УДК 004.41

ВИКОРИСТАННЯ БЕЗСЕРВЕРНИХ FAAS ТЕХНОЛОГІЙ ЯК ЗАСІБ ОПТИМІЗАЦІЇ МІКРОСЕРВІСНИХ АРХІТЕКТУР ДЛЯ ВИСОКОНАВАНТАЖЕНИХ КОМП'ЮТЕРНИХ СИСТЕМ

Зах В.Ю., науковий керівник Коваленко О.Є.

У сучасній індустрії розробки програмного забезпечення спостерігається стрімке зростання попиту на високопродуктивні, масштабовані та економічно ефективні рішення для обробки значних обсягів даних та запитів. Мікросервісна архітектура стала стандартом для створення складних систем, проте навіть вона має певні обмеження при роботі з високими навантаженнями. В останні роки безсерверні технології, зокрема Function as a Service (FaaS), пропонують інноваційний підхід до вирішення цих проблем, змінюючи парадигму розгортання та масштабування додатків.

Безсерверні обчислення (Serverless) представляють модель виконання коду, де розробники зосереджуються виключно на написанні бізнес-логіки, тоді як хмарний провайдер автоматично керує всією інфраструктурою. На відміну від традиційних мікросервісів, що вимагають постійно працюючих серверів та ручного масштабування, FaaS-сервіси, такі як AWS Lambda, Google Cloud Functions або Azure Functions забезпечують автоматичне масштабування, оплату лише за фактично використані ресурси та зменшують операційне навантаження на команди розробників[1]. Такий підхід дозволяє оптимізувати використання ресурсів та суттєво знизити експлуатаційні витрати.

Основна перевага FaaS у високонавантажених системах полягає у їхній здатності миттєво масштабуватися для обробки пікових навантажень. Коли традиційні мікросервіси вимагають попереднього планування ресурсів та часто призводять до їх надмірного виділення, безсерверні функції автоматично створюють нові екземпляри для обробки кожного запиту паралельно. Це особливо цінно для систем з нерівномірним навантаженням, де пікові періоди чергуються з періодами низької активності. Безсерверний підхід забезпечує практично необмежене горизонтальне масштабування без необхідності ручного втручання, що критично важливо для високонавантажених систем.

Інтеграція FaaS технологій у мікросервісну архітектуру відкриває нові можливості для оптимізації. Замість розгортання цілісних мікросервісів для виконання окремих операцій, розробники можуть декомпозувати функціональність на менші, атомарні функції. Це дозволяє точніше розподіляти обчислювальні ресурси та оптимізувати продуктивність системи в цілому. Наприклад, операції обробки зображень, трансформації даних чи валідації можуть бути виділені в окремі функції, які викликаються асинхронно лише за необхідності. Такий підхід знижує навантаження на основні сервіси та підвищує загальну продуктивність системи.

Економічна ефективність є ще одним вагомим аргументом на користь безсерверних технологій. Модель оплати за фактичне використання (pay-per-execution) дозволяє організаціям оптимізувати витрати на інфраструктуру, оскільки ресурси не споживаються в періоди простою. Це особливо актуально для високонавантажених систем з нерівномірним розподілом активності. Крім того, відсутність необхідності управляти серверами знижує операційні витрати та дозволяє командам розробників зосередитися на створенні бізнес-цінності замість адміністрування інфраструктури. Безсерверний підхід також сприяє підвищенню надійності та стійкості до відмов. Хмарні провайдери забезпечують високий рівень доступності та автоматичне відновлення після збоїв. Кожна функція виконується в ізольованому середовищі, що мінімізує ризик каскадних відмов, характерних для монолітних та тісно пов'язаних мікросервісних

архітектур. У випадку збою окремої функції, інші компоненти системи продовжують працювати, що підвищує загальну надійність високонавантажених систем.

Незважаючи на численні переваги, впровадження безсерверних технологій має певні виклики. Холодний старт (cold start) — затримка при першому виклику функції — може негативно впливати на продуктивність систем, критичних до затримок. Обмеження часу виконання функцій (зазвичай до 15 хвилин) ускладнює реалізацію тривалих операцій. Дослідження показало, що правильна архітектура та оптимізація коду можуть мінімізувати вплив цих обмежень, дозволяючи досягти затримки відгуку менше 100 мс навіть при холодному старті AWS Lambda функцій[2].

Дослідження, яке буде проведено, націлене на порівняльний аналіз продуктивності та вартості використання традиційних мікросервісів, розгорнутих на EC2 інстансах, та безсерверних функцій AWS Lambda при різних сценаріях навантаження. Методологія дослідження передбачає створення однакових за функціональністю сервісів, реалізованих за допомогою обох підходів, та тестування їх під декількома рівнями навантаження. Буде вимірюватися час відгуку, використання ресурсів та загальна вартість експлуатації для кожного сценарію.

Особлива увага в дослідженні буде приділена аналізу точок перетину вартості – порогів навантаження, при яких один підхід стає економічно вигіднішим за інший. Передбачається, що для систем з нерівномірним навантаженням та періодами низької активності безсерверні рішення будуть більш економічними, тоді як для постійно завантажених систем традиційні мікросервіси можуть демонструвати кращу вартісну ефективність.

Практичним результатом дослідження стане створення моделі прийняття рішень, яка допоможе архітекторам та розробникам високонавантажених систем визначити оптимальну стратегію використання безсерверних технологій залежно від характеристик навантаження та бюджетних обмежень. Така модель дозволить розробникам приймати обґрунтовані рішення щодо архітектури їхніх систем, максимізуючи продуктивність та мінімізуючи витрати.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Eismann, S., Scheuner, J., Van Eyk, E., Schwinger, M., Grohmann, J., Herbst, N., Abad, C.L. and Iosup, A., (2020). Serverless applications: Why, when, and how?. IEEE software, 38(1), pp.32-39.

2. Arifin, M. A., Satra, R., Syafie, L., & Nidhom , A. M. (2023). Optimizing AWS lambda code execution time in amazon web services. Bulletin of Social Informatics Theory and Application, 7(1), 14–23.

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ АСПЕКТИ РОЗРОБКИ КОМП'ЮТЕРНИХ СИСТЕМ

ЗБІРНИК НАУКОВИХ ПРАЦЬ ЗА МАТЕРІАЛАМИ
VII ВСЕУКРАЇНСЬКОЇ НАУКОВО-ПРАКТИЧНОЇ КОНФЕРЕНЦІЇ
СТУДЕНТІВ І АСПІРАНТІВ
24 квітня 2025 року

Київ 2025

УДК 004

Відповідальна за випуск: М.І. Лендєл

Збірник наукових праць за матеріалами VII Всеукраїнської науково-практичної конференції студентів і аспірантів «ТЕОРЕТИЧНІ ТА ПРИКЛАДНІ АСПЕКТИ РОЗРОБКИ КОМП'ЮТЕРНИХ СИСТЕМ '2025», 24 квітня 2025 року, НУБіП України, Київ. – 452 с. (електронне видання)

Відповідальність за зміст публікацій несуть автори.

Передрук матеріалів, а також використання їх будь-якій формі допускається лише з дозволу авторів

© Національний університет біоресурсів
і природокористування України, 2025