

ОПТИМІЗАЦІЯ НАДАННЯ ОБЧИСЛЮВАЛЬНИХ РЕСУРСІВ ПРИ ВИРІШЕННІ ЗАДАЧ ПАРАЛЕЛЬНОЇ ОБРОБКИ ДАНИХ

Анотація. Розглядаються задачі паралельної обробки даних, які вирішуються в середовищі віртуальних обчислювальних машин, що надаються провайдерами хмарних технологій. Враховуючи вимоги користувачів щодо продуктивності, вартості, тривалості, а також надійності використання відповідних ресурсів можна розглянути задачу оптимізації кількості машин з розподілом задач, передачі обсягів обробки по каналам зв'язку, узгодження рішень.

Ключові слова: паралельна обробка, хмарні технології, віртуальні обчислювальні машини, оптимізація обчислень.

Постановка проблеми. Впровадження та розвиток мережевих технологій, що дозволяють об'єднувати сучасні високопродуктивні обчислювальні машини з різною архітектурою в комп'ютерні мережі створює умови для надання комерційного доступу зовнішнім замовникам до надлишків обчислювальних ресурсів під назвою «технології хмарних обчислень» (Cloud Computing) при обробці даних у вигляді потоків завдань на вирішення задач, серед яких можуть бути окремі незалежні задачі, умовно незалежні частини однієї задачі, задачі послідовних виділених етапів обчислень з використанням конвеєрної обробки.

Характерним прикладом задачі виділення ресурсів для паралельної обробки може бути використання технологій хмарних обчислень для підтримки вирішення задач згідно технологій Grid Computing [1], рис.1.

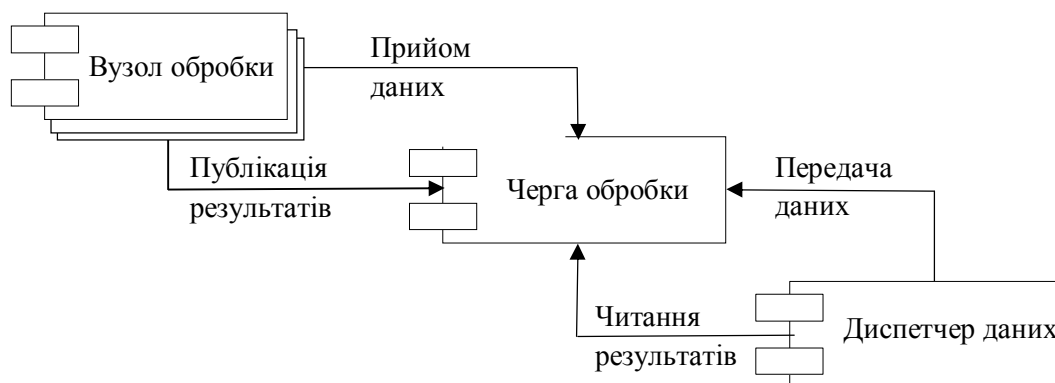


Рис. 1. Архітектура грид-додатків в середовищі хмарних технологій

Обчислювальні ресурси, що надаються такими відомими провайдерами хмарних технологій як Amazon, Google, Microsoft Azure, IBM Cloud [1-3] та ін., мають почасову тарифікацію з врахуванням архітектурних особливостей віртуальних машин, їх технічних характеристик, обсягів робіт по обробці даних та обсягів використання ресурсів.

У зв'язку з цим вирішення задач доцільно виконувати з позицій певних критеріїв ефективності, серед яких можна назвати вартість використання ресурсів, вартість виконання обробки, вартість використання каналів зв'язку, час вирішення задач, надійність, достовірність та ін.

В роботі розглядається задача оптимізації використання обчислювальних ресурсів провайдерів хмарних технологій з побудовою множини ефективних рішень в залежності від складу та важливості обраних критеріїв ефективності, області допустимих значень рішень, що приймаються.

Визначення необхідної топології та архітектури системи обробки даних на основі використання віртуальних машини, яка відповідає сукупностям задач, що вирішуються протягом проміжку часу T , можна сформулювати наступним чином.

Необхідно обрати типи пристроїв обробки (віртуальних машин), визначити необхідні зв'язки між джерелами даних задач та пристроями обробки, що забезпечують вирішення заданих сукупностей завдань та відповідають кращим значенням обраних критеріїв ефективності.

Зазначені рішення потребують використання рішень оптимізаційної задачі математичного програмування з критерієм мінімум витрат та булевими змінними, застосовуючи ефективні алгоритми.

Врахування інших критеріїв, таких як мінімальний час перебування завдання в системі, потребує раціональної організації процесів обробки та вибору більш продуктивних і дорожчих пристроїв обробки. Час обробки і час чекання можуть бути представлені у відповідності з положеннями теорії масового обслуговування залежностями, аналітичний вираз яких здебільшого відомий лише для обмеженої кількості законів розподілу. Тому для оцінки значень компонент критерію доцільним може бути використання імітаційного моделювання з подальшою апроксимацією результатів для оперативного застосування. У якості апроксимуючих функцій вибиралися нелінійні залежності, які враховують основні тенденції впливу аргументів, параметрів та умов взаємодії.

Оскільки критерії ефективності вибору архітектури системи обробки, що розглядаються, суперечливі (наприклад, вартість системи обробки і час обробки, вартість виконання процедур і надійність їх виконання), то важливим питанням стає побудова множини ефективних рішень. Для цього може бути використаний метод узгодження поточних рішень відповідно до поступок по значенням складових критеріїв оптимальності з побудовою множини ефективних рішень.

Висновки. Організація паралельної обробки даних є основним джерелом якості обчислювальних процесів, забезпечуючи зростання продуктивності, зниження вартості обчислень, підвищення надійності і економічності використання. Застосування оперативного розподілу завдань між пристроями обробки з позицій максимального використання доступних ресурсів, обліку їх стану і параметрів використання дозволяє формувати множину ефективних рішень, використання яких призводять до підвищення якості обробки даних і використання обчислювальної системи.

Автор вважає, що в даній роботі новими є такі положення і результати. Отримали подальший розвиток концепції організації паралельної обробки даних в складі обчислювальних систем різної архітектури і топології. Запропоновано підхід до побудови математичних моделей, що дозволяють ефективно формувати топологію і архітектуру обчислювальної системи відповідно до наявних ресурсів і стану обчислювальної системи. В результаті цих дій може бути побудована сукупність рішень, прийнятне рішення серед яких в залежності від рівня топології системи вибирає або ЛПР, або супервізор системи на підставі поставлених цілей, умов їх досягнення, доступних засобів і зіставлення оцінок значень критеріїв оптимальності.

Література

1. High Performance Computing & Grid Computing. Режим доступу: <http://www.its.hku.hk/services/research/hpc-grid>. 11.06.2018.
2. Основи хмарних обчислень – IBM. Режим доступу: <https://www.ibm.com>cl-cloudintro>; 14.03.2012.
3. Google Cloud. . Режим доступу: <https://cloud.google.com>. 12.06.2018.