

УДК 004.32

О. Я. Матов

Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна

Математичні моделі конфліктних утрат продуктивності системи посередників онтології загального використання в GRID-середовищі

Визначено види можливих конфліктів і розроблено аналітичні моделі для дослідження значень коефіцієнта втрат продуктивності наступних структур системи посередників онтології загального використання у GRID-середовищі: однорідних і неоднорідних бібліотек; при відсутності та наявності високопродуктивних кешів; для одно- та багатовхідних бібліотек.

Ключові слова: імовірнісна модель, модель втрат продуктивності, GRID-середовище.

Інфраструктура системи організації рішення задач над множинними інформаційними джерелами належить до класу інформаційної GRID-архітектури. Вона забезпечує розробників стандартними інтерфейсами для включення (plug-in) нових програмних інструментів і баз даних одночасно із публікацією їхніх метаданих стандартним чином. Інфраструктура реалізується в середовищі WEB на основі WEB-сервісів [1].

Основним виконавчим механізмом інфраструктури є система управління потоками робіт, що дозволяє використовувати виклики довільних сервісів разом із запитамися до баз даних і до посередника, як їхні окремі кроки.

Інфраструктура містить також засоби підтримки реєстрів метаданих на основі протоколів Open Archive Initiative (OAI).

Згадана інфраструктура фізично є оболонкою, в яку вбудовуються засоби підтримки посередників періоду виконання. Ці засоби включають:

- репозиторій зберігання метаінформації посередника;
- мову формулювання запитів до посередника;
- засоби переписування запитів до посередника в звернення до зареєстрованих у посереднику інформаційних джерел (шляхом використання їхніх адаптерів);
- засоби планування реалізації запитів у розподіленому середовищі;
- засоби управління посередниками.

Сам посередник реалізується при цьому як WEB-сервіс.

Кожен посередник під час реалізації запитів звертається до окремих онтологій (бібліотек онтологій) загального використання за зчитуванням понять. При цьому розрізняють фазу підготовки запиту до онтології загального використання (ОЗВ), що є проміжком часу від одержання результату попереднього запиту до формування наступного запиту, і фазу обслуговування запиту, яка є проміжком часу від одержання запиту на звертання до ОЗВ до видачі результату. Фаза обслуговування запиту включає час на з'єднання посередника з необхідною бібліотекою та час зчитування необхідного поняття з бібліотеки.

При звертанні великої кількості посередників до бібліотек ОЗВ неминучі ресурсні конфлікти. Вони являють собою ситуацію, при якій два чи більше споживачів вимагають один і той самий ресурс. У СП (система посередників) ОЗВ можливі ресурсні конфлікти двох видів:

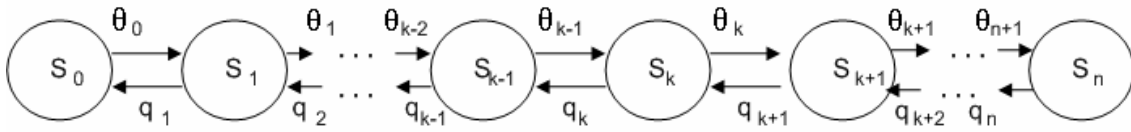
— конфлікти бібліотек при одночасному звертанні двох і більше посередників до однієї бібліотеки;

— конфлікти в операційному вузлі системи управління ОЗВ, що виникають у випадку, якщо число посередників, що одночасно звернулися до ОЗВ, перевищує пропускну здатність цього операційного вузла.

При виникненні конфліктних ситуацій запит посередника повинен очікувати звільнення ресурсу, що знижує продуктивність СП ОЗВ, яка є критичною для систем із динамічним зв'язуванням. Для оцінки ступеню впливу конфліктів на продуктивність необхідно мати модель, що визначає залежність коефіцієнта конфліктних утрат від структури СП ОЗВ, моделі звертань посередників до бібліотек ОЗВ.

Припускається, що СП ОЗВ складається з n однорідних посередників, m однорідних бібліотек, операційного вузла, що забезпечує зв'язок будь-якого посередника з будь-якою бібліотекою. Операційний вузол, на якому розгорнуто систему управління ОЗВ, здатний одночасно обслуговувати l запитів. Потік запитів посередників до блоків пам'яті — пуасонівський з інтенсивністю λ . Посередник генерує наступний запит до бібліотеки тільки після обслуговування попереднього. Розподіл імовірностей запитів посередників до бібліотек задається матрицею $P = \|p_{ij}\|$, де p_{ij} — ймовірність звертання i -го посередника до j -ї бібліотеки. Час обслуговування посередника бібліотекою розподілено за експонентним законом з інтенсивністю μ . Час з'єднання нехтовно малий у порівнянні з часом підготовки й обслуговування запиту, і його можна не враховувати.

При зазначених припущеннях звертання посередників до бібліотек являє собою марковський процес. Станом S_k СП ОЗВ назвемо ситуацію, при якій k процесів звернулися із запитом до бібліотек ОЗВ. Тоді зі стану S_k можливий перехід тільки до сусідніх станів S_{k+1} (якщо один із $(n - k)$ посередників, що залишилися, згенерує запит на звертання в ОЗВ), або до стану S_{k-1} (якщо в одній із бібліотек закінчено обслуговування запиту). При визначенні станів марковський процес переходу зі стану до стану є процесом «загибелі та розмноження» [2], граф переходів якого наведено на рисунку.



Граф переходів зі стану до стану процесу «загибелі та розмноження»

Символом Θ_k на графі позначено інтенсивність переходу зі стану S_k до стану S_{k+1} , а символ q_k — інтенсивність переходу зі стану S_k до стану S_{k-1} . Процеси «загибелі та розмноження» у теорії марковських процесів добре відомі та допускають аналітичне рішення в явній формі для основних характеристик системи.

Для однорідних посередників $\Theta_k = (n - k)\lambda$, а для однорідних бібліотек $q_k = \eta_k \mu$ (η_k — середнє число запитів, що обслуговуються ОЗВ у стані S_k). Очевидно, в стані S_k у конфліктному просторі знаходиться $(k - \eta_k)$ посередників. Коефіцієнт конфліктних утрат продуктивності δ визначаємо з виразу:

$$\delta = \frac{\sum_{k=1}^n (k - \eta_k) P_k}{n}, \quad (1)$$

де P_k — ймовірність перебування системи в стані S_k . Ймовірність P_k знаходимо наступним чином:

$$P_k = P_0 \prod_{i=0}^{k-1} \frac{\Theta_i}{q_{i+1}}, \quad k = \overline{1, n}. \quad (2)$$

Ймовірність P_0 визначаємо з нормуючої умови (сума ймовірностей по всіх станах повинна дорівнювати 1):

$$P_0 = \left(1 + \sum_{k=1}^n \prod_{i=0}^{k-1} \frac{\Theta_i}{q_{i+1}} \right)^{-1}. \quad (3)$$

Для визначення значень параметрів η_k скористаємося наступними міркуваннями. Нехай $P_k^{(j)}$ — ймовірність того, що в стані S_k до j -ї бібліотеки звернувся принаймні один посередник. Тоді:

$$\eta_k = \sum_{j=1}^m P_k^{(j)}. \quad (4)$$

Ймовірність того, що в стані S_k при заданому наборі Ω посередників жоден із них не звернеться до j -ї бібліотеки, дорівнює $\prod_{\gamma \in \Omega} (1 - P_{i\gamma j}) / C_n^k$, де C_n^k — кількість сполучень із n елементів по k . Підсумовуючи цей вираз за всіма можливими наборами посередників $\Omega_1, \Omega_2, \dots, \Omega_t$, одержуємо:

$$p_k^{(j)} = \frac{1}{C_n^k} \sum_{\alpha=1}^t \left[1 - \prod_{\gamma \in \Omega_\alpha} (1 - p_{i\gamma j}) \right]. \quad (5)$$

Визначимо терм v -го порядку як $\prod_{\gamma=1}^v P_{i\gamma i}$. Тоді вираз для ймовірності $P_k^{(j)}$ буде являти собою алгебраїчну суму термів від 1 до k , причому непарні терми входять до неї зі знаком «плюс», а парні — зі знаком «мінус». Терм k -го порядку буде входити рівно C_{n-v}^{k-v} разів до наборів k із n посередників. Позначимо суму v -го порядку j -ї бібліотеки через B_{vj} . Тоді:

$$p_k^{(j)} = \frac{1}{C_n^k} \sum_{v=1}^k (-1)^{v+1} C_{n-v}^{k-v} B_{vj}. \quad (6)$$

Підставляючи значення ймовірності P_k^j у вираз (4), одержимо:

$$\eta_k = \frac{1}{C_n^k} \sum_{v=1}^k (-1)^{v+1} C_{n-v}^{k-v} \sum_{j=1}^m B_{vj}. \quad (7)$$

Вираз (7) справедливий для $k \leq l$. Якщо ж $k > l$, то до бібліотек мають доступ тільки l посередників і $\eta_k = \eta_l$. Підставивши вираз (7) у (1)–(3), одержимо чисельне значення коефіцієнта конфліктних утрат продуктивності δ .

Для рівноймовірних звертань вираз (7) значно спрощується. У цьому випадку $p_{ij} = \frac{1}{m}$ для всіх i та j . Тоді $B_{vj} = C_n^v m^{-v}$, а $\sum_{j=1}^m B_{vj} = C_n^v m^{1-v}$. Підставляючи значення B_{vj} у вираз (7), одержуємо:

$$\eta_k = \frac{1}{C_n^k} \sum_{v=1}^k (-1)^{v+1} C_{n-v}^{k-v} C_n^v m^{1-v}.$$

Неважко переконатися, що $\frac{C_{n-v}^{k-v} C_n^v}{C_n^k} = C_k^v$.

Звідси $\eta_k = \sum_{v=1}^k (-1)^{v+1} C_k^v m^{1-v}$.

Реальна система має, як правило, складнішу структуру, чим розглянута вище однорідна СП ОЗВ. Відмінними рисами такої структури, суттєвими з погляду визначення коефіцієнта конфліктних втрат продуктивності, можуть бути:

— неоднорідність елементів структури — СП ОЗВ може включати різних посередників, що відрізняються один від одного своїми характеристиками, в тому числі й інтенсивністю звертань до ОЗВ λ . Бібліотеки також можуть відрізнятися інтенсивністю обслуговування;

— наявність багатовхідних бібліотек, у яких можливе одночасне обслуговування декількох запитів;

— наявність можливості кешування.

Розглянемо, яким чином в аналітичній моделі можна врахувати зазначені вище особливості структури СП ОЗВ.

Неоднорідність посередників виявляється насамперед у тім, що i -й посередник звертається до ОЗВ з інтенсивністю λ_i . При цьому в аналітичній моделі змінюється вираз для визначення інтенсивностей Θ_k, q_k . Зі стану S_0 до стану S_1 систему може перевести кожний із посередників з інтенсивністю λ_i . Очевидно, що

$\Theta_0 = \sum_{i=1}^n \lambda_i$. Зі стану S_1 до стану S_2 система може перейти із сумарною інтенсивністю, рівною Θ_0 , за винятком інтенсивності звертання того посередника, що в даний момент обслуговується ОЗВ. Імовірність того, що в стані S_1 обслуговується i -й посередник, дорівнює λ_i / Θ_0 . Тоді інтенсивність переходу зі стану S_1 до стану S_2 визначається виразом:

$$\Theta_1 = \sum_{i=1}^n \frac{\lambda_i (\Theta_0 - \lambda_i)}{\Theta_0}.$$

Зі стану S_2 до стану S_3 система може перейти з інтенсивністю Θ_0 , за винятком інтенсивності звертання тих двох посередників, що в даний момент уже звернулися до ОЗВ. Імовірність того, що в стані S_2 до ОЗВ звернулися посередники i_1 та i_2 може бути представлена рівністю:

$$\frac{\lambda_{i_1} \lambda_{i_2}}{\Theta_0 \Theta_1} + \frac{\lambda_{i_2} \lambda_{i_1}}{\Theta_0 \Theta_1} = \frac{2 \lambda_{i_1} \lambda_{i_2}}{\Theta_0 \Theta_1}.$$

Підсумовуючи цей вираз за всіма можливими наборами з двох посередників, одержимо:

$$\Theta_2 = \sum_{i_1=1}^n \sum_{i_2=i_1+1}^n \frac{\lambda_{i_1} \lambda_{i_2}}{\Theta_0 \Theta_1} (\Theta_0 - \lambda_{i_1} - \lambda_{i_2}).$$

Розглядаючи в загальному випадку перехід зі стану S_{k-1} до стану S_k , інтенсивність Θ_k можна визначити в такий спосіб:

$$\Theta_k = \sum_{i_1=1}^n \sum_{i_2=i_1+1}^n \dots \sum_{i_k=i_{k-1}+1}^n \frac{k! \prod_{v=1}^k \lambda_{i_v} (\Theta_0 - \sum_{v=1}^n \lambda_{i_v})}{\prod_{m=0}^{k-1} \Theta_m}. \quad (8)$$

У формулі (8) множник $k!$ визначає число перестановок із k елементів, а множник $k! \prod_{v=1}^k \lambda_{i_v} / \prod_{m=0}^{k-1} \Theta_m$ — імовірність того, що в стані S_k до ОЗВ звернулися посередники i_1, i_2, \dots, i_k . З урахуванням цього виразу (6) прийме наступний вигляд:

$$P_k^{(j)} = \sum_{i_1=1}^n \sum_{i_2=i_1+1}^n \dots \sum_{i_k=i_{k-1}+1}^n \frac{k! \prod_{v=1}^k \lambda_{i_v} [1 - \prod_{v=1}^k (1 - p_{i_v j})]}{\prod_{m=0}^{k-1} \Theta_m}. \quad (9)$$

Підставляючи вираз (9) у формулу (4), можна визначити середнє число запитів, що обслуговуються в ОЗВ, а значить, й інтенсивність переходу зі стану S_k до стану S_{k-1} .

Неоднорідність бібліотек виявляється насамперед у тім, що j -та бібліотека обслуговує запит з інтенсивністю μ_j . При цьому у вихідній аналітичній моделі зміниться вираз для визначення інтенсивності переходу зі стану S_k до стану S_{k-1} .

Очевидно, $g_k = \sum_{j=1}^m P_k^{(j)} \mu_j$.

В обчислювальних системах для підвищення продуктивності застосовують кешування, використовуючи притаманну програмам властивість локальності. Вона означає, що дані, які необхідні в процесі обчислень у даний час, найімовірніше можуть знадобитися в найближчому майбутньому. У зв'язку з цим поточні дані, що використовуються на даний момент посередником, записуються в кеш, замінюючи в ній більш «старі» дані. Як показує практика, існує досить висока ймовірність того, що в наступному циклі обчислень посереднику потрібна частина даних поточного й попереднього циклів. У цьому випадку вони можуть бути зчитані з кешу, що, по-перше, дозволяє знизити інтенсивність звертань до ОЗВ, а по-друге, зменшити цикл звертання до ОЗВ тому, що цикл звертання до кешу приблизно на порядок менше циклу звертання до ОЗВ. Ефективність кешу визначається ймовірністю промаху. Під *промахом* мається на увазі випадок, коли потрібного поняття в кеші немає, і його потрібно витягати з ОЗВ. Імовірність промаху звичайно визначається статистично, на підставі вимірів частоти звертань до ОЗВ, або шляхом імітаційного моделювання. У визначальній ступені вона залежить від обсягу ке-

шу. В аналітичній моделі наявність кешування можна врахувати в такий спосіб. Нехай ймовірність звертання до кешу для всіх посередників однакова і дорівнює P_0 , а інтенсивність обслуговування запитів у кеші — μ_0 . СП ОЗВ у цьому випадку є системою з n посередниками і $(m + n)$ бібліотеки, з яких m перших являють собою бібліотеки з інтенсивністю обслуговування μ , а n останніх — кеші з інтенсивністю μ_0 . Елементи p'_{ij} матриці P' розподілу ймовірностей звертання посередників до бібліотек визначаються в такий спосіб:

$$p'_{ij} = \begin{cases} p_{ij}, & \text{якщо } j \leq m; \\ p_0, & \text{якщо } j = m + i, \\ 0 & \text{в інших випадках.} \end{cases}$$

Така система вписується в аналітичну модель із неоднорідними бібліотеками, і для неї застосовні розглянуті вище розрахункові співвідношення.

Ефективним засобом зниження конфліктних утрат продуктивності є використання багатовхідних бібліотек. Багатовхідною називається бібліотека, що допускає одночасний паралельний доступ декількох посередників.

Розглянемо, які зміни необхідно внести до аналітичної моделі, щоб врахувати наявність у СП ОЗВ багатовхідних бібліотек.

Нехай кожен блок багатовхідної бібліотеки допускає одночасне обслуговування s запитів. Очевидно, це приведе до збільшення середньої кількості запитів η_k , що обслуговуються ОЗВ, для кожного стану s_k . Для визначення параметра η_k необхідно розрахувати середнє число запитів $\eta_k^{(j)}$, що обслуговуються в кожній бібліотеці j в стані s_k :

$$\eta_k^{(j)} = \begin{cases} \sum_{d=1}^s dp_{kd}^{(j)} + \sum_{d=s+1}^k sp_{kd}^{(j)}, & \text{якщо } s < k; \\ \sum_{d=1}^k dp_{kd}^{(j)}, & \text{якщо } s \geq k. \end{cases} \quad (10)$$

Ймовірність того, що при заданому наборі посередників Ω_k до j -ї бібліотеки звернулося d посередників, що складають підмножину Ω_{kd} , дорівнює

$$\prod_{\gamma \in \Omega_{kd}} p_{i\gamma j} \prod_{\gamma \in \Omega_k \setminus \Omega_{kd}} (1 - p_{i\gamma j}).$$

Підсумовуючи цей вираз за всіма можливими підмножинами Ω_{kd} з множини Ω_k і за можливими наборами із k посередників, шукану ймовірність $p_{kd}^{(j)}$ можна представити у вигляді алгебраїчної суми термів порядку від d до k , причому терми порядку $d, d + 2, d + 4, \dots$ будуть входити до цієї суми зі знаком «плюс», а терми порядку $l + 1, l + 3, \dots$ — зі знаком «мінус». Терм v -го порядку буде входити C_v^d

раз у підмножину Ω_k . Число таких підмножин, що містять терм v -го порядку, складає C_{n-v}^{k-v} .

Тоді:

$$p_{kd}^{(j)} = \frac{1}{C_n^k} \sum_{v=d}^k (-1)^{d+v} C_v^d C_{n-v}^{k-v} B_{vj}. \quad (11)$$

Підставляючи вираз (11) у (10), можна для кожної j -ї бібліотеки визначити значення параметра $\eta_k^{(j)}$. Тоді середнє число запитів, що обслуговуються в ОЗВ, у стані S_k можна знайти за наступною формулою:

$$\eta_k = \begin{cases} \sum_{j=1}^m \eta_k^{(j)} & \text{для } k \leq l; \\ \eta_l & \text{для } k > l \end{cases}$$

Підставивши отримане значення η_k в розрахункові формули (1)–(6), одержимо значення коефіцієнта втрат продуктивності δ для СП ОЗВ із багатовхідними бібліотеками.

Використання результатів досліджень можливе при створенні та розвитку інформаційної інфраструктури систем підтримки прийняття рішень у міжорганізаційних багаторівневих організаційних структурах управління.

1. *Матов О.Я.* Проблеми горизонтальної інтеграції інформаційних ресурсів у багаторівневих організаційних структурах з динамічною конфігурацією / О.Я. Матов, І.О. Храмова // Реєстрація, зберігання і оброб. даних. — 2007. — Т. 9, № 3. — С. 88–97.

2. Организация вычислительных процессов в АСУ: [Учебное пособие] / [Матов А.Я., Шпилев В.Н., Комов А.Д. и др.]; под ред. А.Я. Матова. — К.: КВИРТУ ПВО, 1989. — 200 с.: ил. — Библиогр.: С. 193–195.

Надійшла до редакції 09.07.2009