

А. А. Яровий, канд. техн. наук, доц.

ОБРОБКА ЗОБРАЖЕНЬ МЕТОДОМ ПРЯМОГО ПАРАЛЕЛЬНО-ІЄРАРХІЧНОГО ПЕРЕТВОРЕННЯ НА БАЗІ КЛАСТЕРНИХ СИСТЕМ З РОЗПОДІЛЕНОЮ ОБРОБКОЮ ІНФОРМАЦІЇ

Здійснено аналіз методологічних і прикладних особливостей реалізації паралельно-ієрархічного перетворення, пірамідального оброблення інформаційних середовищ на базі кластерних систем з розподіленою обробкою інформації. Наведено результати розробки розподіленої системи управління обчислювальним кластером із використанням обраної технології міжкомпонентної мережевої взаємодії, в якому може виконуватись пряме паралельно-ієрархічне перетворення зображень, що надає можливість забезпечення швидкої обробки у разі збільшення обсягів розмірності зображень.

Вступ

Проблема обробки, розпізнавання, ідентифікації зображень та виявлення їх автентичності із постійним збільшенням обсягів інформаційних потоків та підвищенням вимог до динаміки їх оброблення в реальному часі постає особливо гостро. Із зростанням розмірності зображень, які необхідно ідентифікувати, обчислювальна складність зростає експоненційно. Саме тому останнім часом щораз більше актуалізуються паралельні та розподілені обчислення, зокрема, GPGPU (General-Purpose Computations on Graphical Processing Unit — обчислення загального призначення на відео-адаптері), кластерні, хмаринкові (Cloud computing) та GRID-обчислення [1—4].

В раніше опублікованих роботах неодноразово було відзначено, що паралельно-ієрархічне перетворення в поєднанні з можливостями інформаційної оптоелектроніки відкриває широкі перспективи поліпшення функціонування обчислювальних систем новаторської архітектури, відомих за назвою систем паралельної обробки. Паралельно-ієрархічне перетворення застосовується для виділення характерних ознак зображень, їх кодування і скорочення розмірності під час виконання обчислень. Добра збіжність паралельно-ієрархічного перетворення використовується в структурах паралельної пам'яті, аналізу і розпізнавання зображень, під час кодування і ущільнення даних [5, 6].

Аналізуючи останні тенденції щодо широкого впровадження паралелізму обробки інформації, можна зробити висновок, що ця тематика є надзвичайно актуальною та затребуваною. Поєднання інноваційного паралельно-ієрархічного методу обробки зображень та розподілених кластерних обчислень можна розцінювати як логічний та цілком доцільний крок, який дозволить виконувати алгоритми паралельно-ієрархічного перетворення за прийнятний для користувача час [4, 7].

Метою роботи є створення розподіленої системи управління обчислювальним кластером, в якій може виконуватись пряме паралельно-ієрархічне перетворення зображень для покращення часу обробки та збільшення обсягів розмірності зображень.

Об'єктом роботи є процеси розпаралелювання задач з великою обчислювальною складністю на кластерних системах.

Предметом роботи є системи розподіленої інтелектуальної обробки інформації.

Постановка проблеми

В останні роки визначилась чітка тенденція до росту обсягів перетворення, обробки, передачі і збереження відеоінформації. На основі цього актуальною виявилась розробка ефективних методів і способів інтеграції пристроїв візуалізації та відображення інформації з підсистемами попередньої обробки, процесорних структур для введення і перетворення відеоінформації, інформаційно-пошукових систем, які працюють у «картинних» та образних категоріях.

У загальному вигляді концепцію багатоетапності стосовно обробки зображень можна сформулювати в виконанні такої послідовності основних етапів. Аналіз зображення полягає в послідовному перетворенні збіжних і виявленні (фільтрації) незбіжних у часі складових зображення у разі

переходу елементів нейроподібної паралельно-ієрархічної мережі з поточних енергетичних станів із одними просторовими координатами в стани з меншою енергією з іншими просторовими координатами. Такий процес аналізу зображення відбувається на багатьох етапах, кожний із яких включає виконання вищевказаної процедури. Умовою переходу складового зображення на вищий рівень є наявність динаміки взаємного збігу проміжних результатів обробки в часі в паралельних каналах нижнього рівня. Результат аналізу зображення формується з ізольованих у просторово-часовій області складових зображення.

Необхідно зазначити, що алгоритм обробки зображень за допомогою паралельно-ієрархічного перетворення дозволяє обробляти зображення частинами, вікнами певної розмірності. При цьому різні вікна обробляються незалежно, що дозволяє виконувати таке перетворення паралельно. У проведеному дослідженні для подальших обчислень вирішено використовувати обчислювальний кластер — групу комп'ютерів, об'єднаних високошвидкісними каналами зв'язку, яка є з точки зору користувача єдиним апаратним ресурсом. Кластери досить давно та ефективно використовуються в обчислювальних цілях, зокрема в наукових дослідженнях. Для обчислювальних кластерів істотними показниками є висока продуктивність процесора в операціях над числами з плаваючою крапкою (flops) та низька латентність мережі, і менш істотними — швидкість операцій введення—виведення, яка більшою мірою важлива для баз даних і веб-сервісів. Обчислювальні кластери дозволяють зменшити час розрахунків в порівнянні з одиничним комп'ютером, розбиваючи завдання на гілки, що виконуються паралельно, які обмінюються даними по локальній мережі. Особливої ж актуальності для задач організації паралельних обчислень великих масивів динамічних даних останнім часом набувають високопродуктивні кластери (high-performance computing cluster) [8].

Сформулюємо основну задачу обробки зображень методами паралельно-ієрархічного перетворення на базі кластерних систем. Головна проблема ідентифікації зображень полягає у складності визначення певної унікальної характеристики для зображення. Обчислення такої характеристики може забезпечити, наприклад, функція хешування. Проте у хеш-функцій є значний недолік — незворотність дії: із обчисленого значення хеш-характеристики зазвичай неможливо відновити початкові дані. Такий недолік усуває використання для обчислення хеш-характеристики пряме паралельно-ієрархічне перетворення, оскільки однією із важливих його властивостей є можливість відновлення початкових даних (оборотність).

Отже, основною задачею оброблення зображень методами паралельно-ієрархічного перетворення на базі кластерних систем є виконання над вхідним зображенням прямого паралельно-ієрархічного перетворення, в результаті якого отримується набір хвостових елементів, який однозначно характеризуватиме дане зображення. Пряме паралельно-ієрархічне перетворення повинно виконуватись на кластерній системі довільної топології [9]. Вхідними даними у системі є цифрова інформація, яку можна подати двовимірною матрицею, — зокрема, зображення надвеликої розмірності. Вихідними даними є набір хвостових елементів — результат прямого паралельно-ієрархічного перетворення, — який однозначно характеризує вхідний набір даних.

Характеристика підходу пірамідального оброблення інформаційних середовищ

У разі проблемно-орієнтованого сприйняття використовується не взаємно-однозначне, а багато-однозначне відображення інформаційного поля (ІП) у поле сприйняття, тобто інформація, яка аналізується системою, є узагальненням реальної ситуації. Прикладами обмежених внутрішніх представлень у системах цього класу є контурні препарати зображення, обмежений набір коефіцієнтів двовимірного перетворення тощо [5].

Застосований підхід [5, 6] пірамідального оброблення інформаційного середовища для організації паралельно-ієрархічного перетворення полягає в тому, що, починаючи з вихідних елементів ІП $x_0(i, j)$, які підлягають кодуванню, визначається елемент ІП із мінімальною розмірністю $x_1(i, j)$. В термінах обробки зображень це не що інше, як «низькочастотна версія зображення». Ця «низькочастотна версія зображення» з частотою зрізу f_1 може розглядатися як прогнозування для $x_0(i, j)$. Далі описана вище операція ітеративним чином повторюється. Тобто з $x_1(i, j)$ вибирається елемент ІП із мінімальним розміром $x_2(i, j)$. Після n ітерацій утворюється послідовність поточних ІП, обумовлених виразом $x_{i-1}(i, j) - x_1(i, j)$. На кожній ітерації розмірність поточних ІП

скорочується в кількість раз, що рівна $\frac{n_i}{n_{i+1}}$, де n_i та n_{i+1} — відповідно кількість ненульових елементів П. Якщо ці поточні П розглядати як поставлені одне на інше, то в результаті утвориться пірамідальна структура даних [5, 10, 11]. Пірамідальна структура обробки зображення на будь-якому рівні цієї структури являє собою результат згортки двох гаусоподібних функцій із вихідним зображенням. Зображення гарної якості у разі пірамідального кодування утворюється за середнього коефіцієнта ущільнення близько 10. Якщо встановити вищі коефіцієнти ущільнення (до 70), то виникають значні перекошування від послідовних децимацій [5, 12]. Дані перекошування можна усунути за рахунок збільшення обсягів обчислень. Таким чином, П за пірамідального кодування описується упорядкованою послідовністю елементів П, що розташовуються одне над іншим. У разі такого подання П розбивається на рівні блоки, потім процедура розбиття повторюється для кожного блока доти, поки його розмір не стане рівним розміру елемента початкових П. Регулярність пірамідальних структур визначає їх зручну реалізацію й ефективне використання — з одного боку, а з іншого боку — ієрархія описів різноманітного ступеня спільності сприяє контекстно-незалежній структуризації даних. Ця структуризація може бути основою для моделювання індуктивних і дедуктивних процесів сприйняття інформації людиною. На відміну від методів на основі локальних операторів, контурно-текстурні методи [5, 12] зводяться до розчленовування зображення на текстурні області, оточені контурами, що кодують окремо. Після процедур сегментації і попередньої обробки області, яка підлягає виділенню, характеризуються деякими властивостями. Від вибору властивості, наприклад, текстурні ознаки Хараліка [5, 13], багато в чому залежить складність методу і точність контурів, отриманих після сегментації. Перевагою методу є суттєве ущільнення даних. Слабкою ж стороною є, використовувана при створенні областей, характерна властивість, яка призводить в результаті до досить великої кількості областей, чий межі необов'язково є контурами об'єктів на зображенні. Для усунення зазначених недоліків застосовують кодування на основі розкладання за напрямками [5, 12], у якому основна увага приділяється виявленню контурів із кращим зберіганням контурної інформації. Тут контурний елемент визначається у вигляді двовимірної східчастої функції заданої довжини і заданого напрямку. Множина контурних елементів, що забезпечує це подання, складає контурну множину.

Одним із основних контурно-текстурних методів є метод кодування на основі створення областей [5, 14]. На першому етапі цього методу елементи зображення розділяються на контурні і текстурні. Ця процедура розбиває зображення на множину областей із тією умовою, щоб зміна рівня градацій *Grayscale* всередині області не давала різких розривів, тобто контурів. Сегментація здійснюється в три рівня: попередня обробка, створення областей і усунення артефактів.

Програмна реалізація прямого паралельно-ієрархічного перетворення на базі кластерних систем для задач обробки зображень

У результаті проведених раніше досліджень [15, 16] було обґрунтовано вибір програмно-апаратної платформи для реалізації паралельно-ієрархічного мережного перетворення, а також проведено тестування та доведено ефективність обраної програмно-апаратної платформи для обчислювальних задач, які базуються на моделі паралельно-ієрархічного мережного перетворення [7, 9, 16—20]. Розглянемо коротко алгоритм прямого паралельно-ієрархічного перетворення, у результаті якого для вхідної матриці A розмірністю $m \times n$ отримується унікальний вектор, який однозначно ідентифікує цю матрицю [5, 6].

Перший етап:

1. Для кожного рядка матриці A виконується його сортування за зростанням, після чого виконується операція *unique* ($A[i]$), результатом якої є два вектори: вектор унікальних значень, видобутих із $A[i]$, та вектор кількості входжень кожного унікального числа у початковий рядок $A[i]$.
2. Над отриманими векторами виконується G -перетворення:

$$G[i][j] = (\text{unique.First}[j] - \text{unique.First}[j - 1]) \cdot (n - \text{unique.Second}[j - 1]), \quad (1)$$

де *unique.First* — вектор унікальних значень; *unique.Second* — вектор кількості входжень кожного унікального числа у початковий масив; n — розмір початкового масиву; i — номер поточного рядка матриці A ; j — номер поточного елемента.

3. Отриманий двомірний масив розмірністю $m \times n$ є вхідними даними для наступного кроку обробки.

На кроці i ($i > 1$):

— Вхідна матриця транспонується.

— Над кожним рядком транспонованої матриці виконується набір дій, ідентичних описаним у п. 1.

— В отриманій у результаті п. 2 матриці кожний рядок, починаючи з другого, зсувається на один елемент відносно попереднього:

$$\begin{pmatrix} a & a & a \\ a & a & a \\ a & a & a \end{pmatrix} \rightarrow \begin{pmatrix} a & a & a & X & X \\ X & a & a & a & X \\ X & X & a & a & a \end{pmatrix}. \quad (2)$$

— Елемент a_{11} отриманої матриці запам'ятовується як перший хвостовий елемент паралельно-ієрархічного перетворення, перший рядок отриманої у результаті п. 3 матриці видаляється.

— Отримана у результаті матриця є вхідними даними для наступного кроку.

Перетворення закінчується, коли розмірність вхідної матриці досягає 1×1 ; це значення запам'ятовується як останній хвостовий елемент.

Схема алгоритму прямого паралельно-ієрархічного перетворення зображена на рис. 1.

Загальний алгоритм функціонування програмного комплексу реалізації прямого паралельно-ієрархічного перетворення на базі кластерних систем для задач обробки зображень повинен бути таким:

1) користувач завантажує вхідне зображення, яке необхідно обробити, а також вказує множину вузлів у локальній мережі, на яких необхідно проводити обчислення;

2) після ініціалізації початку обчислень комплекс автоматично формує пакет завдань для кожного обчислювального вузла та запускає їх на виконання;

3) після завершення обчислень комплекс збирає результати окремих вузлів на головному (користувацькому) сервері та демонструє результати користувачу.

Для реалізації цього алгоритму розіб'ємо структуру обчислювального комплексу на окремі складові:

1) консоль керування — виконує управління роботою кластеру;

2) хост-програма — запускається на кожному обчислювальному вузлі та забезпечує таку функціональність:

а) отримання набору завдань по локальній мережі;

б) виконання (над кожним набором даних із пакету) завдань, які базуються на моделі прямого паралельно-ієрархічного перетворення;

в) формування набору векторів хвостових елементів та передача його на вузол-ініціатор обчислень;

3) Програмна бібліотека, яка забезпечуватиме мережеву взаємодію між вищевказаними компонентами.

Класову діаграму серверу, який надає послуги щодо доступу до сервісу паралельно-ієрархічного перетворення, зображено на рис. 2, UML-діаграма описаних класів показана на рис. 3.



Рис. 1. Узагальнена схема алгоритму прямого паралельно-ієрархічного перетворення

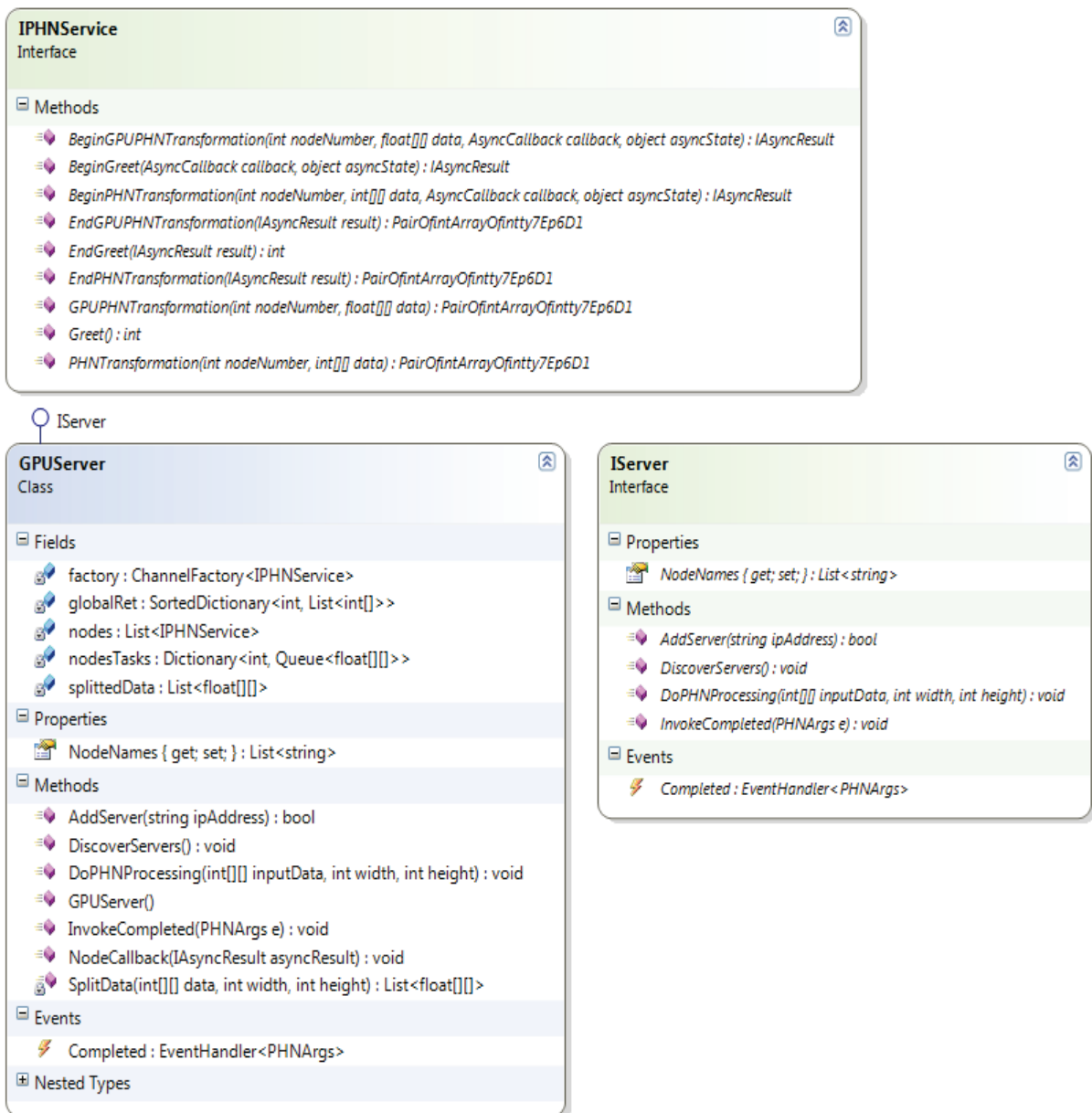


Рис. 2. Класова діаграма сервера, який надає послуги щодо доступу до сервісу паралельно-ієрархічного перетворення

Протестуємо розроблений пакет програм на обчислювальному кластері. Для тестового прикладу виконаємо пряме паралельно-ієрархічне перетворення на обчислювальному кластері із двох обчислювальних вузлів. У якості вхідного зображення оберемо зображення розмірністю 1024×1024 пікселів, заповнене радіальним градієнтом у градаціях сірого, що відповідає ідеалізованому зображенню профілю лазерного пучка.

Порівняємо виконання обчислень на фізичному кластері та на віртуальному кластері, запущеному на одному фізичному комп'ютері; кожен кластер складається із двох вузлів.

На рис. 5 наведено результати прямого паралельно-ієрархічного перетворення, виконаного на віртуальному кластері.

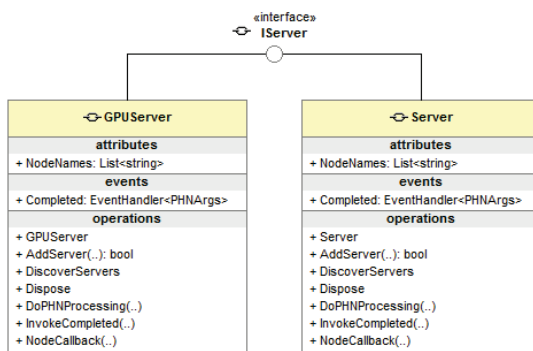


Рис. 3. UML-діаграма серверних класів

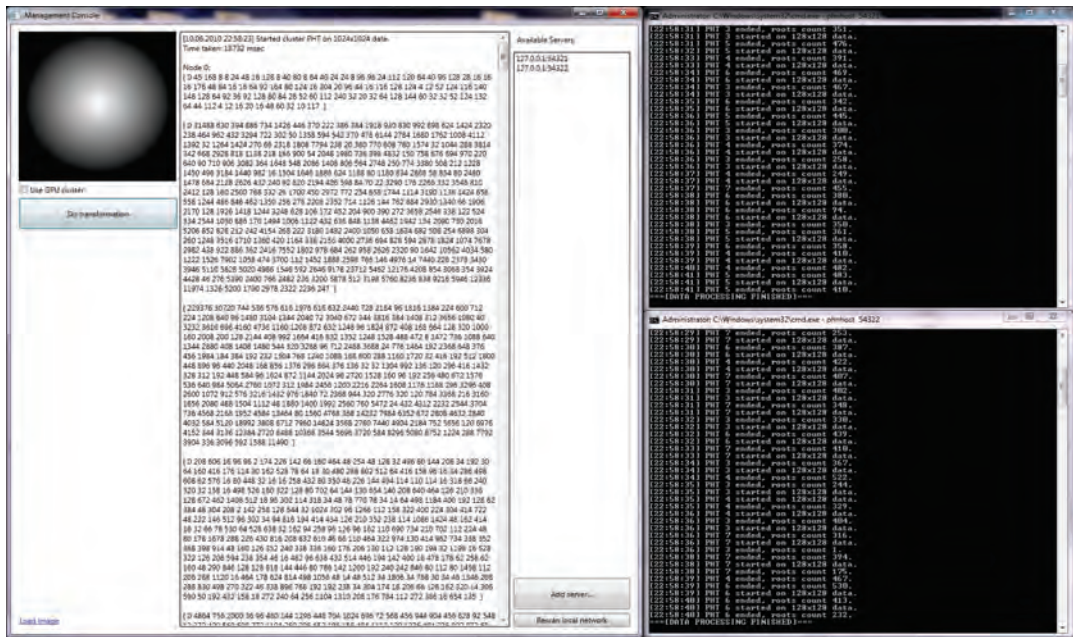


Рис. 5. Емуляція обчислювального кластера із двох обчислювальних вузлів

На рис. 6 наведено виконання прямого паралельно-ієрархічного перетворення на фізичному кластері із двох вузлів. Кожен вузол обробляв власний набір завдань, використовуючи 100 % процесорного часу.

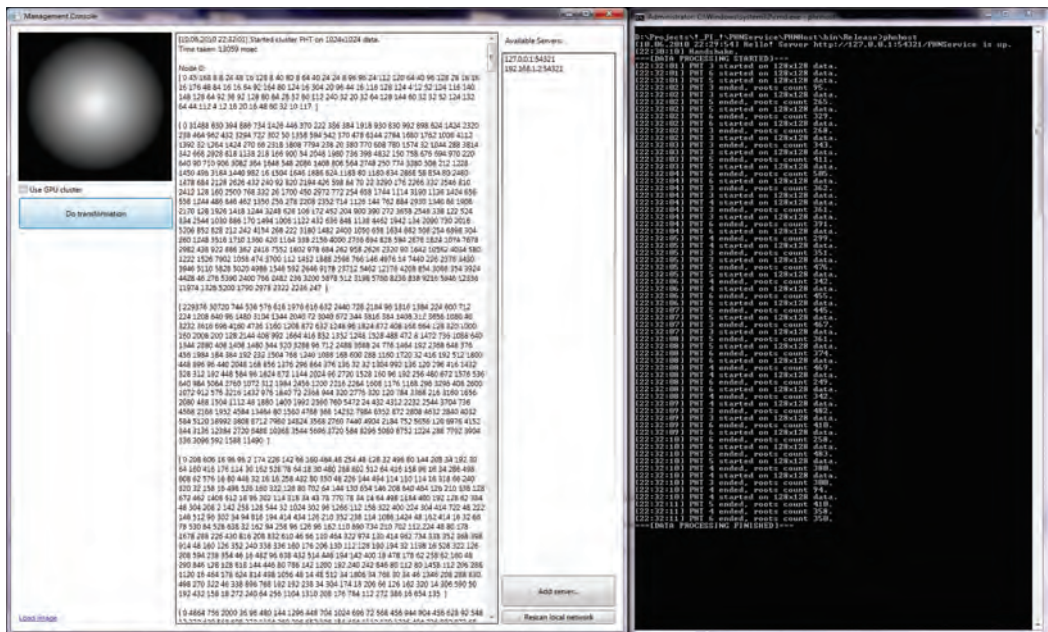


Рис. 6. Виконання прямого паралельно-ієрархічного перетворення на кластері із двох вузлів

В результаті порівняння отриманих результатів обробки констатуємо, що приріст швидкодії відносно обробки на одному вузлі склав 27,7 %.

Висновки

Наведено основні результати розробки розподіленої системи управління обчислювальним кластером із використанням обраної технології міжкомпонентної мережевої асемодії, в якому може виконуватись пряме паралельно-ієрархічне перетворення зображень, що дозволяє зменшення тривалості обробки та можливості забезпечення швидкої обробки при збільшенні обсягів розмірності зображень.

До переваг розробленого пакета програм для реалізації прямого паралельно-ієрархічного пере-

творення на базі кластерних систем для задач обробки зображень необхідно віднести: високу швидкість через використання останніх досягнень у сфері паралелізації; можливість роботи із зображеннями надвеликої розмірності; легкість масштабування у випадку зміни кількості вузлів у кластері; автоматичне сканування локальної мережі на предмет виявлення обчислювальних вузлів. Розроблений пакет програм протестовано на обчислювальному кластері Microsoft IT Academy ВНТУ. Результати тестування дозволяють зробити висновок про правильність та оптимальність проектування. Ці дослідження мають гарні перспективи розвитку. Зокрема, на основі реалізованої програмної бібліотеки, яка здійснює міжкомпонентну взаємодію, можна легко реалізувати потужну обчислювальну парадигму під назвою MapReduce.

Математична модель прямого паралельно-ієрархічного перетворення в порівнянні з відомими числовими методами перетворення (наприклад, розкладання в математичні ряди), простими операціями типу додавання забезпечує складну функціональну обробку сигналів у реальному масштабі часу, а також однозначність і зворотність із гарною збіжністю обчислювального процесу. Причому мережний метод обробки приводить до швидкого ущільнення вхідних масивів інформації.

Галузі застосування мережного методу паралельно-ієрархічного перетворення різноманітні. Це мережні структури паралельної пам'яті, цифрових систем прийому-передачі інформації, цифрових пристроїв ущільнення інформації, систем порівняння багатоградацийних зображень, у тому числі кореляційного порівняння, цифрових пристроїв попередньої обробки зображень, сегментації, кодування, формування ознак для розпізнавання.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Воеводин В. В. Параллельные вычисления / В. В. Воеводин, Вл. В. Воеводин. — СПб. : БХВ-Петербург, 2002. — 608 с.
2. Бройнль Т. Паралельне програмування : навч. посіб. / Бройнль Т. ; пер. з нім. В. А. Святого. — К. : Вища школа, 1997. — 358 с.
3. B. Furht Handbook of Cloud Computing / Borko Furht, Armando Escalante. — New York, Springer Science+Business Media LLC, 2010. — 634 p.
4. Яровий А. А. Методологічні особливості побудови паралельно-ієрархічних та ієрарх-ієрархічних мереж на основі кластерних систем з розподіленою обробкою інформації / А. А. Яровий // Оптико-електронні інформаційно-енергетичні технології. — 2010. — № 1 (19). — С. 69—79.
5. Паралельно-ієрархічне перетворення як системна модель оптико-електронних засобів штучного інтелекту : моног. / [В. П. Кожем'яко, Ю. Ф. Кутаєв, С. В. Свечніков та ін.]. — Вінниця : УНІВЕРСУМ-Вінниця, 2003. — 324 с.
6. Кожем'яко В. П. Паралельно-ієрархічні мережі як структурно-функціональний базис для побудови спеціалізованих моделей образного комп'ютера : моног. / В. П. Кожем'яко, Л. І. Тимченко, А. А. Яровий. — Вінниця : Універсум-Вінниця, 2005. — 161 с. — ISBN 966-641-142-3.
7. Яровий А. А. Прикладні аспекти і перспективи побудови кластерів на основі GPU для реалізації паралельної та паралельно-ієрархічної обробки інформації / А. А. Яровий // Оптико-електронні інформаційно-енергетичні технології. — 2009. — № 2 (18). — С. 119—126.
8. Хьюз К. Параллельное и распределенное программирование на C++. / Хьюз К., Хьюз Т. ; пер. с англ. — М. : ИД «Вильямс», 2004. — 672 с.
9. Яровий А. А. Прикладні аспекти реалізації паралельно-ієрархічного перетворення інформаційних середовищ на основі гетерогенної кластерної системи / А. А. Яровий, Ю. С. Богомолів. // «ІНТЕРНЕТ-ОСВІТА-НАУКА-2010» : зб. мат. VII Міжнародної конференції, 28 вересня — 3 жовтня 2010 року. — Вінниця : ВНТУ, 2010 — С. 243—247.
10. Вступ в алгоритмічну теорію ієрархії і паралелізму нейроподібних обчислювальних середовищ та її застосування до перетворення зображень. Основи теорії пірамідално сільового перетворення зображень / [Кожем'яко В. П., Тимченко Л. І., Кутаєв Ю. Ф., Івасюк І. Д.]. — К. : УМК ВО, 1994. — 272 с.
11. Иерархическая обработка изображений и пирамидальные системы / [Воробьев К. Ю., Тимонькин Г. Н., Харченко В. С., Мельников В. А.] // Зарубежная радиоэлектроника. — 1991. — № 7. — С. 51—61.
12. Мусман С. Г. Достижения в области кодирования изображений / Мусман С. Г., Пирш П., Граллерт Х.-Й. // ТИИЭР. — 1985. — Т. 73, № 4. — С. 23—54.
13. Хорн Б. К. Зрение роботов / Хорн Б. К. — М. : Мир, 1989. — 487 с.
14. Прэтт У. Цифровая обработка изображений / У. Прэтт — (в 2-х книгах). — Москва, «Мир», 1982. — Т. 1. — 310 с., Т. 2. — 790 с.
15. Яровой А. А. Прикладные аспекты программно-аппаратной реализации нейроподобных параллельно-иерархических систем / А. А. Яровой // Сборник научных трудов. [Научная сессия МИФИ — 2009], [XI Всероссийская научно-техническая конференция «Нейроинформатика-2009»], (Москва, 27—30 января 2009 г.), [В 2 частях. Ч. 2.] — М. : МИФИ, 2009. — С. 39—48.
16. Яровий А. А. Прикладна реалізація масштабних нейронних та нейроподібних паралельно-ієрархічних мереж на основі технологій GPGPU [Електронний ресурс] / А. А. Яровий, Ю. С. Богомолів, К. Ю. Вознесенський // Наукові праці Вінницького національного технічного університету. — 2009. — № 2. — С. 1—8. — Режим доступу до журналу : http://www.nbu.gov.ua/e-journals/VNTU/2009_2_ua/2009-2.files/uk/09aayogt_ua.pdf.
17. Свідоцтво про реєстрацію авторського права на твір № 33758. Комп'ютерна програма для реалізації паралельно-

ієрархічного перетворення інформаційних середовищ на основі CPU (Direct PHT::CPU) / Яровий А. А., Богомолів Ю. С. Дата реєстрації Державним Департаментом інтелектуальної власності України 17.06.2010.

18. Свідоцтво про реєстрацію авторського права на твір № 33759. Комп'ютерна програма для реалізації паралельно-ієрархічного перетворення інформаційних середовищ на основі технологій GPGPU (Direct PHT::GPU) / Яровий А. А., Богомолів Ю. С. Дата реєстрації Державним Департаментом інтелектуальної власності України 17.06.2010.

19. Свідоцтво про реєстрацію авторського права на твір № 34123. Комп'ютерна програма «Програмний комплекс для реалізації паралельно-ієрархічного перетворення інформаційних середовищ на основі CPU-орієнтованої кластерної платформи (DirectPHT::Cluster)» / Яровий А. А., Богомолів Ю. С. Дата реєстрації Державним Департаментом інтелектуальної власності України 14.07.2010.

20. Свідоцтво про реєстрацію авторського права на твір № 34122. Комп'ютерна програма «Програмний комплекс для реалізації паралельно-ієрархічного перетворення інформаційних середовищ на основі GPU-орієнтованої кластерної платформи (DirectPHT::GPU-Cluster)» / Яровий А. А., Богомолів Ю. С. Дата реєстрації Державним Департаментом інтелектуальної власності України 14.07.2010.

Рекомендована кафедрою комп'ютерних наук

Стаття надійшла до редакції 23.02.11
Рекомендована до друку 25.03.11

Яровий Андрій Анатолійович — доцент кафедри комп'ютерних наук,
Вінницький національний технічний університет, Вінниця