

УДК 621.391

С.В. Осієвський¹, Ю.С. Долгий¹, К.А. Вельчев²¹Харківський університет Повітряних Сил імені Івана Кожедуба²Командування Повітряних Сил Збройних Сил України, Вінниця

АНАЛІЗ ПІДХОДІВ ЩОДО УПРАВЛІННЯ ПОТОКАМИ ДАНИХ

Проведений аналіз моделей маршрутизації, які використовуються для передачі даних від вузла-джерела повідомлення до вузла, якому це повідомлення повинно бути доставлено.

модель маршрутизації, передача даних, вузол, повідомлення

Вступ

Постановка задачі. На сьогодні інтенсивність обміну інформацією у комп'ютерних мережах постійно зростає, це приводить до появи пікових загрузок у мережі. Для їх виключення необхідно внести обмеження для конкретних вузлів на передаваний об'єм даних. Ця процедура не завжди являється прийнятною, тому що призводить до неефективного використання пропускнуої можливості каналу зв'язку при відсутності пікових загрузок. Виходячи з цього, завдання розробки адаптивних методів управління потоками даних являється актуальним.

Метою статті є аналіз моделей маршрутизації, які використовуються для передачі даних від вузла-джерела повідомлення до вузла, якому це повідомлення повинно бути доставлено.

Викладення основного матеріалу

Алгоритми маршрутизації визначають шлях передачі даних від вузла-джерела повідомлення до вузла, якому це повідомлення повинно бути доставлено. Серед можливих способів рішення даної задачі розрізняють:

- *оптимальні*, що визначають завжди найкоротші шляхи передачі даних і *неоптимальні* алгоритми маршрутизації;

- *детерміновані* і *адаптивні* методи вибору маршрутів (адаптивні алгоритми визначають шляхи передачі даних залежно від існуючого завантаження комунікаційних каналів) [1].

До найбільш поширених оптимальних алгоритмів належить клас *методів покоординатної маршрутизації (dimension-ordered routing)*, в яких пошук шляхів передачі даних здійснюється по черзі для кожної розмірності топології мережі комунікації. Так, для двовимірних ґрат такий підхід призводить до маршрутизації, при якій передача даних спочатку виконується за одним напрямком (наприклад, по горизонталі до досягнення вертикалі процесорів, в якій розташовується процесор призначення), а потім дані передаються вздовж іншого напрямку (дана схема відома під назвою *алгоритму ХУ-маршрутизації*) [3].

Для гіперкуба покоординатна схема маршрутизації може полягати, наприклад, в циклічній передачі даних процесору, визначуваному першою бітовою позицією, що розрізняється, в номерах процесорів, на якому повідомлення розташовується в даний момент часу і на який повідомлення повинно бути передано [2].

В комп'ютерних мережах для організації взаємодії, синхронізації і взаємовиключення паралельно виконуваних процесів використовується передача даних між вузлами обчислювальної мережі. Часові затримки при передачі даних по лініях зв'язку можуть виявитися істотними (в порівнянні з швидкістю процесорів вузлів) і, як результат, *комунікаційна трудомісткість* алгоритму створює істотний вплив на вибір способів рішення мережних задач.

Структура ліній комутації між вузлами обчислювальної системи (*топология мережі передачі даних*) визначається, як правило, з урахуванням можливостей ефективної технічної реалізації. Важливу роль при виборі структури мережі грає і аналіз інтенсивності інформаційних потоків при рішенні найбільш поширених обчислювальних задач. Глобальні і корпоративні мережі, як правило, мають складну структуру, яку можливо описати моделлю гіперкубу (рис. 1).

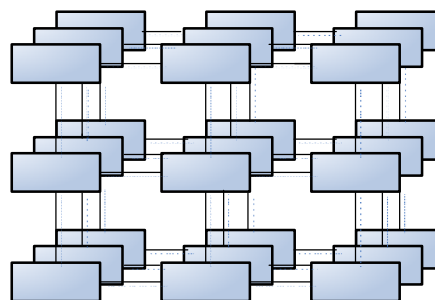


Рис. 1. Структура гіперкубу

Гіперкуб (hypercube) – дана топологія представляє окремий випадок структури ґрат, коли по кожній розмірності сітки є тільки два вузла (тобто гіперкуб містить процесорів при розмірності); даний варіант організації мережі передачі даних досить широко поширений в практиці і характеризується наступним рядом таких ознак:

– два вузла мають з'єднання, якщо двійкові представлення їх номерів мають тільки одну позицію, що розрізняється;

– у N -мірному гіперкубі кожен вузол пов'язаний рівно з сусідами;

– N -мірний гіперкуб може бути розділений на два $(N-1)$ -мірних гіперкуби (всього можливо різного такого розбиття); найкоротший шлях між двома будь-якими процесорами має довжину, співпадаючу з кількістю бітових значень, що розрізняються, в номерах вузлів (дана величина відома як *відстань Хеммінга*).

В процесі обміну даних у великих мережах одночасність виконання декількох комунікаційних операцій є обмеженою – у будь-який момент часу кожен процесор може брати участь тільки в одній операції прийому-передачі даних. Як результат, можуть виконуватися тільки ті комунікаційні операції, в яких взаємодіючі пари вузлів не перетинаються між собою.

Як основні характеристики топології мережі передачі даних найширше використовується наступний ряд показників.

Діаметр – показник, визначуваний як максимальна відстань між двома вузлами мережі (під відстанню звичайно розуміється величина найкоротшого шляху між вузлами); дана величина може характеризувати максимально-необхідний час для передачі даних між вузлами, оскільки час передачі звичайно прямо пропорційний довжині шляху.

Зв'язність (connectivity) – показник, що характеризує наявність різних маршрутів передачі даних між вузлами мережі; конкретний вид даного показника може бути визначений, наприклад, як мінімальна кількість дуг, яку треба видалити для розділення мережі передачі даних на дві незв'язні області.

Ширина бінарного ділення (bisection width) – показник, що визначається як мінімальна кількість дуг, яку треба видалити для розділення мережі передачі даних на дві незв'язні області однакового розміру.

Вартість – показник, який може бути визначений, наприклад, як загальна кількість ліній передачі даних в багатопроцесорній обчислювальній системі.

У табл. 1 наведені характеристики топологій комп'ютерних мереж.

Аналіз результатів показує що найбільш складним по своїй структурі і показникам являється структура мережі повний граф та гіперкуб якими описуються комп'ютерні мережі.

Час передачі даних між процесорами визначає комунікаційну складову (*communication overhead*) тривалості виконання паралельного алгоритму в багатопроцесорній обчислювальній системі. Основний набір параметрів, що описують час передачі даних, складається з наступного ряду величин:

– **час початкової підготовки** (t_n) характеризує тривалість підготовки повідомлення для передачі, пошуку маршруту в мережі і т.п.;

Таблиця 1
Характеристики топологій мережі передачі даних
(ρ – кількість процесорів)

Топологія	Діаметр	Ширина бісекції	Зв'язність	Вартість
Повний граф	1	$\rho^2/4$	$\rho-1$	$\rho(\rho-1)/2$
Зірка	2	1	1	$\rho-1$
Повне двійкове дерево	$2\log((\rho+1)/2)$	1	1	$\rho-1$
Лінійка	$\rho-1$	1	1	$\rho-1$
Кільце	$\rho/2$	2	2	ρ
Решітка $N=2$	$2(\sqrt{\rho}-1)$	$\sqrt{\rho}$	2	$2(\rho-\sqrt{\rho})$
Решітка-тор $N=2$	$2\lfloor\sqrt{\rho/2}\rfloor$	$2\sqrt{\rho}$	4	2ρ
Гіперкуб	$\log_2 \rho$	$\rho/2$	$\log \rho$	$(\rho \log \rho)/2$

– **час передачі службових даних** (t_c) між двома сусідніми процесорами (тобто для процесорів, між якими є фізичний канал передачі даних); до службових даних може відноситися заголовок повідомлення, блок даних для виявлення помилок передачі і т.п.;

– **час передачі одного слова даних** по одному каналу передачі даних (t_k); тривалість подібної передачі визначається смугою пропускання комунікаційних каналів в мережі [5].

До найбільш поширених методів передачі даних належать наступні два основні способи [4]. Перший з них орієнтований на *передачу даних (МПД)* як неподільних (атомарних) блоків інформації (*store-and-forward routing* or *SFR*). При такому підході процесор, що містить повідомлення для передачі, готує весь об'єм даних для передачі, визначає процесор, якому слід направити дані, і запускає операцію пересилки даних. Процесор, якому напрямлено повідомлення, в першу чергу здійснює прийом повністю всіх даних, що пересилаються, і тільки тоді приступає до пересилки прийнятого повідомлення далі по маршруту. Час пересилки даних t_{pd} для методу передачі повідомлення розміром m байт по маршруту завдовжки l визначається виразом

$$t_{pd} = t_n + (t_c + mt_k) \cdot l.$$

При достатньо довгих повідомленнях часом передачі службових даних можна нехтувати і вираз для часу передачі даних може бути записаний в простішому вигляді

$$t_{pd} = t_n + mt_k \cdot l.$$

Другий спосіб ґрунтується на представленні повідомлень, що пересилаються, у вигляді блоків інформації меншого розміру (*пакети*), внаслідок чого передача даних може бути зведена до *передачі пакетів (МПП)*. При такому методі комунікації (*cut-through routing* or *CTR*) приймаючий процесор може здійснювати пересилку даних по подальшому маршруту безпосередньо відразу після прийому чергового пакету, не чекаючи завершення прийому

даних всього повідомлення. Час пересилки даних при використанні методу передачі пакетів визначається виразом [4]

$$t_{\text{пд}} = t_{\text{н}} + mt_{\text{к}} + \ell t_{\text{с}}.$$

Висновки

Порівнюючи наведені вирази, можна відмітити, що для обміну повідомленнями та пакетами необхідно використовувати адаптивний підхід управління потоками даних.

Крім того, даний підхід знижує потребу в пам'яті для зберігання даних, що пересилаються, для організації прийому-передачі повідомлень, а для передачі пакетів можуть використовуватися одночасно різні комунікаційні канали.

Список літератури

1. Олифер В.Г., Олифер Н.А. *Компьютерные сети. Принципы, технологии, протоколы.* – СПб.: Питер, 2006. – 727 с.
2. Гуржий А.Н., Коряк С.Ф., Самсонов В.В., Скляр А.А. *Контроль и управление корпоративными компьютерными сетями.* – К.: СМИТ, 2003. – 194 с.
3. Якубайтис Э.А. *Архитектура вычислительных сетей.* – М.: Статистика, 2000. – 117 с.
4. Бертсекас Д., Галлагер Р. *Сети передачи данных.* – М.: Мир, 2000. – 490 с.
5. Ачясова С.М., Бандман О.Л. *Корректность параллельных вычислительных процессов.* – М.: Наука. – 161 с.

Надійшла до редколегії 19.03.2007

Рецензент: канд. техн. наук, доцент І.В. Рубан, Харківський університет повітряних Сил ім. І. Кожедуба, Харків.