

EVALUAREA PERFORMANTELOR PROCESELOR DE CALCUL PRIN REȚELE PETRI GENERALIZATE STOCASTICE

E. Guțuleac

Universitatea Tehnică a Moldovei

INTRODUCERE

Concepția unui sistem de calcul, bazat pe paralelism și coperarea componentelor de la specificația sa inițială până la verificarea faptului că dezvoltarea propusă satisface cerințelor de performanță, necesită un mediu de dezvoltare. În acest context, este de dorit ca într-un astfel de mediu să fie folosit un singur formalism, care să aibă capacități suficiente pentru descrierea acțiunilor proceselor cooperante și evaluarea performanțelor sistemului analizat [3].

Pentru a descrie adecvat în cadrul aceluiași formalism procesele de calcul, a efectua verificarea și apoi evaluarea caracteristicilor numerice de performanță acestor procese în [1, 2] a fost definită o clasă de rețele Petri generalizate markoviene (RMG), cu capacități negativ-pozitive ale locațiilor și arce reversibile cu automodificare.

În această lucrare este efectuată analiza performanțelor a unor procese de calcul specifice.

1. MODELE MARKOVIENE ALE PROCESELOR DE CALCUL

Clasă de rețele RMG dă posibilitatea de a mări flexibilitatea modelului la descrierea unor fenomene de funcționare ale proceselor concurente, condiționate de starea curentă a sistemelor de calcul paralele/distribuite.

În continuare, vom descrie prin RMG procesele de prelucrare a cererilor de procesare a datelor de către unitatea aritmetică-logică (ALU) a procesorului conform programului memorat în memoria centrală (MC). Pe parcursul executării unui segment de program operanzii și instrucțiunile vor fi transmise din MC în memoria-tampon (MT), stocate și ordonate conform disciplinei de servire primul sosit – primul servit FIFO (sau în stivă cu disciplina primul sosit – ultimul servit LIFO), iar de acolo sunt prelucrate de ALU în ordinea determinată de către acest segment de program [3].

Eventual, la apariția instrucțiunii IF – salt condiționat determinat de rezultatul executării

acestei instrucțiuni, ordinea de prelucrare succesivă a instrucțiunilor poate fi schimbată, iar operanzii din MT pot să aparțină următorului segment de program ce se află îndată după instrucțiunea IF deja executată (rezultatul este “True” și deci direcția secvenței este adoptată) sau să nu aparțină acestui segment (“False”). În acest caz, toți operanzii ce se află în MT sunt anulați și MT devine vidă.

Expresia descriptivă DE1 ce descrie acest proces este următoarea [2]:

$$DE1 = (n+k) p_1 |_{t_1}^{\lambda_1} p_2 |_{t_2}^{\lambda_2} p_3 |_{t_3}^{q_3} p_1 \vee DE'1$$

$$DE'1 = (p_2[M(p_2)] \cdot p_3) |_{t_3}^{q_3} p_1 [1 + M(p_2)].$$

În Fig. 1 este reprezentată maparea lui DE1 în o atare rețea RMG1.

Rețeaua GeN subiacentă RMG1 este o rețea cu arce ce au ponderi automodificabile:

$$W(p_2, t_3) = M(p_2) \text{ și } W(t_3, p_1) = 1 + M(p_2)$$

ce sunt dependente de marcajul curent al locației p_2 , care descrie șirul de așteptare FIFO al memoriei-tampon. Ponderile celorlalte arce sunt egale cu 1.

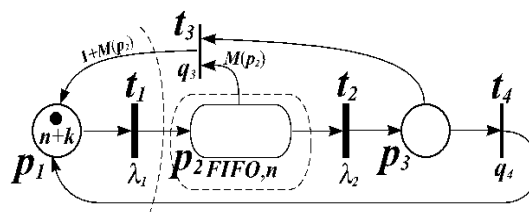


Figura 1. Maparea lui DE1 în RMG1.

Tranzițiile temporizate t_1 și t_2 respectiv modelează procesul de transmitere a operanzilor din MC și procesul prelucrării de către ALU al operanzilor din MT. Rata de declanșare a tranzițiilor t_1 și t_2 respectiv este λ_1 și λ_2 . Locația p_3 și tranzițiile imediate t_3, t_4 descriu selectorul probabilistic în conformitate cu care se va alege modul de executare a cererilor, dacă rezultatul instrucțiunii IF este “True” sau “False”. Cererea, fiind prelucrată de ALU (tranziția temporizată t_2), părăsește sistemul cu probabilitatea q_4 (tranziția

t_4) și cu probabilitatea q_3 (tranziția t_3) poate solicita anularea cererilor stocate în MT și în ALU .

Locația p_1 descrie populația cererilor în sistem. Această rețea are un P -invariant: $M(p_1) + M(p_2) + M(p_3) = n$ și deci ea este mărginită, viabilă și reinițializabilă. Lanțul Markov timp continuu inclus ($LMTTC1$) ergodic [3], asociat cu această rețea este reprezentat în Fig. 2.

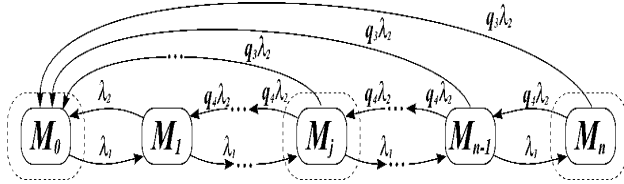


Figura 2. Lanțul $LMTTC1$ inclus al $RMG1$.

Fie π_i sunt probabilitățile staționare de stare ale $LMTTC1$. Pentru tăieturile, respectiv prezentate în Fig. 1 și Fig. 2 obținem următoarele ecuații de echilibru ale fluxurilor de probabilitate în regim staționar:

1. $\lambda_1 \pi_0 = \lambda_2 \pi_1 + q_3 \lambda_2 \sum_{j=1}^n \pi_j$; (1)
2. $(\lambda_1 + \lambda_2) \pi_j = \lambda_1 \pi_{j-1} + q_4 \lambda_2 \pi_{j+1}$, $\forall j = \overline{1, n-1}$;
3. $\lambda_1 \pi_n = \lambda_1 \pi_{n-1}$.

Din ecuația 1 a lui (1), luând în considerație relația de normalitate, obținem: \square

$$\pi_0 = \rho [q_4 \pi_1 + q_3 (1 - \pi_0)], \quad \rho = \lambda_1 / \lambda_2. \quad (2)$$

Ecuația 2 din (1) este o ecuație recurentă de ordinul doi [3], soluția generală a căreia este:

$$b_k = \pi_{n+1-k} = [b_2 - (1 + \rho)b_1] y_{k-1} + b_1 y_k,$$

$$\forall k = \overline{1, n+1}, \text{ unde: } y_k = (x_1^k - x_2^k) / (x_1 - x_2), \quad (3)$$

$$x_{1,2} = 0,5(1 + \rho) \pm \sqrt{0,25(1 + \rho)^2 - \rho q_4} \quad (4)$$

Luând în considerație ecuația 3 a sistemului (1), relațiile (3) și (4) pentru b_k , obținem:

$$b_k = b_1 (y_k - y_{k-1}), \quad \forall k = \overline{1, n+1}. \quad (5)$$

Mărimea $b_1 = 1 / y_{n+1}$ este determinată din ecuația 1 a sistemului (1), luând în considerație faptul că $\pi_1 = b_n$, iar $\pi_0 = b_{n+1}$.

Probabilitatea $\pi_A(n)$ că ALU va fi în stare activă de prelucrare a cererilor și durata medie $\bar{\tau}_s(n)$ de prelucrare a unei cereri de către sistemul $MC-ALU$ sunt determinate de următoarele expresii:

$$\begin{aligned} \pi_A(n) &= 1 - \pi_0 = b_n / b_{n+1}; \\ \bar{\tau}_s(n) &= 1 / (\lambda_2 \pi_A(n)) = b_{n+1} / \lambda_2 b_n. \end{aligned} \quad (6)$$

Notăm că pentru $\lambda_2 < \lambda_1$, ($\rho < 1$), $n \rightarrow \infty$ și $q_3 = 0$, când cererile nu pot fi anulate, atunci probabilitatea asimptotică de aflare a ALU în stare pasivă tinde către zero, adică $\pi_0 = 0$. Acest fapt are loc, fiindcă în regim staționar șirul $FIFO$ cu un număr nelimitat de locuri de așteptare crește spre infinit din cauză că rata de sosire a cererilor este mai mare decât cea de servire și deci ALU totdeauna va fi în stare activă, însă modul de funcționare al sistemului este instabil. Este interesant de a analiza această situație similară, când există cazul anulării cererilor ($q_3 \neq 0$) la apariția instrucțiunilor tip IF .

Pentru aceasta vom împărți la x_1^n numărătorul și numitorul expresiei (3) pentru $\pi_A(n)$, luând în considerație (1) obținem următoarea expresie:

$$\pi_A(n) = (1 - \beta) / (x_1 - x_2 \beta), \text{ unde } \beta = (x_2 / x_1)^n.$$

Din aceste expresii verificăm că pentru $0 < q_3 < 1$, totdeauna are loc relația $x_1, x_2 > 0$ cu $x_1 > x_2$, de unde reiese că:

$$0 < \lim_{n \rightarrow \infty} \pi_A(n) = 1 / x_1 < 1.$$

Remarcăm că pentru $\forall \beta > 0$ și $q_3 > 0$ totdeauna avem $x_1 > 1$. Aceasta înseamnă că, independent de raportul ratelor de sosire și de servire a cererilor în sistem, probabilitatea asimptotică ($\pi_A(\infty) = 1 / x_1$) < 1 că ALU să fie în stare activă este o mărime subunitară.

În modelul $RMG1$ se presupune că șirul $FIFO$ al MT pentru stocarea cererilor este un dispozitiv *neinertial*, pentru care s-a neglijat durata de transfer a operanzilor din MC prin MT către ALU .

În modele fără anularea cererilor această ipoteză este justificată, fiindcă în regim staționar cu $0 < \rho = \lambda_1 / \lambda_2 < 1$ probabilitatea că MT va fi vidă (șirul $FIFO$ este fără cereri) este mică, iar durata de acces-alocare a MT este cu mult mai mică decât durata de servire a cererilor de MC și ALU .

Pentru modelul ce va lua în considerație anularea cererilor din MT probabilitatea ca șirul $FIFO$ să fie *vid* se mărește cu mult și, de asemenea, se măresc și pierderile neproductive de timp, legate de executarea operațiilor de accesare a MT și transmiterea cererilor către ALU , ce nu pot fi neglijate. Astfel, MT devine un dispozitiv *inertial*, deoarece acțiunile de anulare a cererilor au nevoie de o durată de timp pentru analiza cauzei și terminarea servirii fiecărei cereri. De aceea este necesar de a studia influența acestor factori asupra productivității modelului $MC-ALU$.

Fie $\bar{\tau}_{\alpha_j} = (1 + \alpha_j) / \lambda_1$ este durata medie de aflare în starea M_j ce va lua în considerație durata suplimentară pentru analiza cererii anulate, unde parametrul α_j determină pierderile de timp necesare pentru terminarea servirii de MC și introducerea unui operand în MT vidă. Mărimea α_j este funcție de numărul n de registre consecutive, folosite pentru organizarea funcționării șirului de așteptare. Dacă durata de transfer de la un registru la altul este α / λ_1 , atunci $\alpha_j = j \cdot \alpha$, $j = \overline{0, n}$, iar $\bar{\tau}_{\alpha_j} = (1 + j \cdot \alpha) / \lambda_1$ este durata medie necesară pentru analiza cererilor anulate, servirea unei noi cereri de MC și transferul ei prin MT vidă către ALU. Astfel, MT devine un dispozitiv inerțial, care este necesar de a folosi o durată suplimentară pentru trecerea cererilor de la MC către ALU.

Expresia descriptivă DE2 ce descrie acest proces este următoarea:

$$DE2 = np_1 |_{t_1}^{\lambda_1} p_2 |_{t_2}^{\lambda_2} p_3 |_{t_3}^{q_3} p_1 \vee DE'2,$$

$$DE'2 = (p_2[M(p_2)] \cdot p_3) |_{t_3}^{q_3} \vee DE''2,$$

$$DE''2 = |_{t_3}^{q_3} p_4[1 + M(p_2)] \vee DE'''2,$$

$$DE'''2 = p_4[M(p_2)] |_{t_5}^{\lambda_5(m_4)} p_1[M(p_2)].$$

Maparea lui DE2 în formă de rețea RMG2 este prezentată în Fig.3.

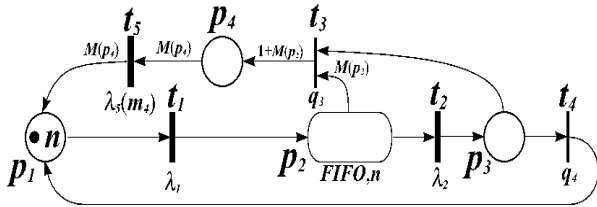


Figura 3. Maparea lui DE2 în RMG2.

Această rețea este obținută în baza rețelei RMG1 din Fig. 2.1 și diferă de ea prin faptul că RMG2 are o locație p_4 și o tranziție temporizată suplimentară t_5 ce modelează procesul inerțial la anularea cererilor din sistem, iar rata declanșării acestei tranziții este dependentă de marcajul $m_4 = M(p_4)$:

$$\lambda_5(m_4) = 1 / \bar{\tau}_{\alpha_j} = \lambda_1 / (1 + j\alpha), \quad j = \overline{0, n}.$$

Ponderea arcelor este aceeași ca și pentru RMG1, iar $W(p_4, p_5) = W(p_5, p_4) = M(p_4)$. Această rețea are un P-invariant: $M(p_1) + M(p_2) + M(p_3) + M(p_4) = n$ și deci ea este mărginită, viabilă și reinițializabilă.

Lanțul LMTc2 ergodic al RMG2 este reprezentat în Fig. 4. El diferă de LMTc2 prin aceea că fiecărei stări M_j îi corespunde o stare ce ea în considerație efectul inerțial al MT.

Ecuatiile de echilibru ale fluxului de probabilitate pentru fiecare stare a LMTc8 sunt:

1. $(\lambda_1 / (1 + j\alpha)) \hat{\pi}_j = \lambda_2 q_3 \pi_j, \quad j = \overline{1, n}$;
2. $(\lambda_1 + \lambda_2) \pi_j = \lambda_1 \pi_{j-1} + \lambda_2 q_4 \pi_{j+1}, \quad j = \overline{0, n-1}$;
3. $\lambda_1 \pi_0 = \lambda_1 \cdot \sum_{j=1}^n \hat{\pi}_j / (1 + j\alpha)$; (7)

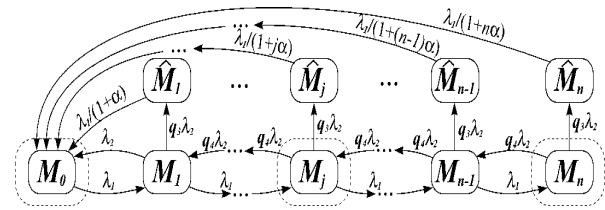


Figura 4. Lanțul LMTc2 inclus al RMG2.

Soluționarea sistemului de ecuații (7), împreună cu relația de normalitate, este efectuată în mod similar cum și pentru LMTc1 [3].

Pentru anumite valori ale parametrilor $\gamma = \lambda_1 / \lambda_2$, q_3 și α funcția $\pi_A(n)$ are un maximum n^* și este o mărime finită.

În Fig. 5 sunt prezentate unele caracteristici ale n^* funcție de q_3 pentru diferite valori ale α care determină pierderile de timp necesare pentru terminarea servirii de MC.

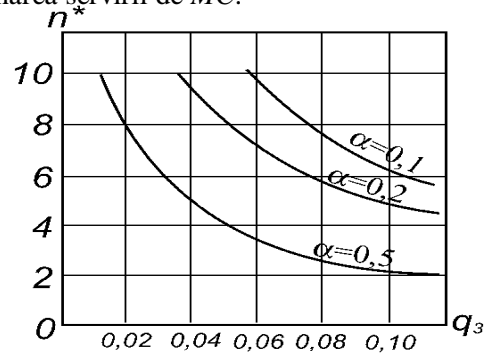


Figura 5. Valorele lui n* funcție de q3

2. MODELAREA PROCESELOR DE CALCUL NEMARKOVIENE

În teoria sistemelor de așteptare, când distribuția duratei de servire a cererilor deseori diferă de cea exponențial-negativă, aceasta poate fi reprezentată prin distribuția tip fază (PH, Phase type) cu faze fictive [3, 4]. În acest caz zona de servire (ZS) a unui server este înlocuită printr-o

subrețea cu mai multe faze de servire, reprezentate prin serveri virtuali, unde cererea obține la fiecare fază o durată de servire distribuită conform legii exponențial-negative cu parametrul respectiv. În ZS , dacă ea este activă, poate să se afle numai o singură cerere, care va vizita fazele respective conform algoritmului de funcționare al acestei ZS .

În general, o lege de distribuție a unei variabile aleatorii X_τ este caracterizată de către coeficientul său de variație (K_V), determinat de către abaterea $\sqrt{D[X_\tau]}$ ce este raportată la speranța matematică $E[X_\tau]$ a acestei variabile X_τ , adică $K_V = \sqrt{D[X_\tau]} / (E[X_\tau])$.

Pentru distribuția exponențial-negativă avem: $F_{\text{exp}}(\tau) = 1 - \exp(-\mu\tau)$, $\tau \geq 0$;

$$\bar{\tau}_s = E[X_\tau] = 1/\lambda, D_s = D[X_\tau] = 1/\lambda^2, K_V = 1.$$

Dacă $K_V \neq 1$, atunci durata de servire a cererilor diferă de cea exponențial-negativă și astfel procesul stocastic ce descrie comportarea sistemului nu totdeauna este un proces Markov. Pentru a analiza aceste cazuri se folosesc diferite metode [3], dintre care este metoda reprezentării prin distribuții tip PH , care descompune durata de servire arbitrar distribuită în faze fictive în durate de servire exponențial-negative cu parametrii respectivi și care trebuie identificați având la dispoziție eşantioane statistice respective ce caracterizează această lege de distribuție [3, 4, 5].

O distribuție tip PH de ordinul k poate fi descrisă cu o subrețea RMG sigură, folosind noțiunea de zonă de declanșare (ZD) a tranziției temporizate $t \in T_\tau$. Algoritmul de funcționare a ZD care descrie declanșarea tranziției $t \in T_\tau$, substituită cu o subrețea RMG sigură tip PH , este similar cu al disciplinei de funcționare a zonei de serviciu. Dacă tranziția $t \in T_\tau$ (substituită cu o ZD tip PH) declanșează, atunci în ZD poate să se afle nu mai mult de un jeton și intrarea unui alt jeton este sincronizată cu ieșirea jetonului curent din ZD . Deci dacă ZD a tranziției $t \in T_\tau$ declanșează, imediat ea va sustrage jetoane respective din toate locațiile de intrare $\forall p \in \bullet t$ și va depune un jeton în locația de intrare a tranziției temporizate ce modelează prima fază de servire, conform legii exponențial-negative cu rata respectivă. După terminarea servirii la prima fază se vor declanșa următoarele tranziții temporizate ce redau fazele respective. La terminarea declanșării tranziției ce redă ultima fază imediat se vor depune în locațiile $p \in t^\bullet$ jetoane

respective și concomitent va fi autorizată intrarea în ZD a unui alt jeton din șirul de așteptare, dacă această tranziție declanșează.

Fie, de exemplu, durata de declanșare a tranziției temporizate $t_2 \in T_\tau$ rețelei $RMG4$ are o distribuție cu coeficientul de variație $0 < K_V < 1$, legea *Erlang* de ordinul k , notată E_k [3]. În acest caz tranziția $t_2 \in T_\tau$ poate fi considerată ca o macrotranziție temporizată, care poate fi substituită cu o subrețea $RMG3$, ce constituie ZD a acestei tranziții, reprezentată în linii întrerupte, precum este indicat în Fig. 6. Expresia descriptivă DE_{RMG3} a rețelei GeN subiacente $RMG4$ este [2, 3]:

$$DE_{RMG3} = (p_{k+2} \cdot p_{k+3})|_{t_{2,k+1}} p_1 \vee DE'_{RMG3},$$

$$DE'_{RMG3} = \bigvee_{j=1}^{k-1} (p_j |_{t_{2,j}} p_{j+1}) \vee p_k |_{t_{2,k}} (p_1 \diamond p_{k+3})$$

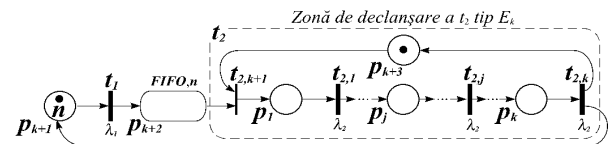


Figura 6. $RMG3$ cu ZD a lui t_2 conform E_k .

Intrarea unui jeton în ZD este controlată de către un semafor, constituit din locația p_{k+3} , ($M(p_{k+3}) = 1$) și tranziția imediată $t_{2,k+1} \in T_0$. Dacă ZD declanșează, atunci imediat declanșează tranziția $t_{2,k+1}$ și din locațiile $\{p_{k+2}, p_{k+3}\} \in \bullet t_{2,k+1}$ se va scoate câte un jeton și se va depune un jeton în locația p_1 , care apoi va parcurge succesiunea de locații p_j și succesiunea de tranziții $t_{2,j} \in T_\tau$, ($j = \overline{1, k}$). Aici tranzițiile temporizate $t_{2,j}$ redau fazele de servire cu aceeași rată de declanșare λ_2 conform legii exponențial-negative. Obținând servirea la ultima fază, redată de către tranziția temporizată $t_{2,k}$, jetonul (cererea) va părăsi ZD și va reactiva semaforul și dacă în șirul *FIFO* (locația p_{k+2}) sunt cereri, atunci cererea care este prima în acest șir va intra imediat în ZD și astfel procesul de declanșare a tranziției va reîncepe, urmând algoritmul de funcționare al ZD .

O altă lege de distribuție tip PH este distribuția *Hiperexponențială* de ordinul k , notată ca distribuția H_k , care are un coeficient de variație $1 < K_V < +\infty$. În acest caz durata de servire a unei cereri conform legii de distribuție $F_{H_k}(\tau)$ poate fi

interpretată ca fiind efectuată de către distribuția exponențial-negativă cu rata λ_j , ce este selectată cu probabilitatea q_j , ($j = \overline{1, k}$), de fiecare dată dintr-un set de k distribuții exponențial-negative cu ratele respective.

Expresia descriptivă DE_{RMG4} a rețelei GeN subiacente $RMG4$ este:

$$DE_{RMG4} = (p_2 \cdot p_4) |_{t_{2,k+1}} p_3 \vee DE'_{RMG4},$$

$$DE'_{RMG4} = \bigvee_{j=1}^k (p_3 |_{t_{2,j}} p_5) \vee p_5 |_{t_{2,k+2}} (p_1 \diamond p_4).$$

Maparea lui DE_{RMG4} în rețeaua $RMG4$ este reprezentată în Fig. 7, unde cu linii întrerupte este redată o zonă de declanșare a tranziției $t_2 \in T_\tau$ conform legii de distribuție $F_{H_k}(\tau)$ ce este descrisă de o subrețea RMG sigură. Această ZD este constituită din tranziții temporizate $t_{2,j}$, ($j = \overline{1, k}$), situate în paralel, durata de declanșare a cărora este distribuită conform legii exponențial-negative cu rata $\lambda_{2,j}$, ($j = \overline{1, k}$), ce redă faza de servire respectivă a cererii. Alegerea tranziției $t_{2,j}$ ce trebuie să fie declanșată este efectuată cu probabilitatea q_j , atașată acestei tranziții și este determinată de către selectorul probabilistic $0 < q_j < 1$, care este redat de către locația p_3 , partajată de către tranzițiile de ieșire din ea, $p_3^* = \{t_{2,j}, j = \overline{1, k}\}$. Intrarea în (respectiv ieșirea din) ZD a cererilor este controlată de către un semafor constituit din locația p_4 și tranziția imediată $t_{2,k+1}$ (respectiv $t_{2,k+2}$).

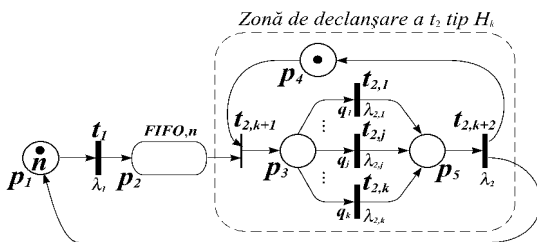


Figura 7. $RMG4$ ce descrie ZD a lui t_2 după H_k .

Expresia descriptivă DE_{RMG5} a rețelei GeN subiacente $RMG4$ este:

$$DE_{RMG5} = (p_2 \cdot p_4) |_{t_{2,k+1}} p_3 \vee DE'_{RMG5},$$

$$DE'_{RMG5} = \bigvee_{j=1}^k (p_3 |_{t_{2,j}} p_5) \vee p_5 |_{t_{2,k+2}} (p_1 \diamond p_4).$$

Dacă legea de distribuție tip PH are un coeficient de variație $0,5 \leq K_V < +\infty$, atunci

această distribuție este $Cox-k$, care poate fi schematizată prin funcționarea ZD în modul următor. Cererea, intră în ZD , va fi servită la prima fază, apoi fie că va părăsi ZD cu probabilitatea b_1 , fie că va continua să fie servită cu probabilitatea a_1 , $a_1 + b_1 = 1$. Astfel, dacă cererea a terminat servirea sa la fază p_j , de fiecare dată se va selecta cu probabilitatea b_j de a părăsi ZD sau cu probabilitatea a_j , $a_j + b_j = 1$, ($j = \overline{1, k}$) de a continua servirea la următoarea fază. La fiecare fază durata de servire este distribuită conform legii exponențial-negative cu rata respectivă λ_j .

În Fig. 8 este reprezentată o rețea $RMG5$ ce descrie un sistem de așteptare de tipul $SA: M|Cox-k|1/n|(FIFO, blocaj)$ cu o subrețea RMG sigură, care redă în linii întrerupte zona de declanșare a macrotranziției temporizate $t_2 \in T_\tau$ conform distribuției $Cox-k$.

Această subrețea $RMG5$ este constituită dintr-o succesiune de tranziții temporizate $t_{2,j} \in T_\tau$, ($j = \overline{1, k}$) și locații $p_j \in \bullet t_{2,j}$ și $\hat{p}_j \in t_{2,j}^*$, care redau fazele respective de servire a cererii. Dacă jetonul ce reprezintă cererea în ZD se află în locația p_j , ($j = \overline{1, k}$), atunci este activată faza respectivă, redată de către $t_{2,j}$, durata declanșării căreia este distribuită conform legii exponențial-negative cu rata $\lambda_{2,j}$.

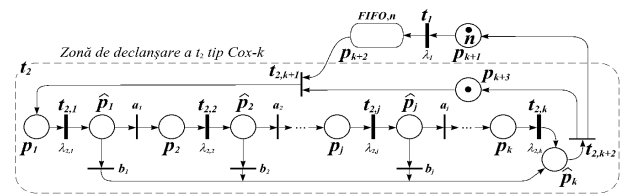


Figura 8. $RMG5$ ce descrie ZD a lui t_2 după $Cox-k$.

După declanșarea tranziției $t_{2,j}$ în locația $\hat{p}_j \in t_{2,j}^*$ va apărea un jeton, care îndată va valida două tranziții imediate ce se află în conflict. Aceste tranziții imediate determină un selector probabilistic $0 < a_j < 1$, $a_j + b_j = 1$, ($j = \overline{1, k-1}$), adică cu probabilitatea b_j cererea (jetonul) poate părăsi ZD și astfel macrotranziția $t_2 \in T_\tau$ substituită de această ZD va termina declanșarea sa sau cu probabilitatea a_j cererea (jetonul) va activa următoarea fază de servire, adică se va valida

tranziția $t_{2,j+1} \in T_\tau$, declanșarea căreia va activa următoarea fază sau declanșarea lui $t_2 \in T_\tau$.

Deseori, utilizatorul care dorește să modeleze activitatea unui server, realmente din statistică nu dispune decât de primele două momente ale distribuției duratei de servire X_τ , adică valoarea medie $\bar{\tau}_s = E[X_\tau]$, și dispersia $D_s = D[X_\tau]$. În măsura în care valorile acestor prime două momente vor fi specificate, utilizatorul, în general, nu va avea alte exigențe, decât tipul legii de distribuție a duratei de servire folosită. Dacă acesta este cazul, atunci majoritatea cerințelor utilizatorului pot fi satisfăcute cu ajutorul unei legi de distribuție tip *Cox-2* de ordinul 2 sau tip *Erlang-k* de ordinul k . Precizăm acest fapt, identificând parametrii acestor legi de distribuție prin valoarea medie $\bar{\tau}_s$ și coeficientul său de variație. Se poate demonstra [4], că pentru orice valoare medie $\bar{\tau}_s$ și pentru orice coeficient de variație K_V , astfel încât $0,5 \leq K_V < +\infty$, este posibil de a determina, cel puțin, o lege *Cox-2* ce posedă o valoare medie și dispersie dorită. Parametrii unei astfel de distribuții *Cox-2* sunt determinați de către următoarele expresii:

- 1) dacă $0,5 \leq K_V < 1$: $\mu_1 = 2/\bar{\tau}_s$,
 $\mu_2 = 1/(\bar{\tau}_s \cdot K_V)$, $a_1 = 2(1 - K_V)$;
- 2) dacă $1 < K_V < +\infty$: $\mu_1 = 1/(\bar{\tau}_s \cdot K_V)$,
 $\mu_2 = 2/\bar{\tau}_s$, $a_1 = 1/(2 \cdot K_V)$.

Aici μ_1 și μ_2 sunt ratele de servire respective la prima și a doua fază, adică pentru ZD reprezentată în Fig. 2.8 pentru RMG6, avem $\lambda_{2,1} = \mu_1$ și $\lambda_{2,2} = \mu_2$, iar $0 < a_1 < 1$ este probabilitatea selectării fazei a doua de servire după ce a fost terminată prima fază de servire.

Dacă coeficientul de variație K_V este inferior 1/2, adică $0 < K_V < 0,5$, atunci avem o altă lege ce diferă de *Cox-2*. În acest caz putem să ne mulțumim de a descrie servirea cererilor cu legea *Erlang-k*, fiindcă $K_V = 1/k$. Dacă nu există un astfel de întreg $k > 0$, încât $1/k$ să fie considerat ca suficient de aproape de K_V , atunci se poate de folosit modelul *Erlang-k generalizat* [4] care este un caz particular al legii *Cox-k*, când $0 < a_1 < 1$, $a_1 + b_1 = 1$ și $a_j = 1$, $b_k = 1$, iar $\lambda_{2,j} = \mu$, $\forall j = 2, k-1$.

Pentru această lege parametrii k , $b_1 = 1 - a_1$ și μ sunt determinați în modul următor:

1) fiindcă este necesar de a avea un număr minim de faze posibile, se va găsi un astfel de întreg $k > 0$, încât:

$$1/k < K_V \leq 1/(k-1);$$

$$b_1 = (2k \cdot K_V + k - 2 - \sqrt{k^2 + 4(1 - k \cdot K_V)})/c,$$

$$c = (2(k-1) \cdot (1 + K_V));$$

$$\mu = b_1 + k(1 - b_1)/\bar{\tau}_s.$$

Astfel, de exemplu, dacă utilizatorul, din prelucrarea unui eșantion statistic experimental al duratei de servire, a determinat valoarea medie $\bar{\tau}_s = 0,5$ și dispersia $D[X_\tau] = 0,5$, atunci pentru parametrii legii *Cox-2* obținem: $K_V = 4$, $\mu_1 = 4$, $\mu_2 = 0,5$, $a_1 = 1/8$ și $b_1 = 7/8$.

5. CONCLUZII

În lucrare sunt considerate unele aspecte de modelare și de evaluare a performanțelor proceselor de calcul prin rețele Petri generalizate stocastice descriptiv-compoziționale. În același context sunt considerate unele aspecte de modelare a proceselor de calcul nemarkoviene prin distribuții tip fază.

Bibliografie

1. **Guțuleac, E.** Aspecte de modelare a proceselor de calcul prin rețele Petri generalizate stocastice cu arce reversibile. *Meridian ingineresc*, Nr. 4, 2007, Chișinău, pp.48-52.
2. **Guțuleac, E.** Descriptive Compositional Construction of Generalized Stochastic Petri Net Models for Performance Evaluation of Computer Systems. *Buletinul Institutului Politehnic din Iași, Tomul LIV, Fasc. 1-4, Automatica și Calculatoare*, 2004, România, pp. 143-159.
3. **Guțuleac, E.** Evaluarea performanțelor sistemelor de calcul prin rețele Petri stocastice. *Editura „Tehnica-Info”, Chișinău*, 2004, 276 p.
4. **Gelenbe E. et al.** *Reseaux de files d'attente. Modelisation et traitement numerique*. Ed.: « Hommes et Technique », 1980. 178 p.
5. **Kleinrock L.** *Queueing Systems, vol.1 "Theory", vol. 2 "Computer applications"*. Wiley&Sons, 1976, 600 p.

Recomandat spre publicare: 26.03.08