*|_{m=7} - C^*|_{m=999} = 4316 \text{ USD, iar } \delta C = \frac{\Delta C(7;999)}{C^*|_{m=7}} \cdot 100\% \approx 41\%.$$

În multe cazuri alegerea din  $m = 7$  tipuri de servere se consideră acceptabilă. Dar, după cum arată calculele efectuate, se poate întâmpla ca soluția să difere, în funcție de valoarea mărimii  $m > 7$ , cu peste 39 %. Și mai mare ar putea fi deosebirea, dacă s-ar lua în considerație o soluție intuitivă și nu cea optimă.

Totodată, comparând datele figurilor 1, 2 și 3 se poate observa:

a) la  $U_{\min} = 2$  Gflops structura fluxului sumari de cereri ale utilizatorilor influențează considerabil soluția – la creșterea ponderii laboriozității cererilor operative ( $i = 1$ ) crește și productivitatea, respectiv costul, setului

Unele din rezultatele obținute sunt prezentate în figurile 1-4, iar la  $U_{\min} = 2$  Gflops și  $m = 999$  valorile optime pentru  $C^*$  sunt:

- 1)  $C^* = 6118$  USD la  $U_{\min 1} = 9U_{\min 2}$ ;
- 2)  $C^* = 4592$  USD la  $U_{\min 1} = U_{\min 2}$ ;
- 3)  $C^* = 3830$  USD la  $U_{\min 2} = 9U_{\min 1}$ .

care se efectuează alegerea. Astfel, la  $U_{\min} = 2$  Gflops (figura 1) și  $U_{\min 1} = 9U_{\min 2}$  avem:

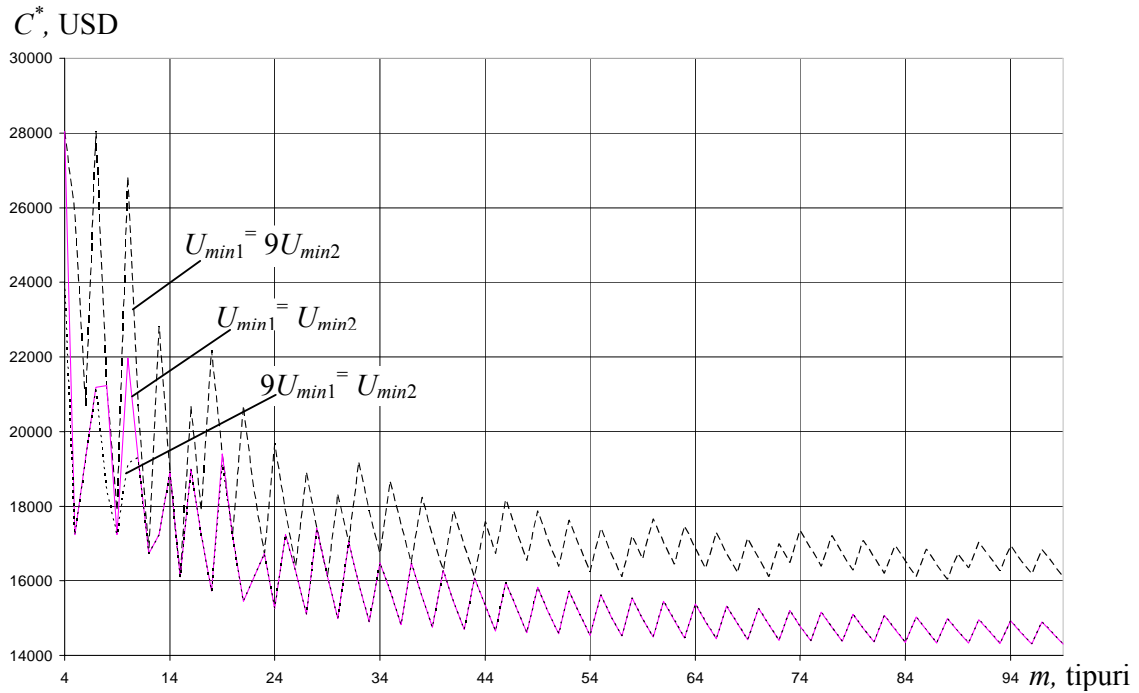
optim de servere necesar pentru deservirea fluxului sumari (vezi fig. 1);

b) la  $U_{\min} = 10$  Gflops influența, menționată în alineatul precedent, devine mai slabă decât la  $U_{\min} = 2$  Gflops, de exemplu pentru cazurile: 2)  $U_{\min 1} = U_{\min 2}$  și 3)  $U_{\min 2} = 9U_{\min 1}$  majoritatea soluțiilor la aceeași valoare a lui  $m$  coincid (vezi fig. 2);

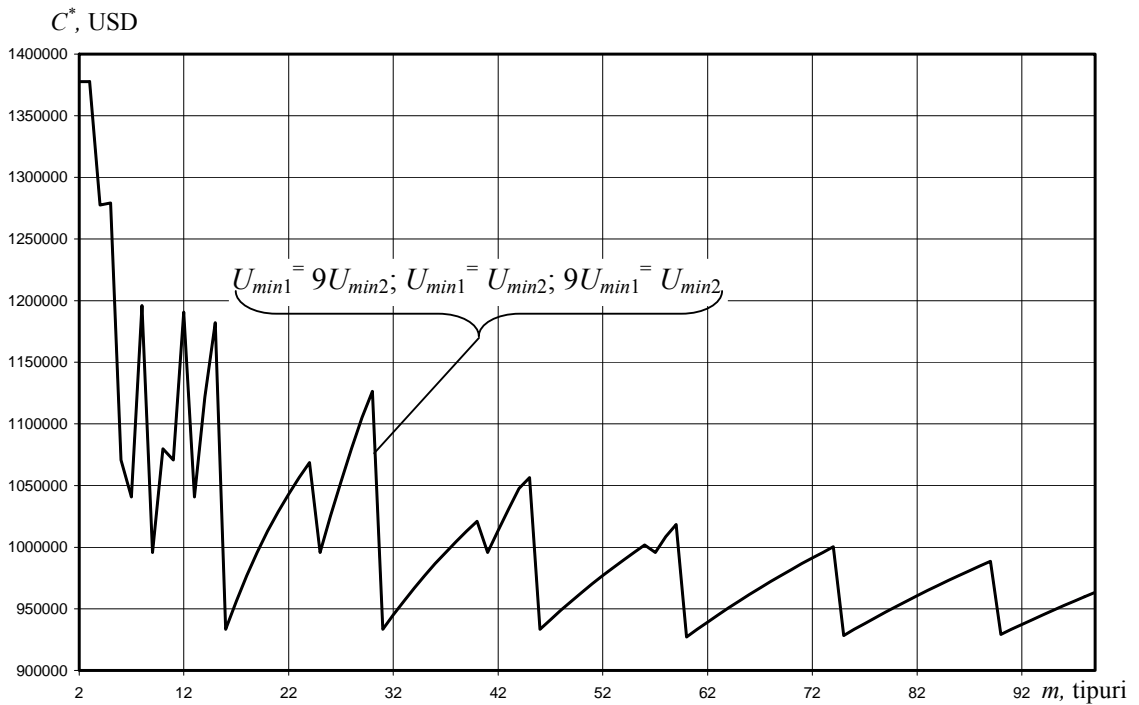
c) la  $U_{\min} = 1250$  Gflops pentru cazurile: 1)  $U_{\min 1} = 9U_{\min 2}$ ; 2)  $U_{\min 1} = U_{\min 2}$  și 3)  $U_{\min 2} = 9U_{\min 1}$  toate soluțiile, la aceeași valoare a lui  $m$ , coincid (vezi fig. 3).

Astfel, se poate conchide că odată cu creșterea laboriozității sumare a fluxurilor de cereri

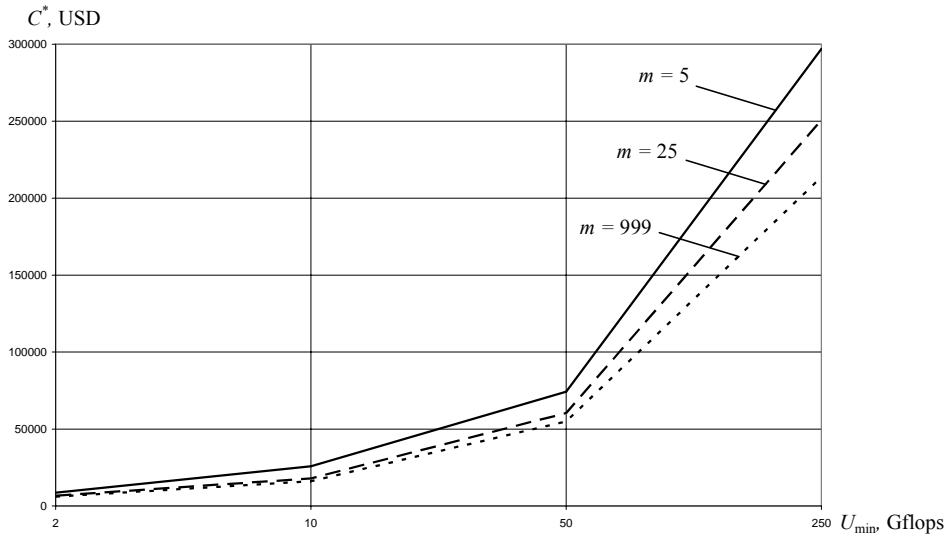
ale utilizatorilor influența structurii fluxului sumar  $\Lambda$  de cereri asupra soluției optime scade și poate fi, începând cu o anumită laboriozitate  $U_{\min}$ , chiar nulă.



**Fig. 2.** Dependența costului  $C^*$  al setului optim de servere de numărul  $m = \{4, 5, \dots, 99\}$  tipuri de servere și structura fluxului sumar  $\Lambda$  de cereri ale utilizatorilor la  $U_{\min} = 10$  Gflops.



**Fig. 3.** Dependența costului  $C^*$  al setului optim de servere de numărul  $m = \{2, 3, \dots, 99\}$  tipuri de servere și structura fluxului sumar  $\Lambda$  de cereri ale utilizatorilor la  $U_{\min} = 1250$  Gflops.



**Fig. 4.** Dependența costului  $C^*$  al setului optim de servere de laboriozitatea  $U_{\min} = \{2, 10, 50$  și  $250\}$  Gflops a fluxului sumar  $\Lambda$  de cereri la  $U_{\min 1} = 9U_{\min 2}$ .

### Concluzii

În baza rezultatelor calculelor efectuate folosind aplicația ConfigServ, se constată o dependență semnificativă a costului optim  $C^*$  a setului omogen de servere de numărul  $m$  de tipuri de servere din care se efectuează selecția; varierea valorii soluției optime poate depăși 39 %. Totodată, influența structurii fluxului sumar  $\Lambda$  de cereri ale utilizatorilor asupra soluției optime este considerabilă la laboriozități  $U_{\min}$  relativ mici ale  $\Lambda$  și scade odată cu creșterea  $U_{\min}$ . Aceasta confirmă necesitatea unor cercetări speciale dependente de caz la configurarea seturilor de servere ale rețelelor de calculatoare.

### Referințe

1. Bolun I., Ciurac A. Configuration of local area network set of servers/ Computer Science Journal of Moldova, vol. 10, No. 2(29), 2002. Chisinau. - pp. 99-124.
2. Ferrari D. Computer systems performance evaluation. New-Jersey: Prentice Hall, 1978.