

АНАЛІЗ МЕТОДІВ ПАРАЛЕЛЬНОЇ ОБРОБКИ ДАНИХ

Могильний С.Б., Котик І.В.

*Інститут телекомунікаційних систем КПІ імені Ігоря Сікорського
E-mail: kotik.iv@gmail.com*

ANALYSIS OF PARALLEL DATA PROCESSING METHODS

The current task is to increase the bandwidth of communication channels, including through the parallel processing of data when compressed, encrypted, etc. In this work we explore the possibilities of multi-core data processing by combining, in our case, several microcomputers into a cluster.

Актуальною є задача збільшення пропускної здатності каналів зв'язку в тому числі за рахунок паралельного оброблення даних при їх стисненні, шифруванні тощо. В даній роботі досліджується можливість багатоядерної обробки даних за допомогою об'єднання, в нашому випадку, кількох мікрокомп'ютерів в кластер.

Паралельні обчислення — спосіб організації комп'ютерних обчислень, при якому програми розробляються як набір взаємодіючих обчислювальних процесів, які працюють паралельно і в архітектурі комп'ютерів реалізуються в основному у формі багатоядерних процесорів.

Ідея розпаралелювання обчислень ґрунтується на тому, що громісткі завдання можуть бути розділені на набір менших завдань, які можуть бути вирішені одночасно.

- Задача розбивається на підзадачі, які можуть виконуватися у один і той самий момент часу.

- Кожна підзадача в свою чергу розбивається на послідовність інструкцій.

- Інструкції кожної підзадачі виконуються одночасно на різних процесорах.

- У процесі обчислень використовується загальний механізм контролю-координації.

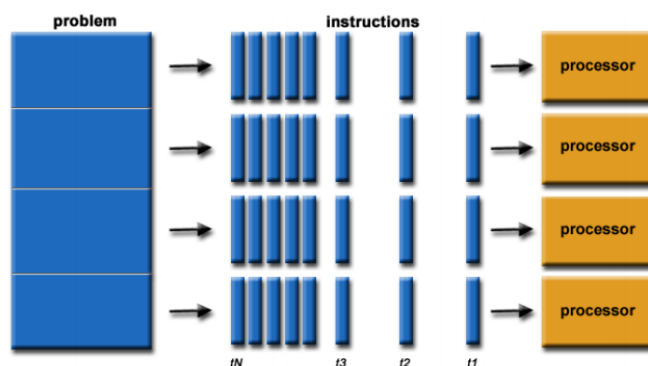


Рис. 1. Схема паралельних обчислень

Обчислювальна задача має:

- Допускати розбиття на незалежні підзадачі, які можна виконувати одночасно.

- З використанням численних обчислювальних ресурсів розв'язуватися за короткий проміжок часу, ніж з використанням одного процесора.

Паралельні обчислення існують таких форм:

- паралельність на рівні інструкцій,
- паралельність даних,
- паралельність завдань.

Паралельність на рівні інструкцій.

Комп'ютерна програма — це потік інструкцій, що виконуються процесором. Даний потік інструкцій можна розділити на групи, які виконуватимуться паралельно, можна змінювати порядок даних інструкцій і це не вплине на результат роботи всієї програми. Такий прийом відомий як паралельність на рівні інструкцій. Просування в розвитку паралельності на рівні інструкцій в архітектурі комп'ютерів відбувалося з середини 1980-х до середини 1990-х.

Паралельність даних.

Основна ідея підходу, заснованого на паралельності даних, полягає в тому, що одна операція виконується відразу над всіма елементами масиву даних. Різні фрагменти такого масиву обробляються на векторному процесорі або на різних процесорах паралельної машини. Розподілом даних між процесорами займається програма. Векторизація або розпаралелювання в цьому випадку найчастіше виконується вже на етапі компіляції – перевід вихідного тексту програми в машинні команди. Роль програміста в цьому випадку зазвичай зводиться до задання налаштувань векторної або паралельної оптимізації компілятора, директив паралельної компіляції, використання спеціалізованих мов для паралельних обчислень.

Паралельність завдань (багатопоточність).

Стиль програмування, заснований на паралельності завдань, має на увазі, що обчислювальне завдання розбивається на декілька відносно самостійних підзадач і кожен процесор завантажується своєю власною підзадачею.

Засоби для здійснення паралельних обчислень.

Засоби, які дають змогу втілити парадигму паралелізму, можна класифікувати наступним чином:

* апаратні

+ засоби для проведення обчислень (обчислювальна техніка)

- обчислювальна техніка, зібрана з стандартних комплектуючих;
- обчислювальна техніка, зібрана з спеціальних комплектуючих;

+ засоби візуалізації;

+ засоби для зберігання і обробки даних;

* програмні

+ програмні засоби загального призначення (операційні системи, стандартні бібліотеки, мови програмування, компілятори, профайлери, дебагери і т.п.);

+ спеціальні програмні засоби: бібліотеки (PVM, MPI); засоби об'єднання ресурсів (Dynamite, Globus та інші).

Паралельні комп'ютери поділяються на:

* Сучасні автономні комп'ютери. Вони є паралельними з точки зору їх внутрішньої будови:

+ Наявність численних функціональних модулів (L1-кеш, L2-кеш, пристрої попереджувального вибору команд (prefetch), decode, floating-point, графічні процесори (GPU), модулі для цілої та дійсної арифметики і т.д.);

+ Багатоядерність чи багатопроцесорність (Multiple execution units/cores);

+ Підтримка використання потоків (Multiple hardware threads)

Будова обчислювального чіпу IBM BG/Q з 18 ядрами (PU) та 16 кеш-модулями 2-го рівня (L2).

* Кластери, які складаються із кількох робочих станцій, з'єднаних мережею (див рис. 2).

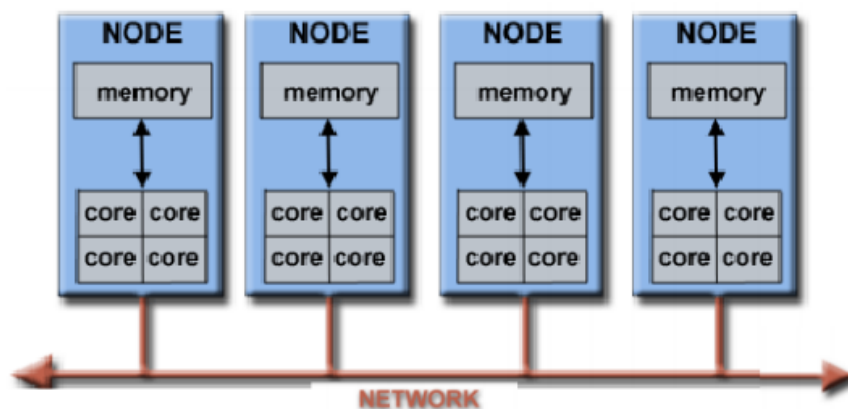


Рис. 2. Схема кластера

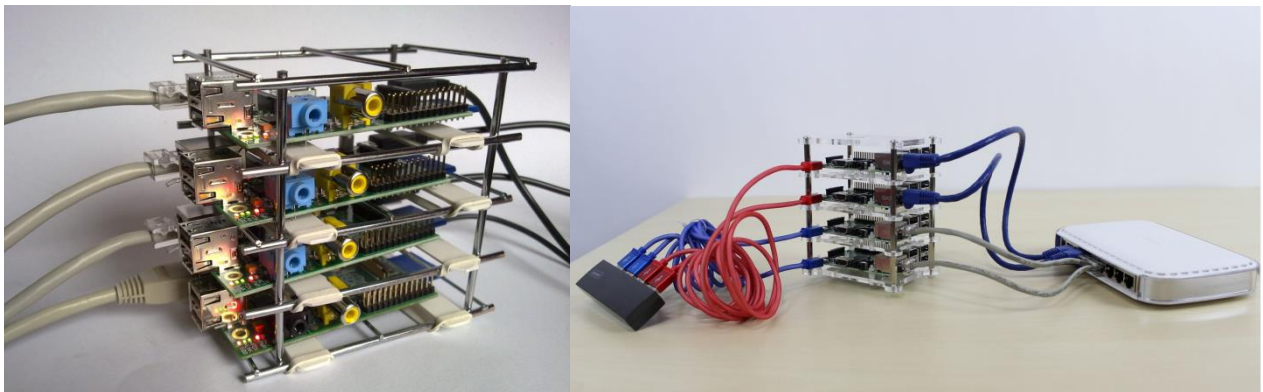


Рис. 3. Кластери з 4-х Raspberry Pi

Висновки. Виконано дослідження дозволяє збільшити ефективність стиснення та шифрування даних при підготовці їх для передачі в каналах зв'язку.

Література

1. Хокні Р., Джессхоуп К. Паралельні ЕОМ. Архітектура, програмування та алгоритми. М.: Радіо і зв'язок. 1986. 392 с.
2. Головін Б.А. Паралельні обчислювальні системи. М.: Наука. 1980. 520 с.