

Виконання завдань розподіленої обробки зображення під управлінням системи CONDOR
Шийка Ю.Я., Шувар Р.Я.

*Львівський національний університет імені Івана Франка,
факультет електроніки, e-mail: shiyka@mail.lviv.ua*

The peculiarities are considered of implementation of distributed computing system under the management of Condor. The project is described of distributed segmentation of images taken by remote sensing system.

Технології паралельних та розподілених обчислень дозволяють виконувати ресурсозатратні завдання та обробляти великі обсяги даних, використовуючи обчислювальні можливості комп'ютерів об'єднаних в обчислювальну мережу (грід, кластер). Такі обчислювальні мережі мають різну топологію і інфраструктуру, залежно від завдань і вимог, які до них ставляться. Грід або грід-інфраструктура – це розподілена програмно-апаратна комп'ютерна мережа з особливою організацією обчислень і управління потоками завдань і даних.

В лабораторії комп'ютерних інформаційних технологій для розподілених обчислень впроваджено систему [Condor](#) під управлінням кластерної системи [Rocks](#). Виконання незалежних підзавдань відбувається в кластерній комп'ютерній мережі, що складається з 14 двохядерних обчислювальних вузлів (нод) та вузла управління. Система Condor дає можливість створювати черги завдань, які виконуються відповідно до заданих правил у строго визначеному чи невизначеному порядку на тих чи інших вузлах кластера.

Розподілені обчислення полягають в обробці незалежних між собою підзавдань для отримання одного кінцевого результату. Розроблено методику розподіленої сегментації растрових зображень ДЗЗ. Згідно цієї методики обробка зображень здійснюється в чотири етапи.

Перший етап – формування незалежних підзавдань. Створено програму генератор незалежних підзавдань, яка велике вхідне зображення нарізає за принципом мозаїки на менші зображення підзавдань; крім того до кожного підзавдання створюється файл з інформацією про розміри і позицію зображення підзавдання у вхідному зображенні а також параметри для виконавчої програми.

Другий етап – обробка незалежних завдань виконавчою програмою в розподіленому режимі. Створена адаптована до розподілених обчислень виконавча програма виконується на обчислювальних нодах. Ця програма за двома вхідними файлами створює два вихідних: перший файл – сегментоване зображення, другий файл – інформація про розміри і позицію зображення у великому вхідному зображенні. Для того, щоб додати набір підзавдань в чергу завдань Condor створено файл опису завдань Condor, в якому вказуються вибрані параметрами та середовище виконання зав-

дань, шлях та вигляд вхідних та вихідних файлів, параметри доступу і пересилання файлів між обчислювальними вузлами і т.п.

Третій етап – формування кінцевого результату після виконання всіх підзавдань. Для цього створено програму, яка утворює результуюче зображення шляхом вставлення зображень результатів підзавдань на відповідні місця у результуючому великому зображенні.

Четвертий етап – постпроцесінг. Після склеювання незалежно сегментованих зображень на границях зображень утворюються зайві границі розділу сегментів. Для того, щоб отримати кінцеве зображення сегменти розділені на краях потрібно об'єднати. Створено програму постпроцесінгу, яка здійснює кінцеву обробку сегментів у результуючому зображенні за інформацією про властивості та характеристики сегментів утворених в результаті сегментації незалежних підзавдань.

Таким чином, в результаті роботи створено систему розподіленої сегментації растрових зображень. Ця система дозволяє здійснювати обробку великих об'ємів даних, що є неможливо на окремих комп'ютерах. Шляхом організації розподілених обчислень обчислювальні потужності окремих вузлів об'єднуються для виконання одної задачі. При виконанні завдань під управлінням системи Condor використовуються тільки вільні ресурси обчислювальних вузлів, що не перешкоджає їх звичайній роботі в складі кластера паралельних обчислень. Впроваджена система є масштабованою і гнучкою. Система дозволяє виконувати різні завдання розподілених обчислень, змінювати параметри і порядок виконуваних завдань. Зміна кількості обчислювальних нод, топології та інфраструктури не впливає на працездатність проектів розподілених обчислень. Існуючі проекти легко налаштовуються відповідно до змін системи.

Отриману кластерну систему та розроблений проект розподілених обчислень використано для обробки даних дистанційного зондування території Шацького національного природного парку.