

МОНІТОРИНГ ІНФРАСТРУКТУРИ ВЕБ-ДОДАТКІВ ХМАРНИХ ОБЧИСЛЕНЬ

Сабурова С.О., Федоров А.С., Кадацька О.І.

Кафедра «Інфокомунікаційної інженерії ім. В.В. Поповського»,
Харківський національний університет радіоелектроніки, Україна

E-mail: svitlana.saburova@nure.ua,
andrii.fedorov@nure.ua,
olha.kadatska@nure.ua

Abstract

Methods of modeling algorithms for controlling the quality parameters of web application services through monitoring the load of cloud computing infrastructure presented. One of the ways to increase the performance and quality parameters of cloud computing web applications, which uses the potential of modern Internet services - networks. Methods of monitoring the quality of infrastructure deployment based on mathematical modeling, methods of calculation and evaluation of load balancing algorithms for access to cloud computing web application services considered.

Основна функція спільного використання ресурсів, програмного забезпечення та інформації через Інтернет є основною зацікавленістю у хмарних обчисленнях з метою зменшення капітальних та експлуатаційних витрат, кращої продуктивності з точки зору часу відгуку та часу обробки даних, підтримання постійності системи та забезпечення адаптації в майбутньому у системі. Отже, існують різні технічні проблеми, які потребують вирішення, такі як перенесення віртуальних машин (ВМ), консолідація серверів, відмовостійкість, висока доступність та масштабованість. Але головним питанням є балансування навантаження трафіку у хмарних обчисленнях у часі. Це механізм розподілу у часі найбільшого навантаження (Business Time – ВТ) між різними вузлами розподіленої системи для покращення як розгортання ресурсів, так і часу реагування на роботу, одночасно уникаючи ситуації, коли деякі вузли мають величезну кількість навантаження, тоді як інші вузли нічого не роблять або простоюють із дуже невеликою роботою. Це також гарантує, що весь процесор в системі або кожен вузол у мережі виконує приблизно однакову кількість роботи в будь-який момент часу [9].

Найбільш підходящий людський агент для потреб замовника може бути обраний на основі типу веб-сторінки, яку переглядає замовник, а також ідентифікації замовника. Мета балансування навантаження у ВТ полягає в наступному: збільшити доступність послуг; збільшити задоволеність користувачів; максимізувати використання ресурсів; скоротити час виконання та очікування завдання, що надходить з іншого місця; для підвищення продуктивності; підтримувати стабільність системи; побудувати систему відмовостійкості; змінити майбутні модифікації;

Проблеми балансування навантаження веб-сторінок хмарних обчислень - це накладні витрати, що застосовуються під ВТ - часи впровадження системи балансування навантаження. Він складається із накладних витрат завдяки руху завдань, міжпроцесорному спілкуванню. Накладні витрати слід зменшити, щоб алгоритм балансування навантаження ВТ працював добре. Пропускна здатність: це кількість завдань, виконаних за фіксований інтервал часів - ВТ. Для підвищення продуктивності система, пропускна здатність повинна бути високою. Продуктивність: Це можна визначити як ефективність системи. Її потрібно вдосконалювати. Використання ресурсів: реалізується для перевірки навантаження ресурсів ВТ. Це повинно бути максимальним для ефективної системи балансування навантаження ресурсів ВТ.

Масштабованість, коли якість обслуговування може бути однаковою, якщо кількість користувачів збільшується. Більшу кількість вузлів можна додати в ресурси ВТ, не впливаючи на службу. Час відгуку: може бути визначено як кількість часу, витраченого на реакцію з допомогою алгоритму балансування навантаження в розподіленої системі. Для кращої роботи цей параметр слід зменшити.

Відмовостійкість: незважаючи на відмову вузла, здатність системи виконувати рівномірний розподіл навантаження. Балансування навантаження ресурсів ВТ є найкращою технікою, що захищає від несправностей.

Точка відмови: проєктована система таким чином, що одна точка відмови не впливає на якість надання послуг. Як і в централізованій системі, якщо один центральний вузол виходить з ладу, тоді вся система вийде з ладу, тому система балансування навантаження ресурсів ВТ повинна бути розроблена для подолання цієї проблеми.

Моніторинг економічно ефективних алгоритмів балансування навантаження для хмарних обчислень веб-сторінок базуються на продуктивності елементів хмарного середовища. Центри обробки даних (ЦОД) неоднорідні за своїм характером із певними характеристиками. Зазвичай центр обробки даних складається з декількох хостів. Хости мають ряд оброблювальних елементів (ОЕ) з характеристиками оперативної пам'яті та пропускну здатності. У хмарному середовищі ці хости віртуалізуються в різну кількість віртуальних машин на основі запиту користувачів. Віртуальні машини також можуть мати неоднорідні характеристики, такі як хости [1].

Хмарний інформаційний сервіс (CIS) - це сховище, яке містить усі ресурси, доступні в хмарному середовищі. Це реєстр центрів обробки даних. Коли центр обробки даних створюється, він повинен зареєструватися в CIS. На рис. 1 представлена загальна архітектура балансування навантаження ресурсів ВТ хмарних обчислень на веб-сторінках.

Покращений алгоритм для ефективного балансування навантаження та планування в хмарі:

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N (PT_1 - PT)^2}, \quad (1)$$

де N - загальна кількість завдань, призначених віртуальній машині, а запропонований алгоритм обчислює час обробки кожного завдання. за допомогою рівняння:

$$PT = \frac{\text{CurrentLoad}}{\text{Capacity}}. \quad (2)$$

Потім рішення про балансування навантаження ресурсів ВТ приймається на основі значення – стандарт відхилення, SD .

Моніторинг CIS збирає інформацію про всі ресурси в центрах обробки даних. На основі цієї інформації хмарний брокер передає ці завдання різним віртуальним машинам у центрі обробки даних хмарних обчислень на веб-сторінках.

У запропонованому методі алгоритм перевіряє перевантажені умови, і він мігрує завдання із перевантажених ВМ в недостатньо завантажені ВМ. Запропонований механізм вирівнювання та планування навантаження ресурсів ВТ працює у чотири різні етапи, як зазначено нижче.

- 1) Розрахунок поточного навантаження ВМ.
- 2) Рішення щодо балансування навантаження та планування.
- 3) Групування ВМ.
- 4) Розклад завдання.

Поточне навантаження на віртуальну машину вимірюється на основі відношення між загальною довжиною завдань, поданих до цієї віртуальної машини, до швидкості обробки цієї віртуальної машини в конкретному екземплярі.

Припустимо, N - загальна кількість завдань, призначених віртуальній машині та L_{en} - це тривалість окремих завдань, а $MIPS$ - це мільйон інструкцій на секунду для цієї ВМ, тоді за допомогою рівняння можна розрахувати поточне навантаження:

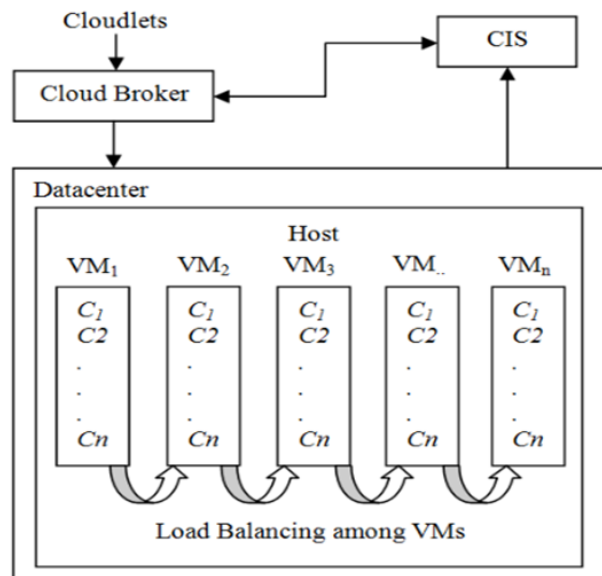


Рис. 1. Архітектура балансування навантаження ресурсів ВТ хмарних обчислень на веб-сторінках

$$Load_{vm} = \frac{N \cdot Len}{MIPS}. \quad (3)$$

Тоді загальне навантаження на центр обробки даних - це сума навантаження на кожен ВМ. Рівняння для загального завантаження центру обробки даних $Load_{dc}$ задається рівнянням:

$$d_{vm} = \frac{N \cdot Len}{MIPS}. \quad (4)$$

Потужність обробки ВМ можна розрахувати, використовуючи рівняння:

$$Load_{DC} \sum_{vm=1}^n Load_{VM}. \quad (5)$$

Ось PE_{num} кількість елементів обробки в конкретній ВМ, PE_{mips} - це обчислювальна потужність PE в MIPS ставка і VM_{bw} - ширина смуги, асоційована для віртуальної машини. Центр обробки даних може мати кілька VMS. Отже, загальна потужність весь центр обробки даних можна розрахувати за допомогою рівняння:

$$Capacity_{DC} = \sum_{vm=1}^n Capacity_{VM}. \quad (6)$$

Тоді час обробки, необхідний центру обробки даних для виконання всіх завдань у ньому, можна обчислити за рівнянням, наведеним нижче

$$PT_{DC} = \frac{Load_{DC}}{Capacity_{DC}}. \quad (7)$$

Стандартне відхилення (SD) є хорошим показником моніторингу відхилень. Запропонований метод використовує SD для моніторингу та вимірювання відхилень робочого навантаження хмарних обчислень на веб-сторінках на кожній ВМ ресурсів ВТ.

Оскільки все більше і більше користувачів вступають у хмарні обчислення, балансування навантаження ресурсів ВТ є важливим завданням в управлінні ресурсами. Оптиміальні алгоритми планування завдань здатні вирішувати проблеми балансування навантаження ресурсів ВТ, а також очікування від вимог системи QoS [2].

Особливості метаевристичних алгоритмів оптимізації роїв (MAOP), присвячені проблемі розміщення віртуальних машин та планування завдань у хмарному середовищі. MAOP – це простий паралельний алгоритм, який можна застосовувати різними способами для вирішення проблем планування завдань. Запропонований алгоритм розглядається як об'єднання алгоритму оптимізації роїв (AOP) та алгоритму пошуку зозулі (АПЗ) називається комбінований алгоритм - АОРТПЗ. Запропонований комбінований алгоритм оцінюється за допомогою Cloudsim Simulator[1].

З іншого боку, алгоритм пошуку зозулі, який розглядається як метаевристичний алгоритм оптимізації, заснований на поведінці птаха-зозулі. Існує деяка схожість між алгоритмом АПЗ та алгоритмами сходження на гірку щодо деякої широкомасштабної рандомізації. Але ці два алгоритми по суті дуже різні.

Міграція завдання виконується лише тоді, коли попит відповідає пропозиції.

1. Економічно ефективно навантаження збалансованого розподілу ресурсів для розділеної хмарної системи. Попереджувальне планування на основі кінцевих термінів (Deadline Based Pre-emptive Scheduling, DBPS) та модель балансування навантаження (Throttled Load Balancing for Cloud, TLBC), засновані на розділенні хмар за допомогою віртуальної машини.

2. Балансування навантаження на основі очікуваного часу закінчення завдань у хмарних обчисленнях. Вдосконалення балансування навантаження за алгоритмом "передбачуваний баланс навантаження часу закінчення", яке враховує поточне навантаження віртуальної машини центру обробки даних та оцінку закінчення обробки часу виконання завдання перед будь-яким розподілом, щоб подолати проблеми, які викликані статичними алгоритмами. Алгоритм дозволяє постачальникам хмарних послуг покращувати продуктивність, доступність та максимально використовувати віртуальні машини у своїх центрах обробки даних.

3. Балансування навантаження за допомогою алгоритму оптимального планування витрат Використовується алгоритм Раунд Роббін (Round Robbin, RR) для планування вхідних завдань, оптимізації вартістості та планування ресурсів на основі вартості. У запропонованому алгоритмі ресурси були згруповані як пакети в кожній ВМ. Коли користувач запитує ресурс, виконується віртуальна

машина, що складається з цього пакета. Цей прийом знижує витрати на виконання постачальника послуг.

4. Балансування навантаження на основі завдань в хмарних обчисленнях за допомогою оптимізації рою частинок. Запропонований метод балансування навантаження системи на основі завдань із використанням оптимізації роєм частинок (Task based System Load Balancing method using Particle Swarm Optimization, TBSLB-PSO), потрібен для того, щоб досягти балансування навантаження системи, лише передаючи додаткові завдання з перевантаженої ВМ, замість того, щоб перенести всю перевантажену ВМ. Для оптимізації додаткових завдань до нових віртуальних машин хоста була використана модель оптимізації рою частинок (Particle Swarm Optimization, PSO).

Результати моделювання оцінюють інтервал дії трьох алгоритмів: випадковий розподіл (BP), оригінальний AOP та комбінований алгоритм AOPTПЗ при використанні 5 віртуальних машин, а завдання 10, 20, 30 та 40, які описані в таблиці 1 та на рис.2.

У таблиці 1 представлений розклад алгоритмів планування завдань BP, AOP та AOPTПЗ з використанням 5 віртуальних машин ВТ.

Таблиця 1. Порівняний алгоритм планування з часом виконання (сек)

Випадковий розподіл	AOP	AOPTПЗ	ВМ	Хмари
3	3,1106937	2,391129032	5	10
4,6	5,230736842	4,268631579	5	20
14,1	12,07002398	9,076738609	5	30
17,1	22,12103448	13,211939655	5	40

Рис. 2 ілюструє порівняння часу ВТ виконання за допомогою алгоритмів. Вісь – X представляє кількість хмарних програм (завдань), а вісь Y – час виконання у секундах.

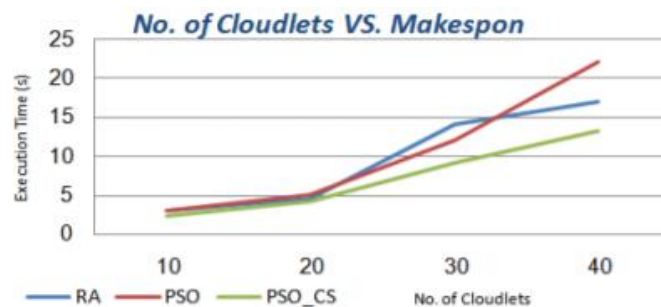


Рис. 2. Графіки залежності продуктивності від часу ВТ виконання хмарних програм

Згідно з результатами на рис. 2, встановлено, що розрахована кількість запропонованих комбінованих алгоритмів – AOPTПЗ час виконання суттєво зменшився на 30% – 31.2 %, враховуючи більшу кількість хмар. Запропонований алгоритм AOPTПЗ забезпечує найкращі результати щодо алгоритмів BP та AOP.

Література

1. Аградж Шарма. Час реакції: Базове балансування навантаження в хмарних обчисленнях/ Шарма Аградж, К.Сатеш//В матеріалах Міжнародної конференції з управління, приладобудування, засобів зв'язку та обчислювальних технологій (ICCSCT), – 2014. – С.С. 1287-1293.
2. Сю Г. Модель балансування навантаження на основі хмарного розділення для публічної хмари/Г. Сю, Панг Дж., Фу Сяодун//Журнал науки і техніки Цінхуа, – 2013. – С.С. 34-39.
3. Шрідхар Джорджія. Оптимальне навантаження у хмарних обчисленнях шляхом ефективного використання віртуальних машин/Джорджія Шрідхар, G.R.M.Редді//В матеріалах Шостої міжнародної конференції IEEE з систем зв'язку та мереж (COMSNETS), – 2014. – С.С. 1-4,