

МЕТОД БАЛАНСУВАННЯ НАВАНТАЖЕННЯ НА ОСНОВІ ІНТЕГРОВАНОЇ АРХІТЕКТУРИ УПРАВЛІННЯ З ВИКОРИСТАННЯМ ФУНКЦІЇ NVF

Стрихалюк Б.М., Шпур О.М., Селюченко М.О.

*Інститут телекомунікацій, радіоелектроніки та електронної техніки
Національний університет «Львівська політехніка», Україна
E-mail: o.shpur@gmail.com, bogdan_str@ukr.net, m.seliuchenko@gmail.com*

The method of load balancing based on integrated management architecture using functions NVF

Load balancing seeks to improve the performance of distributed system, in terms of load distribution between multiple interacting hosts. We propose a new approach to solving the problem of load balancing, by implementing an integrated management architecture based on function NVF. This analysis will be based on the maximum integral index of physical resources. The maximum index of virtual and telecommunication resources will be transferred harmonist, which if necessary will carry out the migration of applications to the less loaded servers.

Хмарні обчислення працюють на основі Інтернету і використовують комп'ютерні технології. Концепція включає в себе і може надавати у користування інфраструктуру як сервіс (IaaS), платформу як послугу (PaaS) і програмне забезпечення як послугу (SaaS). Одна з найважливіших ідей, що лежать в основі ідеї хмарних обчислень є масштабованість, а ключовою технологією, яка робить це можливим є віртуалізація [1]. Віртуалізація дозволяє більш ефективно використовувати сервери шляхом об'єднання кількох операційних систем і додатків на одному комп'ютері із загальним доступом. Віртуалізація також дозволяє онлайн міграцію, особливо коли сервер перевантажений, і екземпляр операційної системи і її додатків, можуть бути перенесені на новий, менш завантажений сервер. В таких випадках гостро постає проблема балансування навантаження, особливо коли міграція відбувається в режимі реального часу.

Балансування навантаження прагне підвищити продуктивність розподіленої системи, в плані розподілу навантаження між безліччю взаємодіючих хостів [2]. Така система може спробувати рівномірно розподілити навантаження на кожен хост і мати дуже малі відхилення від робочого навантаження на всіх інших фізичних хостах, або може спробувати уникнути перевантажень і блокувань на окремих серверах.

Зазвичай рішення про міграцію віртуальної машини приймається на основі даних про завантаженість серверів: наявність вільної смуги пропускання логічного і фізичного каналу, наявність ресурсів віртуальних машин (оперативної пам'яті, можливості процесора і т.п.) [3]. Ми пропонуємо новий підхід для оцінки цих показників, що дасть змогу вирішити проблему балансування навантаження, за допомогою реалізації інтегрованої архітектури управління на основі функції NVF.

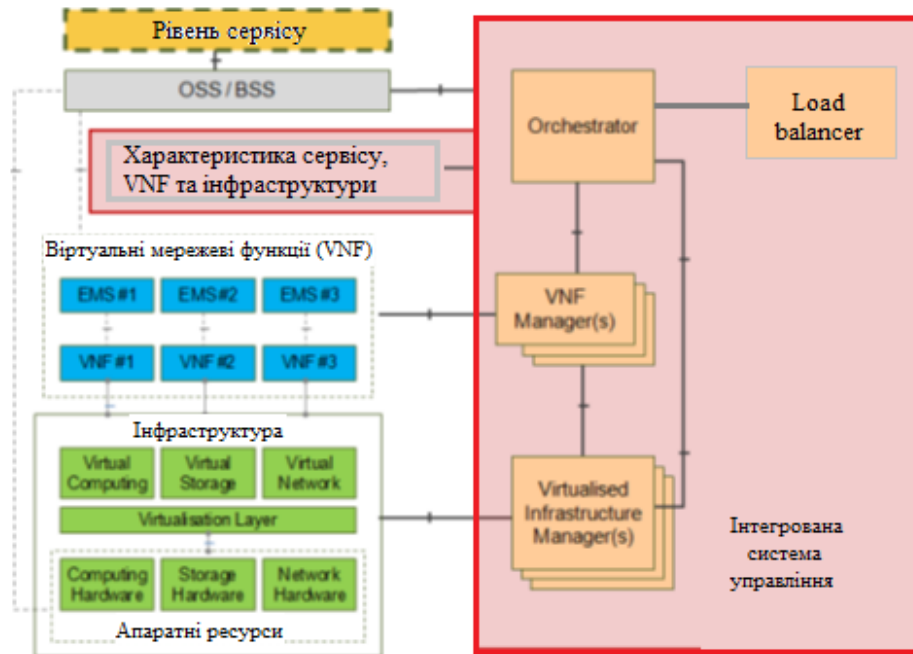


Рис. 1 Архітектура cloud системи на основі функції NFV

Запропонована архітектура управління на основі функції NFV включає в себе:

- Virtual Infrastructure Manager (VIM), який відповідає за управління NFV ресурсів.
- Менеджер VNF - відповідає за управління життєвим циклом примірників VNF (примірника, конфігурації, оновлення, масштабування вгору / вниз). Менеджер VNF несе відповідальність за весь життєвий цикл VNF і через Orchestrator зажадати ресурсів інфраструктури та взаємодіяти з ними, наприклад, встановити програмний компонент або налаштувати VNF
- Orchestrator - несе відповідальність за оркестровку та управління ресурсами NFVI і надає мережеві послуги на NFVI. Orchestrator містить загальний інтерфейс, який слідкує за розгортанням та наданням послуг; обробляє інформацію служб моніторингу та конфігурації; а також на основі даних від LB приймає рішення про міграцію . З іншого боку він має інтерфейс до різних VNF менеджерів, а також до VIM (ів). Крім того, він має доступ до сервісу, інфраструктури модуля VNF, який містить інформацію про пріоритетність послуг.
- Load balancer (LB) – аналізує та визначає доступні віртуальні та фізичні ресурси, та надсилає запити до Orchestrator, який прийме рішення про необхідність міграції віртуальних машин з метою балансування навантаження та оптимального використання ресурсів

LB забезпечуватиме підтримку з'єднань між всіма фізичними хостами через VDE комутатори. Перемикач на такий тип комутатора існує на кожному фізичному сервері. В даній архітектурі Load balancer буде виконувати функцію аналізатора навантаження між фізичними вузлами в хмарі, щоб збалансувати навантаження між серверами в залежності від їх використання процесора або ІО використання та за допомогою VDE перемикачів здійснювати «перемикання» додатків на інші фізичні машини [4]. Такий аналіз буде

здійснюватися на основі максимального інтегрального показника ресурсів фізичної машини. Для його визначення необхідно контролювати наявність не зайнятих апаратних ресурсів та їх доступність. Під доступністю мається на увазі наявність вільних ресурсів віртуальних машин та каналів зв'язку. Параметри не задіяних потужностей кожної віртуальної машини, на якій розташовано M_i додатків розраховуються за формулами:

$$CPU_{pr} = \frac{\sum_{i=1}^k M_i \times CPU_i}{\sum_{i=1}^k M_i}, \quad (1)$$

$$RAM_{pr} = \frac{\sum_{i=1}^k M_i \times RAM_i}{\sum_{i=1}^k M_i}, \quad (2)$$

де M_i – кількість додатків у i -тій VM; CPU_i – тактова частота процесора VM, яку використовує i -тий додаток; RAM_i – об'єм оперативної пам'яті VM, яку використовує i -тий додаток; k – кількість VM cloud-системи.

Для оцінки стану вільних ресурсів пропонуємо наведений у [5] алгоритм. Він дозволить в певній мірі на базі інформації про обчислювальні потужності кожного вузла пришвидшити час обробки запитів та аналізувати ступінь завантаженості кожної машини

Оскільки в cloud систему надходить велика кількість запитів на надання сервісів, які відображають різні маршрути передачі даних використовуючи для цього смугу пропускання каналу, то можна легко визначити яку смугу частот займає той чи інший додаток, беручи за основу вагові коефіцієнти каналів. Вони вказує на відсоткове співвідношення кількості пакетів потоку, що необхідно передати по даному каналу зв'язку. Пріоритет визначається відносно коефіцієнта ефективності використання каналу:

$$k_i = \frac{\int_{t_1}^{t_2} f_i(t) dt}{C_i(t_2 - t_1)} \quad (3)$$

де i – порядковий номер гілки окремого тунелю, $f_i(t)$ – функція, яка описує потік, що проходить через канал i , C_i – смуга пропускання i каналу.

Значення пріоритету додатку P_i виражається обернено пропорційно до коефіцієнта ефективності використання, що у свою чергу спричиняє одночасне підтримання ефективності використання каналу та сприяє оптимальному його завантаженню:

$$P_i = \left(1 - \frac{\int_{t_1}^{t_2} f_i(t) dt}{C_i(t_2 - t_1)}\right) \cdot 100 \quad (4)$$

де множник 100 нормалізує пріоритет відповідно до стобальної шкали.

Визначивши по пріоритету додатку, що певна смуга пропускання каналу використовується неефективно, за допомогою VDE перемикача здійснюється процес вивільнення частини смуги, що не використовується, та відбувається її розподіл між додатками із більшою потребою, що визначається за допомогою прогнозування інтенсивності навантаження

Максимальне значення показників вільних віртуальних та телекомунікаційних ресурсів будуть передаватися Orchestrator, який в разі потреби здійснюватиме міграцію додатків на менш завантажені сервери.

Балансування навантаження прагне підвищити продуктивність розподіленої системи, в плані розподілу навантаження між безліччю взаємодіючих хостів. Ми пропонуємо якісно новий підхід для вирішення проблеми балансування навантаження, за допомогою реалізації інтегрованої архітектури управління на основі функції NVF. Такий аналіз буде здійснюватися на основі максимального інтегрального показника ресурсів фізичної машини. Максимальне значення показника вільних віртуальних та телекомунікаційних ресурсів будуть передаватися Orchestrator, який в разі потреби здійснюватиме міграцію додатків на менш завантажені сервери

Література

1. Стрихалюк Б.М., Алгоритми пошуку шляху за критерієм мінімальної затримки для центру обробки даних //Стрихалюк Б.М., Шпур О.М., Селюченко М.О., Андрухів Т.В.// Вісник Національного університету «Львівська політехніка» №796. Радіоелектроніка та телекомунікації. – Львів. - 2014.с.176-181
2. Strykhalyuk B., Service provisioning by using a structure stability algorithm in a virtualized data center based on cloud technology // Bogdan Strykhalyuk, Olga Shpur, Andriy Masiuk // Computational Problems of Electrical Engineering, vol. 4, №1, 2014, p.83-88
3. Стрихалюк Б.М., Віртуалізація мобільних систем зв'язку на основі технології NFV та моделей cloud-сервісів// Стрихалюк Б.М., Шпур О.М., Масюк А.Р.// Сучасні проблеми телекомунікацій та підготовка фахівців в галузі телекомунікацій: матеріали конференції (30 жовтня – 02 листопада 2014 р. м. Львів), 2014р., с.21-24,
4. C Clark, KFraser, S Hand, J Hansen, and E Jul., “Live migration of virtual machines”, Proceedings of the 2nd ACM/USENIX Symposium on Networked Systems Design and Implementation (NSDI), pages: 273-286
5. Климаш М. М. Модель надання сервісів на основі методу адаптації логічної структури cloud-системи / Климаш М. М., Стрихалюк Б. М., Шпур О. М., Бешлей М. І./ Наукові записки Українського науково-дослідного інституту зв'язку. – 2014. – №5(33) с. 27-36