

МАТЕМАТИЧНА МОДЕЛЬ ОБРОБЛЕННЯ ІНФОРМАЦІЇ СКЛАДНОЇ ІНФОРМАЦІЙНОЇ СИСТЕМИ

© Бойченко О.В., 2011

Розглядається питання моделювання оброблення інформації складної інформаційної системи для підвищення ефективності інформаційного забезпечення правоохоронних органів. Запропоновано застосування кластеризації обчислювальних вузлів для моделювання планування оброблення інформації в розподіленій гетерогенній обчислювальній системі.

Ключові слова: оброблення інформації, розподілена інформаційна система, модель кластеризації.

The article deals with the matter of modeling of information processing of the sophisticated information system for the improvement of effectiveness of information support of the law enforcement authorities. It was suggested to use the clusterization of computational nodes to simulate the planning of data processing within the distributed heterogeneous computing system.

Key words: information procession, distributed information system, clusterization model.

Постановка задачі

Стрімке впровадження інформаційно-телекомунікаційних технологій управління в усі сфери життєдіяльності суспільства, широке застосування цифрової технології оброблення та обміну даними в практику діяльності організацій та установ, зумовлює необхідність пошуку нових підходів до розроблення алгоритмів функціонування системи інформаційного забезпечення підприємств та закладів.

Найактуальнішого значення зазначена обставина набуває в діяльності органів внутрішніх справ, які функціонують в умовах оброблення значних обсягів даних для ефективного управління підрозділами міліції щодо виконання завдань охорони громадського порядку та протидії злочинності. Тому науковий пошук у напрямку удосконалення способів та методів оброблення даних у складі складної інформаційної системи спеціального призначення є достатньо обґрунтованим.

Дослідженням організаційно-технічних проблем інформаційного забезпечення правоохоронних органів свого часу займалися такі вчені, як Р.А. Калюжний, Я.Ю. Кондратьєв, Л.І. Петрачук, В.А. Саницький, Б.Б. Шавиркін, В.Г. Хахановський [1–5] та інші. Поряд з цим зазначимо, що питанням моделювання оброблення інформації складних розподілених ієрархічних систем спеціального призначення у наукових розробках приділено не достатньо уваги.

Мета роботи полягає у проведенні аналізу наукових підходів до моделювання систем оброблення інформації складної інформаційної системи для оптимізації комплексу заходів інформаційного забезпечення в діяльності органів внутрішніх справ (ОВС) України.

Виклад основного матеріалу

Інформаційні підсистеми, як основні складові частини системи інформаційного забезпечення, призначені для збору, накопичення, зберігання та оброблення інформації певних обліків, які орієнтовані на використання в діяльності багатьох служб. Найбільша ефективність автоматизованих інформаційних систем досягається під час оптимізації планів роботи, швидкого вибору оперативних рішень, чіткого маневрування матеріальними і фінансовими ресурсами тощо.

Найприйнятнішою для забезпечення виконання завдань ОВС є ієрархічна структура інформаційної системи, яка має зіркоподібну топологію (рис. 1) [3, 4, 6].

Описуючи конкретніше інформаційну систему ОВС (рис. 2) [6], потрібно зазначити, що основу системи збирання, контролю та використання інформації становить нижчий (третій) рівень. На цьому рівні забезпечується первинне накопичення інформації, ведення територіальних банків даних, захист інформації, актуалізація інформаційних обліків та передавання інформації до банків даних другого і першого рівнів.

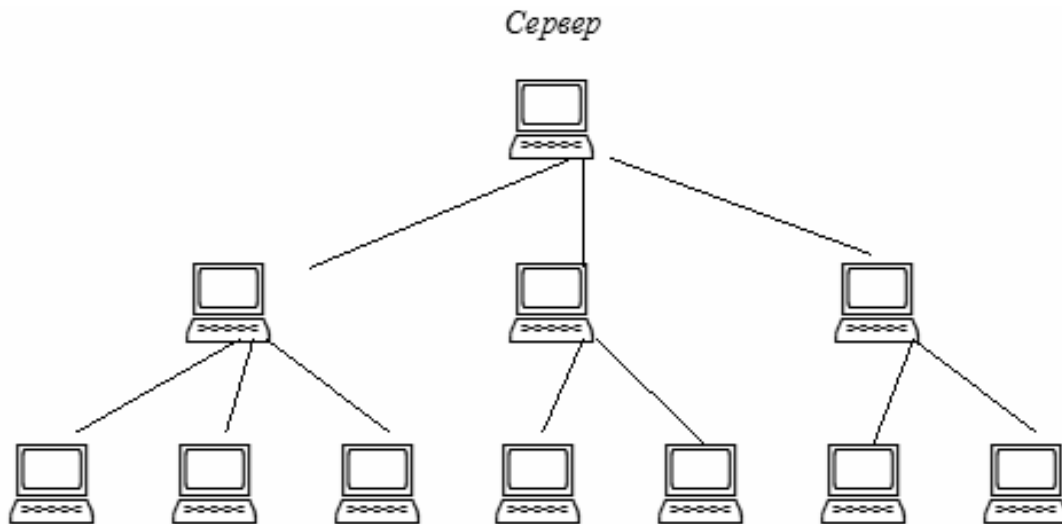


Рис. 1. Ієрархічна структура інформаційної системи ОВС



Рис. 2. Система інформаційного забезпечення ОВС України

На другому рівні формують регіональні банки даних, опрацювують інформацію, забезпечують доступ територіальних підрозділів для авторизованого використання інформації, інформаційний зв'язок із територіальними правоохоронними органами та іншими установами держави. На першому рівні забезпечується інтеграція та опрацювання інформації для формування банків даних інформаційних підсистем, міжвідомчий та міжрегіональний інформаційних зв'язок, управління системою інформаційного забезпечення ОВС та дотримання стратегії її розвитку, розроблення нормативно-правової бази та інформаційних технологій.

Розглядаючи процедуру опрацювання інформації і складних ієрархічних системах ОВС (рис. 3), потрібно зазначити наявність низки проблем, пов'язаних зі значною кількістю хибних інформаційних пакетів, поява яких визначена причинами недосконалої програмного забезпечення, низькою кваліфікацією персоналу тощо.



Рис. 3. Технологія опрацювання інформації в інформаційних підсистемах ОВС (БД – база даних; ЛБД – локальна база даних)

У такому разі для вирішення зазначеної проблеми доцільним є застосування моделі обчислювальної системи та вхідного потоку інформаційних даних (запитів) в складній інформаційній системі, а також алгоритму оптимізації. Тоді ефективність планування опрацювання інформації визначатиметься адекватністю використаних моделей в обчислювальній системі [7].

Наявні моделі обчислювальних систем, які використовуються для планування опрацювання інформації в складних інформаційних системах, можна розглядати з трьох позицій:

- 1) адекватність моделі реальному об'єкту інформатизації;
- 2) ресурсна ємність алгоритму оптимізації у разі використання моделі;
- 3) ресурсна ємність отримання інформації стану.

Особливістю процесу моделювання є те, що під час визначення класу моделі обчислювальної системи для використання в плануванні опрацювання інформації на частину показників накладаються обмеження, а інші оптимізуються. Окремим випадком зазначеного є більшість систем планування опрацювання інформації [8], де обчислювальний вузол моделюється процесом, який моделює інформаційний потік у монопольному режимі, що істотно знижує адекватність моделі.

Тому в умовах застосування складних розподілених ієрархічних інформаційних систем доцільно застосовувати розподілену гетерогенну обчислювальну систему [9], яка обробляє вхідний нестационарний інформаційний потік. Метою функціонування системи планування опрацювання інформації є максимізація задоволення потреб користувачів з своєчасності отримання результатів.

У такому разі цільова функція має вигляд

$$\sum_{i=0}^n \xi^i(t^i_{os}) \rightarrow \max,$$

де n – кількість отриманих запитів за інтервал часу планування.

Виконання завдання оптимізації моделі інформаційної системи для використання з метою планування опрацювання інформації передбачає наявність таких вимог мінімізації:

1) розбіжність вихідних реакцій реальної системи та її моделі під час опрацювання інформації

$$Z = \sum_i |(R(Q_i) - R'(Q_i))| \rightarrow \min, \quad (1)$$

де $R(Q_i)$ – витрати ресурсів усіх видів на оброблення i -го запиту [10];

2) витрати ресурсів на забезпечення актуальності моделі та планування

$$Z = r^p + r^a \rightarrow \min, \quad (2)$$

де r^p – ресурсна ємкість планування; r^a – ресурсна ємкість підтримки актуальності моделі.

Виконання завдання (1), (2) можливе з застосуванням такої моделі обчислювальної системи:

$$S_{ГРОС} = \{S_{OC}, S_K, P\},$$

де S_{OC} – модель обчислювальної системи; S_K – модель комунікаційної системи; P – протоколи взаємодії елементів системи.

Модель обчислювальної системи S_{OC} задається набором груп обчислювальних вузлів G : $S_{OC} = \{G_1, \dots, G_m\}$. Кожна група обчислювальних вузлів призначена для опрацювання інформації одного типу.

Модель обчислювального вузла $V_j \in G_k$ дає змогу описувати як однопроцесорні ЕОМ, так і багатопроцесорні системи, які перебувають під управлінням операційної системи кластерного типу (наприклад, Windows Computer Cluster [11]). Для цього пропонується модель обчислювального вузла такого вигляду:

$$V = \langle n, p, o, C, os, Arh, PO \rangle, \quad (3)$$

де n – кількість процесорів (ядер) обчислювального вузла; p – умовна виробнича потужність з оброблення найприватніших запитів; o – обсяг оперативної пам'яті; $C = \langle k, c, t_p \rangle$ – комунікаційна підсистема (кількість, виробнича потужність і тип мережених інтерфейсів); os – тип операційної системи; Arh – архітектура обчислювальної системи; PO – множина встановленого програмного забезпечення.

Облік виробничої потужності обчислювального вузла на основі запитів, які трапляються найчастіше, дає змогу прогнозувати час завершення оброблення запиту, що мінімізує вираз (1).

Для обліку поточного стану обчислювальних вузлів пропонується обліковувати очікуваний час оброблення запиту (запитів) $t = t_1, \dots, t_s$, завантаженість $\eta = \eta_m, \eta_o, \eta_c$, а також список ідентифікаторів запитів $Q = Q_{UD_1}, \dots, Q_{UD_s}$. Тоді модель обчислювального вузла набуває такого вигляду:

$$V_j = \langle t, \eta, Q \rangle.$$

Для мінімізації ресурсних витрат (2) пропонується визначати вираз (3) тільки один раз під час під'єднання обчислювального вузла до системи. У разі зміни конфігурації вузла в систему планування повинна передаватися тільки зміна параметрів ΔV . Для зменшення ресурсної ємкості планування пропонується об'єднувати обчислювальні вузли, які оброблюють запити одного типу з приблизно однаковою виробничою потужністю в групі обчислювальних вузлів G .

Об'єднувати обчислювальні вузли в групи пропонується на основі близькості вихідних впливів під час оброблення запитів з використанням методу k -середніх [12]. Для цього використовується набір моделей обчислювальних вузлів (3) і його вихідна реакція на оброблення, запиту $R(Q_i)$, яка визначається обсягом витрат ресурсів. Кожна група обчислювальних вузлів визначається моделлю:

$$G = \langle \theta, \bar{p}, \bar{o}, \bar{C}, os, Arh, PO \rangle,$$

де θ – функція прогнозування часу завершення оброблення запиту; \bar{p} , \bar{o} , \bar{C} – середні значення відповідних параметрів.

У результаті кластеризації зменшується кількість параметрів моделі обчислювального вузла:

$$V = \langle n, p, o, C, os, Arh, PO \rangle \Leftrightarrow V = \langle n, k_p, k_o, k_C, G \rangle,$$

де $k_p = \frac{p}{\bar{p}}$, $k_o = \frac{o}{\bar{o}}$, $k_C = \frac{C}{\bar{C}}$.

Отже, модель обчислювального вузла являє собою об'єднання моделей V та V^o :

$$Vp = \{V, V^o\} = \left\{ \langle n, k_p, k_o, k_C, G \rangle, \langle t, \eta, Q \rangle \right\}.$$

Щодо моделі комунікаційної системи, то її зображають у вигляді графа, вершинами якого є обчислювальні вузли, а ребрами – комунікаційні лінії:

$$S_k = (\{V\}, \{C\}).$$

Кожне ребро має вагову характеристику, яка відображає пропускну спроможність та латентність лінії. Протоколи взаємозв'язку P елементів системи визначаються у вигляді набору записів, які визначають правила та механізми взаємозв'язку.

Для перевірки ефективності моделі доцільно застосовувати експериментальний метод порівняння підходів до моделювання за показниками адекватності моделі обчислювальної системи, ресурсної ємкості виконання завдання планування, а також ресурсної ємкості підтримки моделі в актуальному стані.

Зокрема оцінюють адекватність моделі за показником похибки прогнозування часу оброблення запиту

$$\varepsilon = \frac{1}{n} \sum_{i=0}^n \left| \frac{t - t_i}{t_i} \right|.$$

Висновки

Підсумовуючи, потрібно зазначити, що застосування запропонованої моделі дає змогу підвищити адекватність моделі та зменшити її ресурсну ємкість. Зазначене є запорукою підвищення ефективності опрацювання інформації складної інформаційної системи спеціального призначення.

1. Калюжний Р.А. Автоматизовані системи і засоби обчислювальної техніки у організації і функціонуванні органів внутрішніх справ / Калюжний Р.А. – К.: КНУВС, 1990. – 193 с.
2. Кондратьєв Я.Ю. Нормативно-правова база інформаційно-аналітичного забезпечення діяльності оперативних підрозділів міліції / Я. Ю. Кондратьєв, В.Г. Хахановський // Науковий вісник НАВС України, 2003. – № 3, ч. 2. – С. 53–59.
3. Саницький В.А. Комп'ютеризовані інформаційні підсистеми МВС України у боротьбі зі злочинністю: монографія / В.А. Саницький. – К.: АНТЕКС, 1998. – 130 с.
4. Саницький В. А. Система інформаційного забезпечення ОВС України: навчально-практичний посібник [для курс. вищ. навч. закл. МВС] / В.А. Саницький. – К.: АНТЕКС, 2000. – 144 с.
5. Хахановський В.Г. Правові та організаційно – технічні проблеми захисту конфіденційної інформації в автоматизованих системах ОВС України / В.Г. Хахановський // Науковий вісник НАВС України, 2002. – № 1, ч. 1. – С. 39–45.
6. Шавиркін Б.Б. Застосування ЕОМ в ОВС: навчальний посібник [для курс. вищ. навч. закл. МВС] / Б.Б. Шавиркін, Л.І. Петрачук. – Донецьк: ДЮІ, 2007. – 178 с.
7. Топорков В.В. Модели распределенных вычислений: монографія / В.В. Топорков. – М.: Физматлит, 2004. – 185 с.
8. Якововский М.В. Распределенные системы и сети: учеб. пособие [для студ. высш. учебн. завед.] / М.В. Якововский. – М.: МГТУ “Станкин”, 2000. – 275 с.
9. Лебеденко Е. В. Оптимизация модели распределенной вычислительной системы, используемой для планирования обработки запросов / Е.В. Лебеденко, И.В. Логвинов // Информатика и системы управления. – Благовещенск: АГУ, 2009. – № 3(21). – С. 118–124.
10. Alur R. Model-checking for real-time systems / Alur R., Coucorbetis C., Dill D. // Proceedings of LICS 90. – 1990. – P. 414–425.
11. Windows Computer Cluster Server 2003 [Електронний ресурс]. – Режим доступу: <http://www.microsoft.com/rus/win-dowsserver2003/ccs/default.mspx>.
12. Мандель И.Д. Кластерный анализ: монографія / И.Д. Мандель. – М.: Финансы и статистика, 1988. – 174 с.