

УДК 004.75:510.644.4

ОПТИМІЗАЦІЯ РОЗПОДІЛУ РЕСУРСІВ У ХМАРНИХ СЕРВЕРНИХ СИСТЕМАХ НА ОСНОВІ МЕТОДІВ НЕЧІТКОЇ ЛОГІКИ

Чабаненко А.О.

e-mail: andrii.chabanenko@nure.ua

Науковий керівник – ст. викл. Путятіна О.Є.

Харківський національний університет радіоелектроніки, каф. ІНФ
м. Харків, Україна

This research addresses inefficient resource utilization and SLA non-compliance caused by volatile traffic in cloud environments. We propose a fuzzy logic-based control mechanism that evaluates real-time CPU and memory metrics to optimize scaling decisions. Unlike rigid threshold-based strategies, the fuzzy inference system provides fluid transitions, minimizing instability in virtual machine management. Simulations demonstrate that this approach boosts energy efficiency by 15% while sustaining high availability.

Розвиток парадигми хмарних обчислень вимагає від провайдерів забезпечення високої доступності сервісів при одночасній мінімізації операційних витрат. Сучасні хмарні дата-центри стикаються з проблемою динамічної зміни навантаження, яка має стохастичний характер. Традиційні стратегії управління, що базуються на статичних порогах (threshold-based), демонструють низьку ефективність у випадках різких сплесків трафіку. Надлишкове резервування потужностей (over-provisioning) призводить до неефективного використання енергії, тоді як недовантаження (under-provisioning) спричиняє деградацію продуктивності та порушення умов SLA. Використання апарату нечіткої логіки дозволяє моделювати процеси прийняття рішень в умовах невизначеності, забезпечуючи плавну адаптацію обчислювальних ресурсів до поточних потреб.

Метою дослідження є розробка інтелектуальної системи управління, яка б забезпечувала баланс між продуктивністю та витратами. Для цього необхідно розрахувати оптимальний коефіцієнт зміни ресурсів:

$$K_{res} \in [0,1],$$

де значення 0 відповідає мінімальному виділенню потужностей, а 1 – максимальному

Процес прийняття рішень реалізовано за алгоритмом нечіткого виведення Мамдані. На етапі фазифікації вхідні чіткі дані x_1 (завантаження CPU, %) та x_2 (заповненість RAM, %) перетворюються у нечіткі терм-множини $T = \{\text{Low, Normal, High, Critical}\}$. Для забезпечення плавності керування та уникнення різких коливань при масштабуванні використовуються гаусові функції належності:

$$\mu_A(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right),$$

де σ -центр функції (математичне сподівання), σ -коефіцієнт концентрації (середньоквадратичне відхилення).

Формування бази знань базується на експертному досвіді управління серверною інфраструктурою. Система правил дозволяє адекватно реагувати на критичні стани та оптимізувати роботу в періоди низької активності. Приклад логічних конструкцій:

1. IF (CPU is Critical) AND (RAM is High) THEN (Action is Urgent_Scale_Up);
2. IF (CPU is Low) AND (RAM is Low) THEN (Action is Scale_Down);
3. IF (CPU is Normal) THEN (Action is No_Change).

Завершальним етапом є дефазифікація — перетворення нечіткої множини у конкретне числове значення керуючого сигналу. Для цього використано метод центру площі (Centroid method):

$$S_{out} = \frac{\sum_{i=1}^n \mu(y_i) y_i}{\sum_{i=1}^n \mu(y_i)}$$

Моделювання проводилося на базі фреймворку CloudSim. Було сконструйовано датацентр з 10 фізичними хостами та 50 віртуальними машинами. Навантаження генерувалося згідно з реальним профілем трафіку сервісу з добовими піками. Результати експерименту продемонстрували наступні показники ефективності порівняно з класичними методами:

Таблиця 1 – Показники ефективності системи

Метрика	Статичний метод	Пороговий метод	Нечітка логіка
Середня утилізація ЦП	45%	68%	82%
Економія енергії	0%	18%	27%
Кількість порушень SLA	0.5%	4.2%	1.1%

Результати продемонстрували, що нечіткий контролер підвищує утилізацію ЦП до 82% та забезпечує економію енергії на рівні 27%. Використання плавних функцій належності дозволило мінімізувати ефект «пінг-понгу» при міграції віртуальних машин.

Список використаних джерел:

1. Zadeh L. A. Fuzzy Sets. *Information and Control*. 1965. Vol. 8, No. 3. P. 338–353.
2. Buyya R., Broberg J., Goscinski A. *Cloud Computing: Principles and Paradigms*. Hoboken: Wiley, 2011. 664 p.
3. Passino K. M., Yurkovich S. *Fuzzy Control*. Menlo Park: Addison-Wesley, 1998. 522 p.
4. Dhiman R., Gupta R. K. Resource management in cloud computing using fuzzy logic: A review. *Journal of King Saud University – Computer and Information Sciences*. 2022. Vol. 34, No. 6. P. 3328–3342