

УДК 519.7:004.032.24

**ОСОБЛИВОСТІ РОБОТИ ОБЧИСЛЮВАЛЬНИХ КЛАСТЕРІВ  
У ВІРТУАЛЬНОМУ СЕРЕДОВИЩІ**

*Нікітченко В.В.*

*Одеська національна академія зв'язку ім. О.С. Попова,  
65029, Україна, м. Одеса, вул. Ковальська, 1.  
vncentaurus@gmail.com*

**ОСОБЕННОСТИ РАБОТЫ ВЫЧИСЛИТЕЛЬНЫХ КЛАСТЕРОВ  
В ВИРТУАЛЬНОЙ СРЕДЕ**

*Никитченко В.В.*

*Одесская национальная академия связи им. А.С. Попова,  
65029, Украина, г. Одесса, ул. Кузнечная, 1.  
vncentaurus@gmail.com*

**FEATURES OF WORK COMPUTING CLUSTERS IN A VIRTUAL ENVIRONMENT**

*Nikitchenko V.V.*

*O.S. Popov Odessa national academy of telecommunications,  
1 Kovalska St., Odessa, 65029, Ukraine  
vncentaurus@gmail.com*

**Анотація.** Запропонована модифікація математичної моделі обчислювального кластера з урахуванням можливості його побудови у віртуальному середовищі. Отримано формальні вирази для типових характеристик роботи кластера, таких як завантаженість і прискорення.

**Ключові слова:** обчислювальний кластер, віртуальний функційний пристрій, дійсний функційний пристрій, завантаженість, прискорення.

**Аннотация.** Предложена модификация математической модели вычислительного кластера с учетом возможности его построения в виртуальной среде. Получены формальные выражения для типовых характеристик работы кластера, таких как загруженность и ускорение.

**Ключевые слова:** вычислительный кластер, виртуальное функциональное устройство, действительное функциональное устройство, загруженность, ускорение.

**Abstract.** Proposed by modification of the mathematical model of a computing cluster, with the possibility of constructing it in a virtual environment. Obtained formal expressions for the typical characteristics of the cluster, such as the load and acceleration.

**Key words:** computing cluster, virtual functional device, real functional device, load, acceleration.

На сьогодні застосування паралельної форми обчислень є одним з головних напрямів розвитку інформаційних технологій, що обумовлюється декількома причинами. Насамперед, це бажання зменшити сумарний час, необхідний для виконання деякої процедури або сукупності таких процедур. Серед іншого можна назвати необхідність розв'язувати "великі" задачі, що не завжди є доцільним та навіть досяжним на звичайних ЕОМ з послідовною формою виконання інструкцій, а також прагнення використати в процесі обчислень доступні мережні ресурси.

Серед різноманітних форм паралельних обчислень особлива увага приділяється обчислювальним кластерам. По-перше, кластерні системи – це один з найпоширеніших способів втілення паралельних обчислень. Це підтверджує щорічний рейтинг суперкомп'ютерів TOP500[1]: більшість представлених суперкомп'ютерів – це саме кластери типу HPC (High Performance Computing). По-друге, популярність кластерних технологій зумовлена їх доступністю та високою продуктивністю. Останнє суттєво для навчальних і наукових закладів, які потребують високопродуктивних обчислень, але звичайно обмежені в ресурсах.

**Метою даної статті** є дослідження деяких особливостей побудови і функціонування обчислювальних кластерів. Основою для створення кластера слугує сукупність звичайних комп'ютерів або робочих станцій (далі – обчислювальних вузлів), що поєднуються мережним сегментом вибраної топології. Раніше говорилося про порівняно невисоку вартість такого обладнання, проте деякі кошти на його придбання все ж потрібні. Альтернативою можна вважати використання вже існуючих комп'ютерних систем шляхом встановлення на них додаткового програмного забезпечення, та можливо, зміною конфігурації мережних зв'язків. Це, у свою чергу, потребує узгодженого графіка використання таких систем, тобто винесення означених робіт на вечірній час, що не завжди зручно. Крім того, нічних годин може виявитись недостатньо для завершення складного обчислювального процесу, проте ми будемо вимушені його перервати та повернути машини до обслуговування їх звичного навантаження.

З кластером на базі дійсного апаратного обладнання пов'язані й інші умови, такі як виділення площі для його фізичного розміщення або суттєве споживання електроенергії. З іншого боку, констатуємо інтенсивний розвиток і збільшення кількості обчислювальних хмар, що в їхній основі також лежить сукупність обчислювальних вузлів, а наявність технологій віртуалізації дозволяє поставити у відповідність декільком віртуальним пристроям один фізичний.

Для організації НРС-кластера потрібна сукупність обчислювальних вузлів, проте ніщо не говорить про походження цих вузлів – фізичне або віртуальне. В даному випадку можемо використати одну з моделей надання хмарних послуг для створення множини віртуальних обчислювальних машин, пов'язати їх сегментом віртуальної мережі, та встановлюючи системне і спеціальне програмне забезпечення, розгорнути кластер. Така форма організації обчислювального кластера ніяк не вплине на кінцевих користувачів послуг, оскільки вони здебільшого взаємодіють лише з його програмним середовищем. Тут під вибраною моделлю надання хмарних послуг розуміємо IaaS (Infrastructure as a Service) – інфраструктура як послуга.

Звичайно, що така ідея не є оригінальною, і подібні рішення присутні вже деякий час. Та вибираючи між фізичною або віртуальною платформами при створенні обчислювального кластера, варто зауважити декілька аспектів. Найпростіше це питання вирішується в економічній площині. Підраховуючи сукупні витрати на створення та подальшу експлуатацію фізичного кластера з одного боку та вартість початкового замовлення і оренди віртуальних обчислювальних вузлів з іншого, виносимо рішення про перевагу однієї з платформ.

Іншим важливим моментом є питання безпеки даних. Звичайно, модель IaaS передбачає цілковитий адміністративний контроль над програмним забезпеченням з боку користувача. В той же час всі апаратні пристрої накопичення та зберігання інформації знаходяться на боці обчислювальної хмари. Ця обставина робить таке рішення потенційно небезпечним, адже нескладно отримати потрібну інформацію за наявності фізичного доступу до її носія, тоді як повне шифрування даних файлової системи віртуального пристрою підходить не у всіх випадках.

Окремо слід виділити питання продуктивності, і саме на них переважно акцентуємо увагу. З продуктивністю в даному випадку непрямо будемо пов'язувати два показники – час, який витрачається процесорними елементами на виконання деякої сукупності обчислювальних операцій, а також час, що необхідний для обміну даними між вузлами кластера. Дана робота фокусується на встановленні формальних значень характеристик, пов'язаних з роботою процесорних елементів, при їх використанні для створення віртуального НРС-кластера в обчислювальній хмарі типу IaaS.

Модель хмарних послуг IaaS звичайно будується на основі технології апаратної віртуалізації, яка передбачає реалізацію множини VCPU (Virtual CPU) як спеціального режиму роботи фізичного процесорного елемента. Власне спеціальний режим втілюється

через додатковий набір інструкцій, що залежно від виробника називаються VT-x і AMD-V, для процесорів Intel та AMD відповідно [2, 3]. Ідея апаратної віртуалізації – напряму використовувати ресурси пристроїв в віртуальних машинах. Маємо два режими роботи – Root Mode і Non-Root Mode, де в першому з них функціонує спеціальне програмне забезпечення віртуалізації, а в другому виконується код «гостьових» систем. Обидва режими при цьому підтримують традиційний розподіл повноважень у вигляді так званих кілець захисту, від Ring0 до Ring3, а перемикання між режимами реалізується інструкціями VMENTRY і VMEXIT [4]. Програмною складовою такого типу віртуалізації є гіпервізор, який серед іншого здійснює керування розподілом ресурсів між декількома операційними системами обчислювального вузла.

Для дослідження обчислювального кластера довільної природи застосуємо модель паралельного процесу у вигляді ациклічного орієнтованого графа  $G(V, E)$  [5]. Тут  $V$  є скінченна множина вершин графа, де їм ставиться у відповідність множина операцій паралельного алгоритму. Множина  $E$  двохелементних підмножин  $\{a, b\}$  множини  $V$  являє собою сукупність дуг орієнтованого графа. Якщо відповідно до змісту алгоритму результат виконання операції  $a$  є аргументом для наступної операції  $b$ , відповідні вершини поєднуються дугою  $\{a, b\} \in E$  [6]. При цьому з однієї вершини може виходити декілька дуг, якщо результат поточної операції використовується багатьма наступними операціями, вершина може зовсім не мати вхідних дуг –  $indeg(v) = 0$  – якщо це початкові дані, або вихідних дуг –  $outdeg(v) = 0$  – якщо це кінцевий результат.

Розглядаючи віртуальний кластер, створений в обчислювальній хмарі, окремо від дійсних фізичних пристроїв можемо зазначити, що за своїми властивостями він практично нічим не відрізняється від традиційного кластера. Для нього справедливі всі ті співвідношення, що вони отримані в [6,7] для паралельних обчислювальних систем. Типові характеристики, такі як завантаженість, прискорення, номінальна та реальна продуктивність, цілковито зберігають своє змістовне навантаження, та при сумісному розгляданні віртуального кластера і фізичних пристроїв формальні вирази для таких характеристик повинні бути доопрацьовані.

Поняття функційного пристрою (ФП) широко використовується при дослідженні паралельних систем. Для можливості опису паралельних систем з віртуалізацією розширимо поняття ФП та введемо до розгляду два нових елементи – дійсний функційний пристрій та віртуальний функційний пристрій.

*Визначення 1.* Дійсним функційним пристроєм (ДФП) будемо називати довільний пристрій фізичної природи, який входить до складу обчислювальної системи. Залежно від контексту це може бути процесорний елемент, якщо ми досліджуємо виконання обчислювальних операцій, пристрій зберігання даних при розгляді систем з пам'яттю та ін.

*Визначення 2.* Віртуальним функційним пристроєм (ВФП) будемо називати довільний пристрій, який існує у віртуальній площині та здатний до відображення на деякий дійсний пристрій з відповідними можливостями. Під віртуальною площиною розуміємо абстракцію, створену програмними, програмно-апаратними або суто апаратними засобами віртуалізації ресурсів. Тут можемо провести аналогію до поняття віртуального процесора VCPU, проте змістовне навантаження ВФП більш широке, і передбачає можливість опису будь-якого пристрою, а саме для виконання обчислень, зберігання даних, мережної взаємодії тощо.

Можливе існування однієї або більше сукупності віртуальних пристроїв, поєднаних зв'язками вибраної топології у своїй віртуальній площині, проте реальні дії звичайно виконуються на ДФП. Тому необхідно знову встановити зміст таких характеристик, як завантаженість та прискорення, враховуючи особливості віртуалізації відповідних пристроїв.

Позначимо через  $S^{virt}$  кількість віртуальних процесорних елементів, а через  $S^{phy}$  – кількість фізичних процесорних елементів. Співвідношення  $S^{virt} / S^{phy} < 1$  теоретично можливе, проте не має практичної цінності, оскільки зменшуючи кількість доступних

процесорів обчислювальної хмари, може бути зведене до випадку  $S^{virt} / S^{phy} \geq 1$ . За наявності  $n$  віртуальних площин пристроїв, загальний вираз для конфігурації змістовних варіантів співвідношення набуває вигляду  $\sum_{i=1}^n S_i^{virt} / S^{phy} \geq 1$ .

Грунтуючись на методиці застосування теорії графів у дослідженні паралельних систем і алгоритмів, а також техніці апаратної віртуалізації процесорних елементів, побудуємо модель віртуального кластера. Завантаженість окремого пристрою паралельної системи на вказаному проміжку часу є «відношення вартості дійсно виконаних операцій до максимально можливої вартості операцій за даний час» [6]. Оскільки вартість операцій ототожнюємо з часом їх виконання, можемо записати вираз для завантаженості довільного  $i$ -го ВФП

$$p_i^{VFD} = \frac{t_i^w}{t_i^w + t_i^{gos}}, \quad (1)$$

де  $p_i^{VFD}$  – завантаженість  $i$ -го віртуального пристрою;  $t_i^w$  – час виконання користувацького процесу, що знаходиться у віртуальній площині;  $t_i^{gos}$  – час виконання службових операцій «гостьової» операційної системи, враховуючи час простоювання.

Завантаження кластера із  $m$  віртуальних пристроїв встановлюється виразом  $p^{VFD} = \sum_{i=1}^m \eta_i p_i^{VFD}$  [6]. Коефіцієнт зваженої суми загалом визначається як  $\eta_i = \alpha_i \times \left( \sum_{i=1}^m \alpha_i \right)^{-1}$ , де  $\alpha_i$  – номінальна продуктивність відповідного пристрою. Враховуючи, що хмарою IaaS звичайно поставляються типові VCPU, та вважаючи ВФП простими, тобто  $\alpha_i = \alpha_j = \alpha = 1$ , вираз загальної продуктивності віртуального кластера можемо переписати у більш зручному для практичного використання вигляді:  $p^{VFD} = \frac{1}{m} \sum_{i=1}^m t_i^w / (t_i^w + t_i^{gos})$ .

Вищезазначені формальні вирази мають зміст для розгляду віртуального кластера без урахування його відображення на фізичні пристрої. Проте головним завданням є побудова такої моделі віртуального кластера, яка буде містити дві складові – віртуальну та дійсну, з кінцевою метою отримати характеристики для ДФП. Отже, визначимо завантаженість для ДФП, де на ньому може бути відображена множина ВФП. Спочатку розглянемо найпростіший випадок, коли лише одному ВФП ставиться у відповідність деякий ДФП. Враховуючи наявність віртуалізації, та аналогічно до (1), запишемо вираз для завантаженості  $i$ -го ДФП у вигляді

$$p_i^{RFD} = \frac{t_i^w}{t_i^w + t_i^{gos} + t_i^{hyp}}, \quad (2)$$

де  $p_i^{RFD}$  – завантаженість  $i$ -го фізичного пристрою з урахуванням віртуальної складової; значення  $t_i^w$  і  $t_i^{gos}$  аналогічні таким з (1), а  $t_i^{hyp}$  – час роботи гіпервізора на дійсному обладнанні, який містить також час простоювання ДФП. Слід зауважити, що такий вираз справедливий лише для систем з апаратною віртуалізацією, в яких можливо розрізняти окремі користувацькі процеси віртуальних машин. Для природної віртуалізації, яка здійснюється виключно програмними засобами, все навантаження віртуальної машини представляється одним процесом.

Шляхом тотожних перетворень можемо виразити  $p_i^{RFD}$  через  $p_i^{VFD}$ :

$$p_i^{RFD} = \frac{t_i^w}{t_i^w + t_i^{gos}} \times \frac{t_i^w + t_i^{gos}}{t_i^w + t_i^{gos} + t_i^{hyp}} = p_i^{VFD} \times \frac{t_i^w + t_i^{gos}}{t_i^w + t_i^{gos} + t_i^{hyp}}, \quad (3)$$

де другий множник кінцевого виразу також має практичний зміст. Користуючись визначенням завантаженості, стверджуємо, що це є завантаженість фізичного процесора в цілому, розглядаючи всю його роботу в режимі Non-Root Mode як користувацьке навантаження, за прикладом природної віртуалізації. Позначимо його через  $p_i^{VFD/RFD}$ , запишемо цей вираз окремо, а також врахуємо це в (3):

$$p_i^{VFD/RFD} = \frac{t_i^w + t_i^{gos}}{t_i^w + t_i^{gos} + t_i^{hyp}}, \quad (4)$$

$$p_i^{RFD} = p_i^{VFD} \times p_i^{VFD/RFD}. \quad (5)$$

У загальному випадку маємо розподілені в часі відображення  $n$  віртуальних пристроїв на один ДФП. В такому разі вираз для  $p_i^{RFD}$  набуває вигляду

$$p_i^{RFD} = \frac{\sum_{j=1}^n t_{ij}^w}{\sum_{j=1}^n (t_{ij}^w + t_{ij}^{gos})} \times \frac{\sum_{j=1}^n (t_{ij}^w + t_{ij}^{gos})}{\sum_{j=1}^n (t_{ij}^w + t_{ij}^{gos}) + t_i^{hyp}}. \quad (6)$$

З урахуванням (5) можемо записати вираз для завантаженості кластера із  $m$  ДФП [8]:

$$p^{RFD} = \sum_{i=1}^m \theta_i p_i^{VFD} p_i^{VFD/RFD}, \quad (7)$$

де  $\theta_i$  – коефіцієнт зваженої суми для  $i$ -го ДФП;  $\theta_i = \beta_i \times \left( \sum_{i=1}^m \beta_i \right)^{-1}$ ;  $\beta_i$  – номінальна продуктивність  $i$ -го ДФП.

При цьому необхідно розуміти, що значення номінальних продуктивностей  $\beta_i$  відповідних дійсних функційних пристроїв не досягають одиниці, оскільки потрібно не менш однієї операції перемикавання між режимами Root Mode і Non-Root Mode. Тобто, на відміну від твердження  $0 \leq p^{VFD} \leq 1$  для ВФП, для дійсного пристрою навіть за інших ідеальних умов маємо  $0 \leq p^{RFD} < 1$ . Звичайно, це справедливо лише за обов'язкової наявності режиму віртуалізації, та порівняно з аналогічним пристроєм без віртуалізації.

Під прискоренням розуміємо відношення часу виконання деякого алгоритму на одному простому функційному пристрої до часу виконання означеного алгоритму на системі таких пристроїв (у нашому випадку – обчислювальному кластері) [7]. Позначивши прискорення для сукупності ДФП як  $R^{RFD}$ , отримуємо [8]:

$$R^{(RFD)} = \frac{\sum_{i=1}^m p_i^{VFD} p_i^{VFD/RFD} \beta_i}{\max_{1 \leq i \leq m} \beta_i}. \quad (8)$$

У підсумку зазначаємо, що в статті модифікуються деякі положення графової моделі паралельної системи, враховуючи можливу наявність віртуальної складової. Введено поняття де дано визначення віртуальному і дійсному функційним пристроям, а також отримано формальні вирази для завантаженості і прискорення обчислювального кластера з урахуванням віртуалізації ресурсів.

ЛІТЕРАТУРА:

1. "June 2014." Top 500 supercomputer sites. [Електронний ресурс] – 2014. – Режим доступу: <http://www.top500.org/lists/2014/06/>.
2. Torres, Gabriel. Everything You Need to Know About the Intel Virtualization Technology | Hardware Secrets. [Електронний ресурс] – 2014. – Режим доступу: <http://www.hardwaresecrets.com/article/everything-you-need-to-know-about-the-intel-virtualization-technology/263>.
3. "AMD64 Virtualization Codenamed "Pacifica" Technology." [Електронний ресурс] – 2014. – Режим доступу: <http://www.mimuw.edu.pl/~vincent/lecture6/sources/amd-pacifica-specification.pdf>.

4. Wahlig, Elsie. Hardware Based Virtualization Technologies. [Електронний ресурс] – 2014. Режим доступу:  
<https://www.redhat.com/f/summitfiles/presentation/June2/ECO/Red%20Hat%20Summit%20Breakout%20A%20Elsie.pdf>.
5. Андерсон Дж.А. Графы, ориентированные графы и деревья. [Текст] / Андерсон Дж.А. // Дискретная математика и комбинаторика. – К.: Издательский дом "Вильямс", 2004. – С. 244-298.
6. Воеводин В.В. Математические модели и методы в параллельных процессах. [Текст] / Воеводин В.В. // Математические модели и методы в параллельных процессах. – М.: Наука, Главная редакция физико-математической литературы, 1986.
7. Воеводин В.В. Параллельные вычисления. [Текст] / В.В. Воеводин, Вл.В. Воеводин // Параллельные вычисления. – СПб.: БХВ-Петербург, 2002.
8. Vitalii Nikitchenko. Application graph model to describe the virtual clusters of the HPC type in the cloud computing. / Vitalii Nikitchenko. // Матеріали дванадцятій міжнародній конференції TCSET'2014 "Modern problems of radio engineering, telecommunications and computer science". – Львів — Славське, 2014. – С. 75.

REFERENCES:

1. "June 2014." *Top 500 supercomputer sites*. N.p., n.d. Web. 29 Oct. 2014. <<http://www.top500.org/lists/2014/06/>>.
2. Torres, Gabriel. "Everything You Need to Know About the Intel Virtualization Technology | Hardware Secrets." *Everything You Need to Know About the Intel Virtualization Technology | Hardware Secrets*. N.p., n.d. Web. 29 Oct. 2014. <<http://www.hardwaresecrets.com/article/everything-you-need-to-know-about-the-intel-virtualization-technology/263>>.
3. "AMD64 Virtualization Codenamed "Pacifica" Technology." University of Warsaw. N.p., n.d. Web. 29 Oct. 2014. <<http://www.mimuw.edu.pl/~vincent/lecture6/sources/amd-pacifica-specification.pdf>>.
4. Wahlig, Elsie. "Hardware Based Virtualization Technologies." *RedHat*. N.p., n.d. Web. 29 Oct. 2014. <<https://www.redhat.com/f/summitfiles/presentation/June2/ECO/Red%20Hat%20Summit%20Breakout%20A%20Elsie.pdf>>.
5. Anderson, James A. "Графы, ориентированные графы и деревья." *Дискретная математика и комбинаторика*. Kiev: Izdatel'skii dom "Williams", 2004. 244-298. Print.
6. Voevodin, V.V.. *Matematicheskie modeli i metody v parallel'nykh protsessakh*. Moskva: "Nauka", Glavnaya redaktsiya fiziko-matematicheskoi literatury, 1986. Print.
7. Voevodin, V.V., and V.I.V. Voevodin. *Parallel'nye vychisleniya*. Sankt-Peterburg: BKhV-Peterburg, 2002. Print.
8. Nikitchenko, Vitalii. "Application Graph Model to Describe the Virtual Clusters of the HPC Type in the Cloud Computing." *Modern Problems of Radio Engineering, Telecommunications and Computer Science*. Lviv - Slavske, Ukraine. 25 Feb. 2014: 75. Print.