

УДК 621.391

**Климаш М. М.**, докт. техн. наук, професор (Тел.: +380 50 431 98 07. E-mail: mklimash@lp.edu.ua)

**Стрихалюк Б. М.**, канд. техн. наук, доцент (Тел.: +380 93 920 06 98. E-mail: bogdan\_str@ukr.net)

**Шпур О. М.**, аспірант (Тел.: +380 95 698 11 01. E-mail: o.shpur@mail.ru)

**Бешлей М. І.**, аспірант (Тел.: +380 93 710 32 44. E-mail: beshlembi@gmail.com)

(Національний університет «Львівська політехніка», м. Львів)

## МОДЕЛЬ НАДАННЯ СЕРВІСІВ НА ОСНОВІ МЕТОДУ АДАПТАЦІЇ ЛОГІЧНОЇ СТРУКТУРИ CLOUD СИСТЕМИ

**Климаш М. М., Стрихалюк Б. М., Шпур О. М., Бешлей М. І. Модель надання сервісів на основі методу адаптації логічної структури cloud-системи.** Запропоновано модель надання сервісів на основі методу адаптації логічної структури cloud-системи при міграції віртуальних машин. Здійснено оцінку стійкості структури у площині надання сервісу з врахуванням фізичних та апаратних ресурсів сервера. Зменшено загальну затримку надання сервісу у cloud-системі.

**Ключові слова:** cloud-система, хмарні обчислення, віртуальна машина, логічна структура, центр обробки даних

**Климаш М. Н., Стрихалюк Б. М., Шпур О. Н., Бешлей Н. И. Модель предоставления сервисов на основе метода адаптации логической структуры cloud-системы.** Предложена модель предоставления сервисов на основе метода адаптации логической структуры cloud-системы при миграции виртуальных машин. Осуществлена оценка стойкости структуры в плоскости предоставления сервиса с учетом физических и аппаратных ресурсов сервера. Уменьшено общую задержку предоставления сервиса в cloud-системе.

**Ключевые слова:** cloud-система, облачные вычисления, виртуальная машина, логическая структура, центр обработки данных

**Klymash M. M., Strykhaluyuk B. M., Shpur O. M., Beshley M. I. Model of services grant based on the method of adaptation cloud-system logical structure.** Model of cloud services that is based on adaptation of the logical structure in cloud systems In this paper the model of cloud services that is based on adaptation of the logical structure in cloud systems during migration of virtual machines has been proposed. Here, the stability of the structure in the service plane taking into account physical hardware and server resources has been assessed and overall delay of the service in the cloud system has been decreased.

**Keywords:** cloud-system, cloudy calculations, virtual machine, logical structure, processing data center

**Вступ.** Однією з основних реорганізацій, яка у світовому масштабі впливає на всі інформаційні та комунікаційні галузі, є перехід до використання cloud-технологій (cloud – хмара). Ця технологія передбачає використання розподіленої системи пристроїв та обчислювальних ресурсів як пулу для надання сервісу багатьом користувачам одночасно. Така cloud-система має потужну обчислювальну інфраструктуру, в основі якої лежить надшвидкісна мережа. Завдяки динамічному середовищу cloud-системи спільне використання ресурсів стало зручним та економічно вигідним. Cloud-системи використовуються як платформи для створення будь-якого сервісу [1].

На даний момент процес стандартизації в області хмарних обчислень продовжується, тобто в плані розробленості цієї галузі можна впевнено сказати, що вона лише почала свій розвиток [2]. Відповідно недостатньо досліджено надання сервісів з наукової точки зору.

Cloud -системи є складовою при розподілених обчисленнях, які за допомогою реплікації компонентів сервісу та великим обчислювальним ресурсам забезпечують виконання будь-яких запитів на надання сервісу як у найкоротшій рядка. Різноманітність таких запитів та їх обслуговування було б неможливим без функції віртуалізації. У хмарних системах віртуалізація - це технологія, що дозволяє абстрагуватися від апаратних засобів до точки, де програмні стеки можуть бути розгорнуті і смороду будуть задіяні, не будучи прив'язаними до конкретного фізичного сервера. Віртуалізація дозволяє створити динамічний центр обробки даних (ЦОД), де сервери забезпечують пул ресурсів, які будуть використані в разі споживи, і де додатки для обчислення, зберігання і мережеві ресурси, змінюватимуться динамічно з метою задоволення потреб. У роботі [3] визначений склад сервісів і служб

інфраструктури додатків ЦОД, сформульовані основні вимоги до вказаних служб по параметрах доступності, безпеки, масштабованості, керуваності, консолідації і взаємодії.

Кожен сервіс генерує потоки даних з особливими параметрами, при цьому кожен сервіс має свої вимоги до якості обслуговування, які потребує користувач мережі. З позиції користувача, якість обслуговування розглядається як найшвидший процес обробки його запитів на надання сервісу. Тому розроблення моделі гарантованого надання сервісу в умовах їх різноманітності з врахуванням адаптації логічної структури при міграції віртуальної частини центру обробки даних cloud-системи є актуально науковою задачею.

**Постановка задачі.** Перевагою cloud-систем при виконанні розподілених обчислень є використання можливості розгортання віртуальних машин на серверах в залежності від клієнтських запитів, які надходять до ЦОД та можливість міграції віртуальних машин (Virtual Machine – VM) з одного сервера на інший. Центр обробки даних – це цілий комплекс інженерних та ІТ-систем, який є невід'ємною частиною множини телекомунікаційних структур, він повинен забезпечити єдиний інформаційний ресурс з гарантованими рівнями достовірності, доступності та безпеки даних (Рис.1).

Важливою особливістю хмарних обчислень є можливість міграції навантаження між серверами, як у середині центру обробки даних так і крізь центри обробки даних (навіть в різних географічних регіонах). Під міграцією розуміється можливість переміщення цілих VM (з їхніми операційними системами і додатками) з однієї машини на іншу і продовжити роботу віртуальної машини на другій машині, тобто здійснити її ініціалізацію з присвоєнням, найчастіше, іншої IP-адреси [4]. Ця міграція може бути необхідною через вартість (дешевша вартість для запуску робочих навантажень в центрах обробки даних в іншій країні, в залежності від часу доби або споживаної потужності), або з міркування ефективності (наприклад, пропускна здатність мережі).

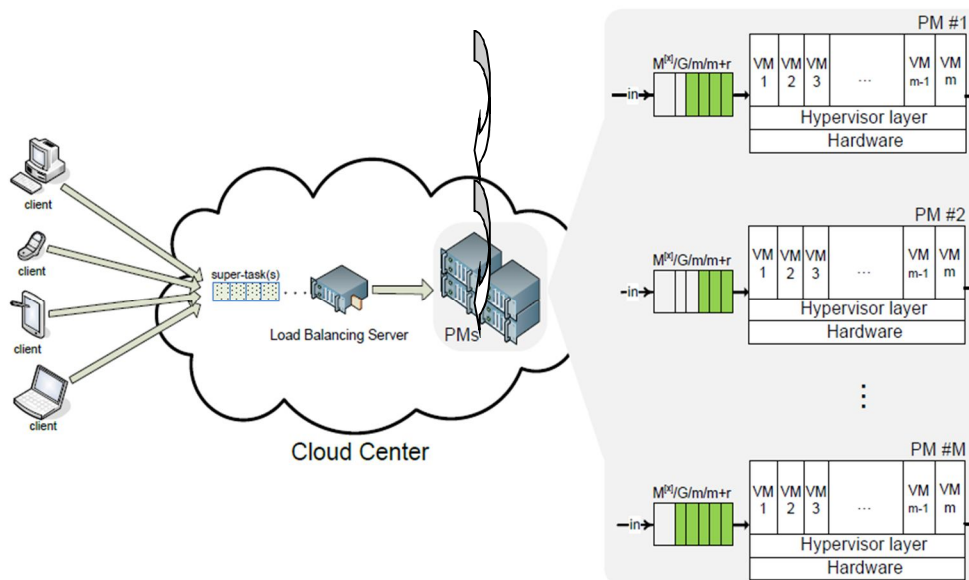


Рис. 1. Структура cloud системи

При міграції VM запити на надання сервісу перенаправляються на інші логічні, а й часом і фізичні канали. Зміна структури ЦОД вносить затримки при обслуговуванні запитів, що безперечно впливатиме на загальний час надання сервісу. Отож, постає проблема зменшення часу обслуговування запитів, які надходять до центру обробки даних, з врахуванням топологічної структури такого центру. Однак врахування структури ЦОД недостатньо, важливим фактором при цьому виступає живучість такої структури, адже чим стійкіша структура, тим система швидше виконує та перенаправляє користувачу необхідний сервіс.

Дослідженням даної проблеми займалися багато вчених. Так у роботі [5] проведена оцінка і аналіз живучості “типових” структур (зірка, кільце), в яких процес появи нових вузлів не здійснює значного впливу на процес надання сервісу кінцевому користувачу. У роботі [6] запропонована модель для оцінки ефективності високо віртуалізованого хмарного центру з Пуасонівським розподілом надходження задач та звичайним розподілом розміру задач.

Модель базується на двоступеневій техніці апроксимації, де основний немарківський процес спочатку моделюється як вбудований напівмарківський процес, який потім моделюється як апроксимований процес Маркова, але тільки в моменти надходження надзадач. Однак дана модель не передбачає зміни положення віртуальних машин і, як наслідок, їх вплив на час надання сервісу кінцевому користувачу

Для цього у даній статті пропонується новий підхід надання сервісу на основі методу оцінки стійкості структури віртуалізованої частини ЦОД з адаптацією логічної її складової в залежності від зміни інтенсивності поступлення запитів, що надходять до VM, завантаженості мережевих пристроїв провайдера cloud послуг та корпоративних компаній, що дасть змогу на підставі даних про стійкість структури в конкретні моменти часу не здійснювати повторний перерахунок оптимального шляху передавання, що призведе до зменшення загальної затримки.

**Математичне представлення моделі.** При наданні cloud послуг провайдери змушені обробляти запити різного типу. Користувач потребуючи сервіс навіть не усвідомлює всю складність процесу передавання його запиту до центру обробки даних. Перш за все запит повинен надійти у площину корпоративної мережі, де вихідні шлюзи перенаправляють його на площину обслуговування провайдера. Площина провайдера об'єднує в собі цілий комплекс мережевих пристроїв та компонентів, що складають окрему структуру мережі, які виконують не тільки функцію маршрутизації, але й здійснюють процес балансування, пріорітезації та управління трафіком. Кожна корпорація обирає сама провайдера cloud послуг, головне – чи зможуть канали провайдера забезпечити необхідну якість надання сервісу на усіх рівнях мережі.

Під якістю надання сервісу мається на увазі надання сервісу чи додатку з мінімальним часом затримки для кінцевого користувача. Далі ISP (Internet Service Provider – ISP) надсилає запит до хмарного центру обробки даних. ЦОД ініціалізує виділення фізичного сервера РМ (Physical Machine – РМ), який містить необхідний тип сервісу чи додатку (Рис. 2).

При надходженні декількох запитів на одну і ту ж РМ продуктивність таких сервісів зменшується, оскільки доступ сервісів до ресурсів фізичної машини відбувається за методом часового розділення. Як наслідок при збільшенні кількості сервісів на одній РМ та при збільшенні інтенсивності надходження запитів на дану РМ, а також при збільшенні інтенсивності надходження запитів на кожен сервіс збільшується тривалість обслуговування запитів. У випадку зменшення продуктивності сервісу система керування переносить цей сервіс на іншу VM чи РМ. В такому випадку загальний час передавання сервісу від користувача до ЦОД і в зворотньому напрямку розраховуватиметься:

$$t_{nep} = \sum_1^n t_{комут.} + \sum_1^{n-1} t_{н.к.з.} + t_{обр.} \cdot \quad (1)$$

де  $n$  – кількість запитів;

$t_{комут.}$  – час проходження запиту через систему комутації;

$t_{н.к.з.}$  – час пошуку каналів, по яких буде здійснюватися передавання;

$t_{обр.}$  – час обробки запиту, який є сумою часів обробки запиту сервісом, який складається з  $k$  атомарних сервісів:

$$t_{обр.} = \sum_1^k t_{a.c.} \cdot \quad (2)$$

Оскільки оптимальний шлях передавання змінюється, то це призводить до збільшення часу пошуку логічних, а часом і фізичних каналів, по яких буде здійснюватися передача –  $t_{п.к.з.}$ , що в свою чергу призведе до збільшення загального часу передачі. Для зменшення та оптимізації цього часу пропонуємо використовувати алгоритм пошуку шляху за критерієм мінімального часу проходження [7]. Даний алгоритм враховує пропускну здатність кожного каналу передачі, завантаженість кожної ланки мережі, сумарна затримка на інтерфейсах, надійність шляху, по якому буде надаватися сервіс кінцевому користувачеві.

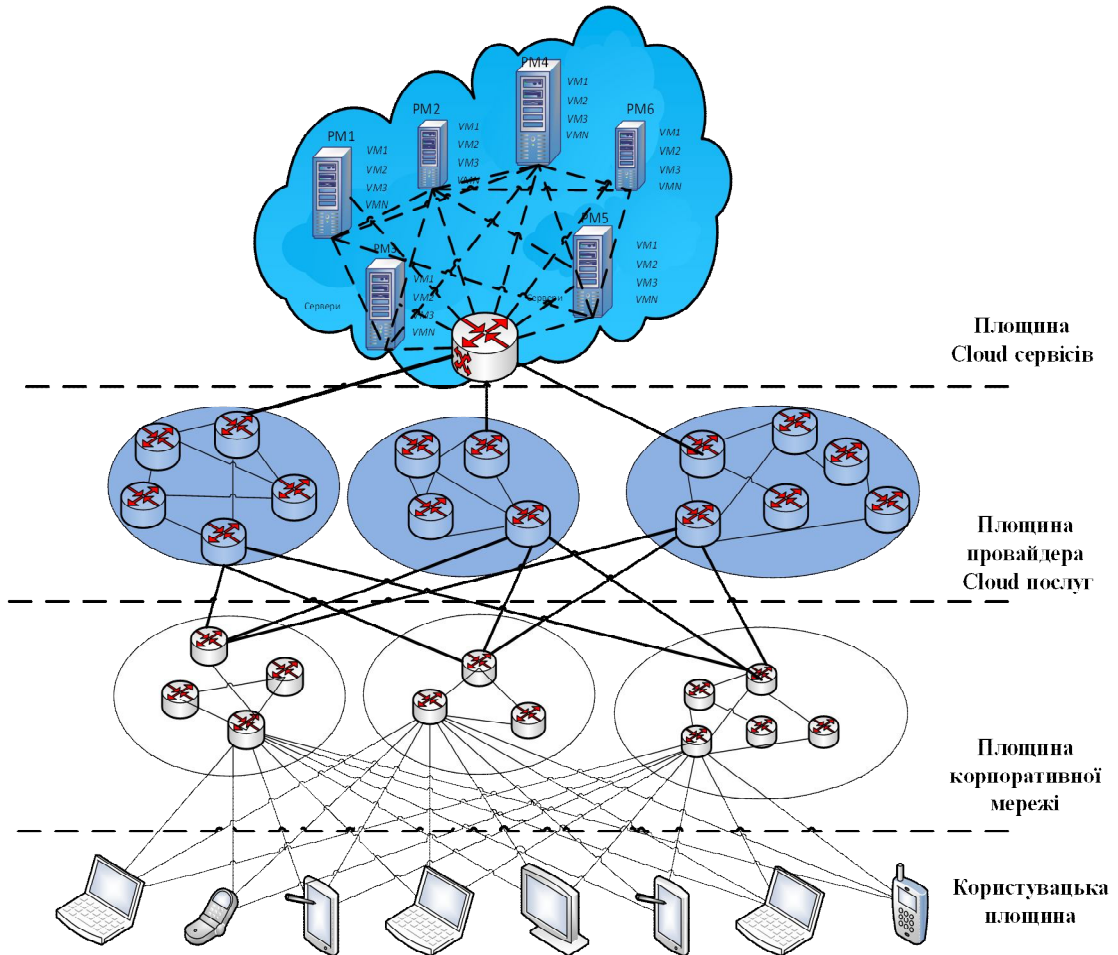


Рис. 2. Структура глобальної мережі з наданням cloud послуг

Затримка на інтерфейсах виникає внаслідок зміни топологічної структури мережі, де відбувається обслуговування і надання сервісу – ЦОД. Затримка може бути зменшена тільки у тому випадку, коли структура центру обробки даних буде стійкою на протязі часу  $T$ , тобто відбудеться логічна адаптація потоків запитів на зміну структури.

$$d = \frac{1}{P_{cm}} \sum_{Deg.n} \frac{w_T(n)}{C - w_T(n)} \leq d_{зад} \quad (3)$$

Така зміна буде відбуватися не тільки у площині надання сервісу, а й у площині провайдерів послуг. Всіма процесами у площині надання сервісів управляє гіпервізор [6 8]. Кожен провайдер повинен забезпечити користувачів якісним і швидким наданням послуги. Однак не треба забувати і про готовність вузлів мережі провайдера та й самих фізичних машин центру обробки даних до виконання запитів на надання сервісу. Під готовністю вузла будемо розуміти, що вузол перебуватиме в працездатному стані в довільний момент часу, коли він не зайнятий виконанням запитів. Працездатність вузла буде характеризуватися наявністю необхідних параметрів та потужностей процесорів, які оброблятимуть запити.

Параметри потужностей сервера, на якому розташовано  $M_i$  віртуальних машин розраховуються за формулами:

$$CPU_{pr} = \frac{\sum_{i=1}^k M_i \times CPU_i}{\sum_{i=1}^k M_i}, \quad (4)$$

$$RAM_{pr} = \frac{\sum_{i=1}^k M_i \times RAM_i}{\sum_{i=1}^k M_i}, \quad (5)$$

де  $M_i$  – кількість VM у  $i$ -тій PM;  
 $CPU_i$  – тактова частота процесора VM  $i$ -тої PM;  
 $RAM_i$  – об'єм оперативної пам'яті VM  $i$ -тої PM у;  
 $k$  – кількість PM cloud-системи.

Для оцінки стану ресурсів готовності вузла до обробки запитів пропонуємо наведений на Рис. 3 алгоритм.

Він дозволить в певній мірі відслідкувати процес міграції та на базі інформації про обчислювальні потужності кожного вузла пришвидшити час обробки запитів.

Стійкість структури (живучість) cloud-системи пропонуємо оцінити на основі методу, що включає в себе оцінку структури мережі (моделі мережі), завантаженості кожної віртуальної машини в момент часу  $T$ , коефіцієнту готовності вузла, а також взаємопов'язаність вузлів між собою [9]:

$$P_{cm} = 1 - \sum_{i=1}^n w_i(t) + K_{\Gamma} + P_{зв_i}, \quad (6)$$

де  $w_i(t)$  – завантаженість вузлів в момент часу  $t$ ;

$P_{зв_i}$  – ймовірність зв'язності вузлів;

$K_{\Gamma}$  – коефіцієнт готовності

Завантаженість вузлів до площини надання сервісу в момент часу  $t$  буде оцінюватися [10]:

$$T = \sum_{k=1}^K 2 \left( \sum_{k=1}^M \tau_{ik} + \sum_{j=0}^N \sum_{m=0}^{i-1} \rho \frac{\rho^{1-H}}{\rho^{1-H}} U_{викор(i-m)_j} \frac{l_{сер(i-m)_j} \times 8}{C_j} \right) \quad (7)$$

де  $M$  – загальне число каналів зв'язку між клієнтом і сервісом;

$k$  – порядковий номер запиту;

$\tau_{ik}$  – величина часу затримки розповсюдження пакету сервісу по  $k$ -каналу зв'язку;

$N$  – загальне число комутаційних пристроїв (вузлів) розміщених між клієнтом і сервісом;

$C$  – пропускна здатність  $j$ -го каналу;

$U_{викор(i-m)_j}$  – коефіцієнт використання каналу сервісом  $i$ -го пріоритету в  $j$ -му вузлі;

$l_{сер(i-m)_j}$  – середня довжина пакету послуги  $i$ -го пріоритету в  $j$ -му вузлі;

$\rho$  – коефіцієнт завантаження пристрою.

Загальна затримка надання сервісу кінцевому користувачеві буде визначатися як сума часів затримок до площини cloud-сервісів та часу затримки у сервісній площині, при реконфігурації структури центру обробки даних:

$$D = T + d. \quad (8)$$

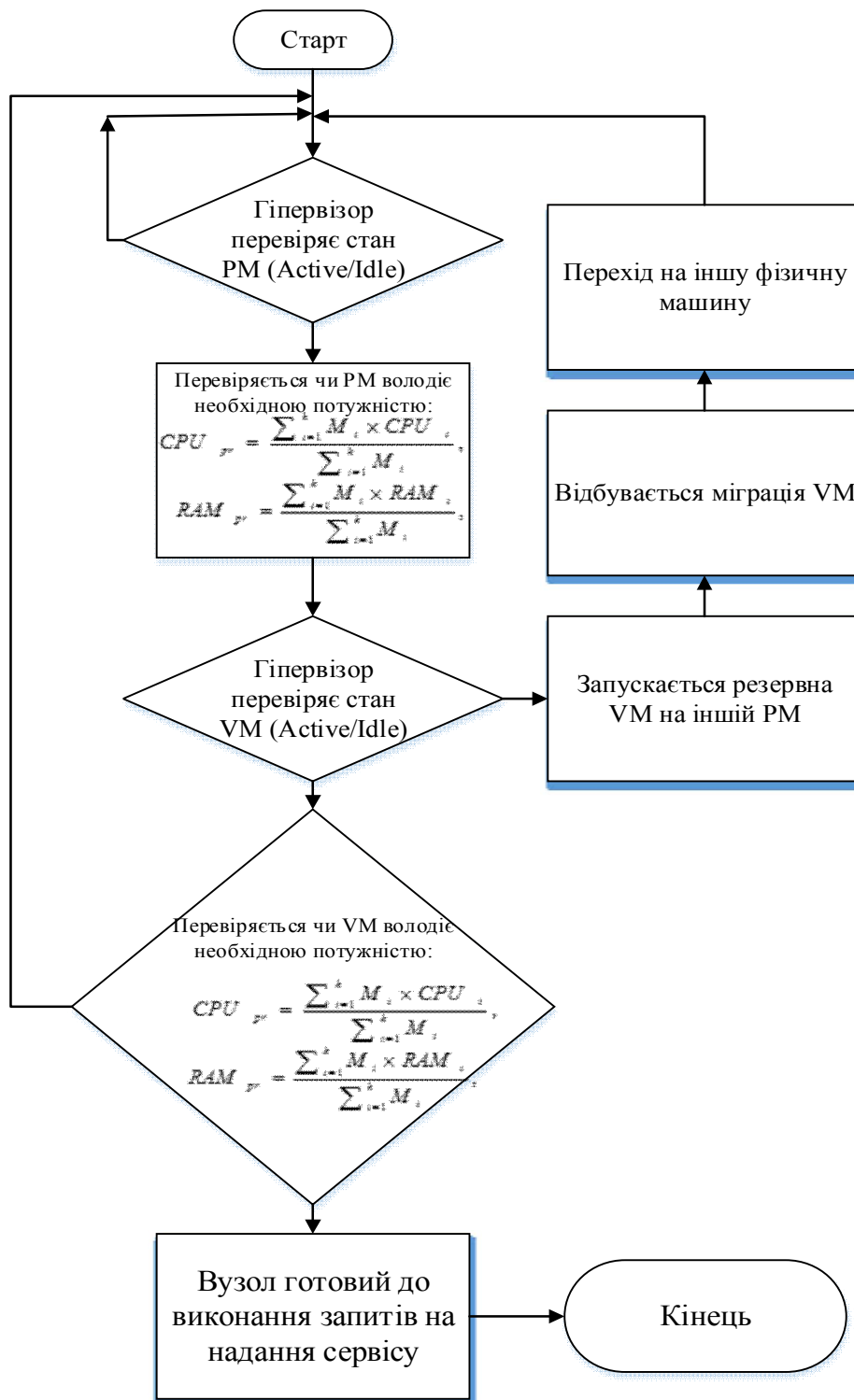


Рис. 3 Алгоритм оцінки стану готовності вузла

**Аналіз результатів моделювання.** Припустимо, що запити на надання сервісів до площини cloud надходять по найкоротшому шляху визначеному протоколом маршрутизації, наприклад, як на Рис. 4. Кожен із провайдерів надає свій окремо визначений тип сервісу [11, 12].

Як видно з Рис. 4 модель побудована так, щоб під час проходження пакету від джерела до адресата врахувати різні випадки коефіцієнтів завантаження вузлів. Прогнозування здійснювалось для трафіку створюваного домашньою групою користувачів, які користуються різними послугами в різному обсязі.

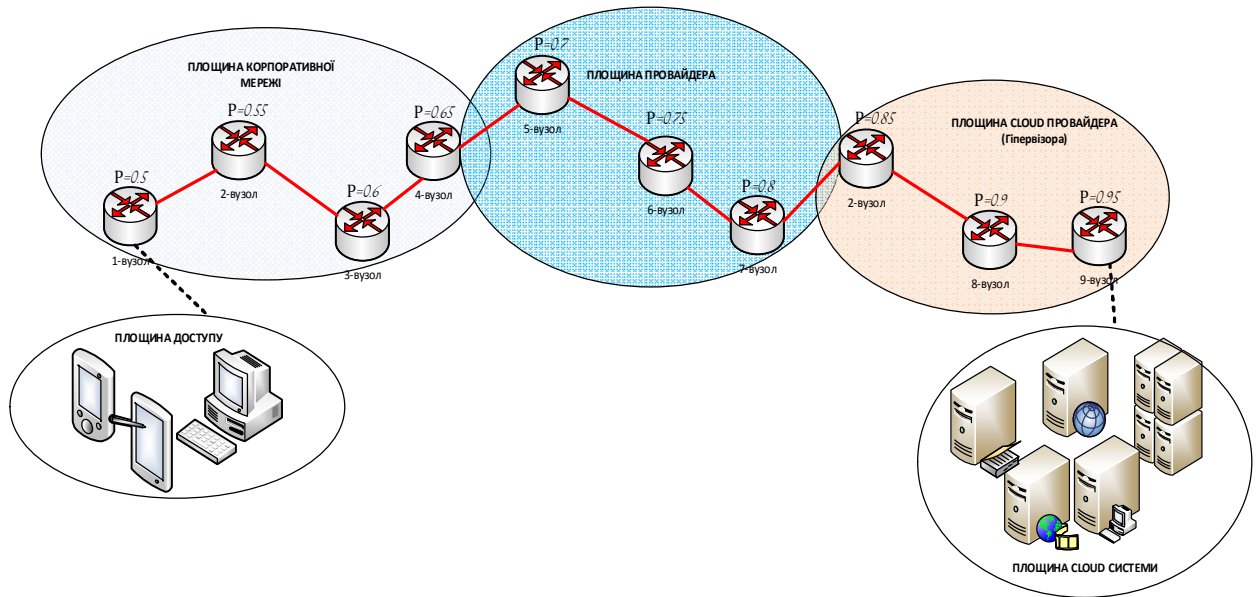


Рис. 4. Відправка запиту по визначеному найкоротшому маршруту

Для кожної із групи показано графіки прогнозування тривалості затримок (Рис. 5).

При реалізації даного алгоритму в програмному середовищі Matlab отримано залежності, які вказують на вплив оцінки стійкості структури мережі на затримку при обслуговуванні запитів та на час пошуку каналів, по яких буде здійснюватися передача. Для здійснення моделювання було прийнято, що один атомарний сервіс виконується віртуальною машиною  $t_{a.c.} = 0,005$  мс.

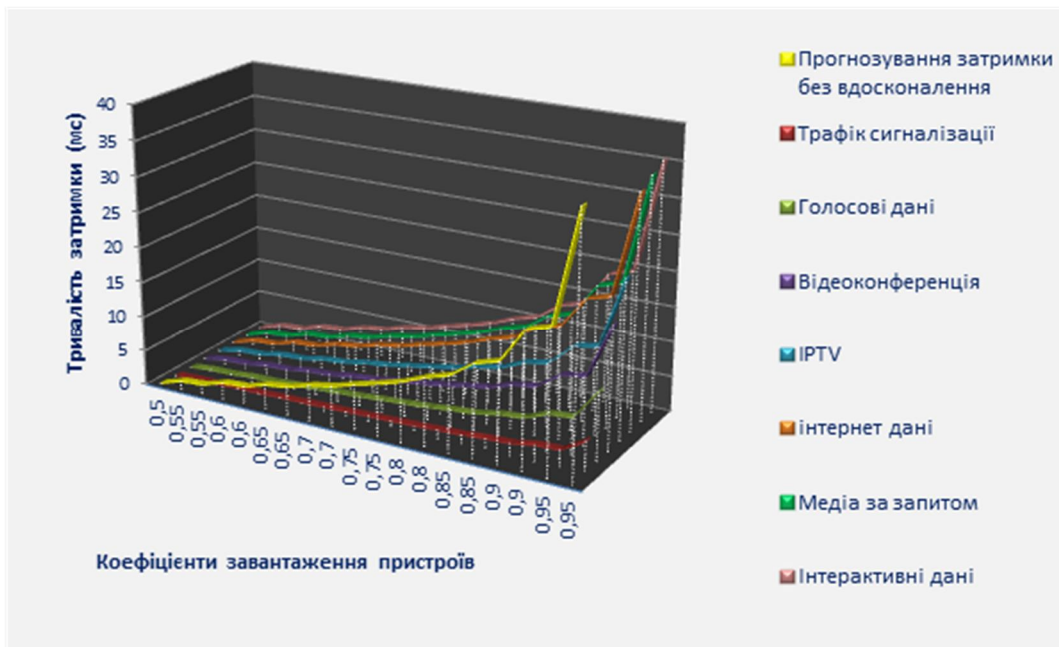


Рис. 5. Прогнозування тривалості затримок для домашніх користувачів

На Рис. 6 наведена залежність, яка показує, зменшення часу затримки при врахуванні додаткових умов, що призводить до пришвидшення процесу надання сервісу кінцевому користувачу. Під лінією позначеною 1 показано затримку після застосування вище описаних умов.

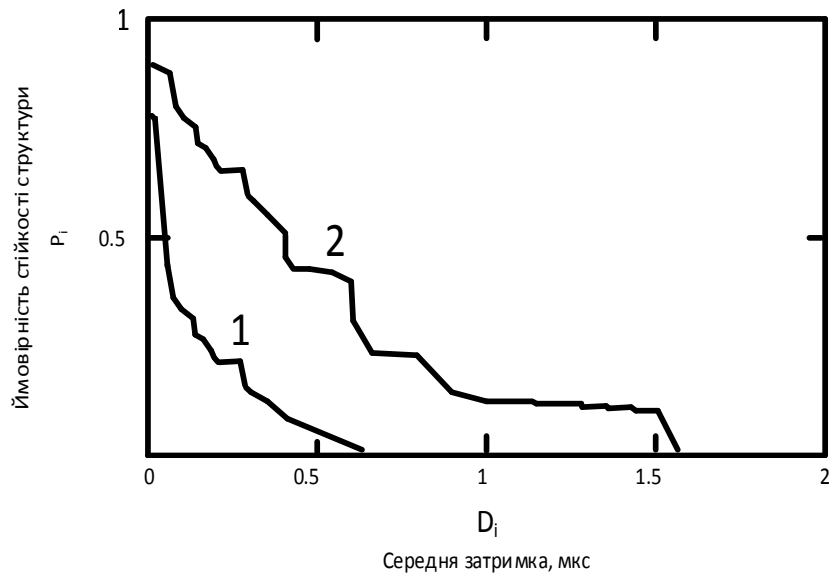


Рис. 6. Залежність ймовірності стійкості структури від середньої сумарної затримки

В результаті моделювання встановлено, що при зменшенні затримки час пошуку каналів, по яких буде здійснюватися передавання запитів, зменшується, що в результаті призведе і до зменшення загального часу передавання сервісу до ЦОД і назад.

В результаті отриманих залежностей можна говорити про загальне зменшення затримки (для домашніх користувачів) та підвищення якості надання сервісів користувачам cloud-мережі (Рис. 7).



Рис. 7. Прогнозування тривалості затримки пакетів голосового сервісу дослідженої мережі для домашніх користувачів з використанням запропонованого методу

**Висновки.** Хмарні обчислення є однією з найбільш прогресивних технологій розподілених обчислень, які надають користувачу можливість отримати всі необхідні ресурси. У роботі пропонується зменшити час надання сервісу, за допомогою методу адаптації логічних структур віртуалізованого ЦОД з врахуванням інваріантності запитів, що надходять до VM, коефіцієнту готовності вузлів. Це дасть змогу на підставі даних про адаптацію структури в конкретні моменти часу не здійснювати повторний перерахунок оптимального шляху передання, що призведе до зменшення затримки при наданні сервісу. Встановлено, що від стійкості структури та її переконфігурації залежить середня затримка, що дає змогу зменшити час надання сервісу. У подальших дослідженнях велику увагу буде



зосереджено на оптимізацію використання ресурсів кожної складової ЦОД, що в результаті дозволить ще більше вплинути на час обробки запитів у ЦОД

### **Література**

1. Ramesh B. Mobile applications in multimedia cloud computing / B. Ramesh, N. Savitha, A. E. Manjunath // *Int.J.Computer Technology & Applications*. – Jan-Feb 2013. – Vol 4(1). – P. 97-103.
2. Копейка О. В. Архитектура системы управления ИТ-инфраструктурой в современных Дата-центрах / О. В.Копейка // *Наукові записки Українського науково-дослідного інституту зв'язку*. – 2014. – № 1(29). – С.29-37.
3. Копейка О. В. Проектирование сервисов инфраструктуры приложений в дата-центрах / О. В. Копейка // *Телекомунікаційні та інформаційні технології*. – 2014. – № 1. – С. 19-27.
4. Klymash M. Model of network resources management on the basis of services priorities association / Mykhailo Klymash, Mykola Beshley, Orest Lavriv // *Proceedings of international conference CADSM'2013. Polyana-Svalyava*. – 2013. – P. 172-173.
5. Птицын Г. А. Вероятностные модели распада динамических сетей / Г. А. Птицын // *Электротехнические и информационные комплексы и системы*. – 2006. – Т. 2, №4.
6. Khazaei H. Performance of Cloud Centers with High Degree of Virtualization under Batch Task Arrivals / H. Khazaei, J. Mišić, and V. B. Mišić // *IEEE Transactions on Parallel and Distributed Systems*. – 2012. – Vol. 10, no. 5. – P. 1.
7. Стрихалюк Б. М. Алгоритми пошуку шляху за критерієм мінімальної затримки для центру обробки даних // Б. М. Стрихалюк, О. М. Шпур, М. О. Селюченко, Т. В. Андрухів // *Вісник Національного університету «Львівська політехніка». Радіоелектроніка та телекомунікації*. – 2014. – №796. – С.176-181.
8. Klymash M. Service Quality Oriented Method of Multiservice Telecommunication Networks Design / Mykhailo Klymash, Orest Lavriv, Bohdan Buhyl, Yuriy Danik // *11th International Conference Modern Problems of Radio Engineering, Telecommunications and Computer Science Dedicated to the 60th Anniversary of Radio Department at Lviv Polytechnic National University TCSET'2012. February 21-24, 2012 Lviv-Slavske, Ukraine*. – Publishing House of Lviv Polytechnic. – P. 235-236.
9. 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*. – 2014. – Vol. 4, №1. – P.83-88.
10. Стрихалюк Б. М. Віртуалізація мобільних систем зв'язку на основі технології NFV та моделей cloud-сервісів / Б. М. Стрихалюк, О. М. Шпур, А. Р. Масюк // *Сучасні проблеми телекомунікацій та підготовка фахівців в галузі телекомунікацій: матеріали конференції (30 жовтня – 02 листопада 2014 р., м. Львів)*. – 2014р. – С. 21-24.
11. Beshley M. Research and Development the Methods of Quality of Service Provision in Mobile Cloud Systems / M. Beshley, T. Maksymyuk, B. Strykhaluk, M. Klymash // *IEEE International Conference [Black Sea Conference on Communications and Networking (BlackSeaCom'2014)]*, May 27-30, 2014, Odessa, Ukraine. – 2014. – P. 165-169.
12. Klymash M. A Novel Approach of Optimum Multi-criteria Vertical Handoff Algorithm for Heterogeneous Wireless Networks / M. Klymash, B. Stryhaluk, I. Demydov, M. Beshley, M. Seliuchenko // *International Journal of Engineering and Innovative Technology (IJEIT)*. – November 2014. – Volume 4, Issue 5. – P. 42-52.