УДК 004.75:004.421.2

БАГАТОКРИТЕРІАЛЬНА МОДЕЛЬ РОЗПОДІЛУ НАВАНТАЖЕННЯ МІЖ ДАТА-ЦЕНТРАМИ ІНТРАНЕТ-МЕРЕЖ

Ільїн С. А.
Інститут програмних систем Національної академії наук України
Україна, Київ
sergey.ilin.gsmart@gmail.com

Анотація. *Стрімке зростання обсягів даних і розвиток розподілених обчислень потребують удосконалення механізмів управління навантаженням між центрами обробки даних у межах інтранет-мереж. Зміна топології, коливання трафіку та відмови вузлів ускладнюють забезпечення стабільної продуктивності систем. Класичні алгоритми балансування, такі як Round Robin чи Least Connections, не враховують динаміку мережевих характеристик, що призводить до неефективного використання ресурсів. У роботі запропоновано адаптивну багатокритеріальну модель розподілу навантаження, яка враховує затримки передавання даних, завантаження вузлів, енергоспоживання та надійність каналів зв'язку. Модель реалізує принцип динамічного оновлення параметрів у процесі роботи системи та оптимізує баланс між швидкодією, енергоефективністю й стійкістю мережі. Імітаційне моделювання продемонструвало зниження середньої затримки на 30% і підвищення енергоефективності на 15–18% порівняно з класичними підходами. Отримані результати підтверджують перспективність впровадження розробленої моделі в програмно-визначених мережах (SDN) та корпоративних дата-центрах із динамічною структурою.*

Ключові слова: *балансування навантаження; дата-центр; інтранет-мережа; багатокритеріальна оптимізація; адаптивна модель; енергоефективність; SDN.*

Вступна частина. Підвищення обчислювальної складності сучасних інформаційних систем і динамічне збільшення обсягів даних висувають нові вимоги до механізмів балансування навантаження між центрами обробки даних в інтранет-середовищах. Забезпечення стабільної продуктивності, мінімізації затримок та оптимального використання енергоресурсів стає критичним завданням для корпоративних і державних IT-інфраструктур. У сучасних умовах динамічної топології мереж — спричиненої масштабуванням сервісів, віртуалізацією ресурсів і коливаннями трафіку — класичні алгоритми балансування, такі як Round Robin, Least Connections або Weighted Response Time, втрачають ефективність, оскільки не враховують реальні зміни стану мережі та пропускної здатності каналів [1–3].

Актуальні дослідження [4–6] демонструють, що адаптивні підходи до балансування навантаження, які враховують параметри мережевої динаміки та енергоспоживання, дозволяють істотно підвищити ефективність розподілу ресурсів. Зокрема, активно розвиваються методи network-aware та energy-aware load balancing, що використовують аналітичні моделі оптимізації і алгоритми машинного навчання для прогнозування стану мережі. Проте більшість існуючих рішень зосереджена або на продуктивності, або на енергетичній ефективності, тоді як узгодження кількох суперечливих критеріїв залишається відкритою проблемою.

Таким чином, постає завдання створення багатокритеріальної моделі розподілу навантаження між дата-центрами інтранет-мереж, яка враховує часові затримки, завантаження

вузлів, пропускну здатність каналів і енергоспоживання. Такий підхід дозволить забезпечити баланс між швидкодією, енергоефективністю та стійкістю до змін топології.

Метою роботи є розроблення математичної моделі багатокритеріальної оптимізації розподілу навантаження між дата-центрами інтранет-мереж, що враховує динаміку топології, затримки передавання даних та енергоспоживання, з метою підвищення ефективності функціонування розподілених обчислювальних систем.

Завдання дослідження

1. Проаналізувати сучасні моделі та алгоритми розподілу навантаження в інтранет-мережах і визначити фактори, що впливають на їх ефективність у динамічних середовищах.
2. Визначити ключові параметри, які впливають на продуктивність дата-центрів — затримку передавання даних, пропускну здатність, енергоспоживання та надійність каналів.
3. Розробити багатокритеріальну математичну модель розподілу навантаження між дата-центрами з урахуванням змін топології мережі.
4. Реалізувати імітаційний експеримент для аналізу стійкості системи при зміні трафіку, кількості вузлів і пропускної здатності.
5. Провести порівняльну оцінку запропонованої моделі з традиційними алгоритмами (Round Robin, Least Connections, Weighted Response Time) та визначити потенціал її практичного застосування.

Основна частина.

Ключові параметри впливу на ефективність системи. Для формалізації процесу балансування визначено чотири групи параметрів, що найбільше впливають на продуктивність дата-центрів (табл. 1).

Таблиця 1.

Визначальні параметри ефективності системи

Позначення	**Параметр**	**Одиниця вимірювання**	**Характер впливу**
τ_{ij}	Середня затримка передавання запиту	мс	Визначає швидкодію системи
η_j	Коефіцієнт завантаження дата-центру	%	Відображає баланс використання ресурсів
R_j	Потужність споживання енергії вузлом	кВт	Характеризує енергоефективність

Залежно від умов роботи системи ваги впливу цих параметрів можуть змінюватися, що робить задачу динамічною.

Формування адаптивної багатокритеріальної моделі/ Запропонована модель оптимізації розглядає задачу як мінімізацію комбінованої функції витрат:

$$Z = \alpha \cdot T + \beta \cdot E + \gamma \cdot (1 - R),$$

де $T -$ середня затримка обробки запитів;

$E -$ питомі енергетичні витрати;

$R -$ коефіцієнт надійності мережі;

$\alpha, \beta, \gamma -$ вагові коефіцієнти важливості критеріїв.

На відміну від статичних підходів, у цій моделі параметри оновлюються на кожному кроці моделювання залежно від змін у матриці зв’язності мережі:

$$C(t) = [c_{jk}(t)],$$

де $c_{jk}(t) \in [0,1]$ визначає поточну доступність зв'язку між вузлами d_j і d_k.

Алгоритм прийняття рішення про перерозподіл навантаження виконується за принципом адаптивної корекції: якщо упродовж інтервалу часу середня затримка перевищує поріг $T_{\max}$, трафік перенаправляється до вузла з мінімальним відношенням

$$\frac{\tau_{ij}}{\eta_j \cdot R_j}.$$

Імітаційна модель та результати дослідження. Для експерименту змодельовано мережу з чотирьох дата-центрів, з'єднаних каналами із затримками 25–80 мс та пропускною здатністю 200–700 Мбіт/с. Динамічні зміни топології реалізовано через випадкове відключення вузлів з імовірністю 0,1 на інтервал.

Оптимізація виконувалася методом Adaptive Evolutionary Load Balancing (AELB), який поєднує стохастичний вибір маршрутів із корекцією вагових коефіцієнтів за результатами попередніх ітерацій.

Таблиця 2.

Порівняльна характеристика алгоритмів балансування

Алгоритм	**Середня затримка, мс**	**Середнє завантаження вузлів, %**	**Енергоспоживання, кВт**	**Надійність, відн. од.**	**Інтегральний показник**
Round Robin	57	83	4,2	0,88	0,59
Least Connections	49	77	4,0	0,90	0,64
Weighted Response	45	73	3,8	0,91	0,67
AELB (запроп.)	36	68	3,3	0,94	0,76

Результати показують, що нова модель дозволяє знизити затримку на **близько 30%**, покращити енергоефективність на **понад 15%** і підвищити надійність каналів на **5–7%** у порівнянні з традиційними алгоритмами.

Практичне застосування моделі. Запропонована стратегія може бути реалізована у програмно-визначених мережах (SDN) або у системах управління корпоративними дата-центрами. Вона забезпечує автоматичну реакцію на зміни трафіку, відмови вузлів і зменшення енергоспоживання без втрати пропускної здатності.

Модель може бути використана як частина інтелектуального контролера балансування у хмарних і гібридних середовищах, що потребують високої надійності та мінімальних затримок у критичних системах.

Висновки. У результаті проведеного дослідження розроблено адаптивну багатокритеріальну модель розподілу навантаження між дата-центрами інтранет-мереж, що враховує змінність топології та динаміку трафіку. На відміну від класичних алгоритмів (Round Robin, Least Connections, Weighted Response Time), запропонований підхід забезпечує комплексну оптимізацію кількох критеріїв — часу затримки, енергоспоживання та надійності каналів зв'язку.

Імітаційне моделювання показало, що використання адаптивної еволюційної стратегії (AELB) дозволяє зменшити середню затримку передавання даних приблизно на 30%,

підвищити енергоефективність на 15–18% та покращити рівномірність завантаження вузлів до 10% у порівнянні з традиційними методами балансування.

Запропонована модель є стійкою до топологічних змін, здатною до самоналаштування у режимі реального часу та може бути впроваджена в системах програмно-визначених мереж (SDN), корпоративних дата-центрах і гібридних хмарних інфраструктурах. Подальші дослідження планується спрямувати на інтеграцію методів машинного навчання для прогнозування навантаження та автоматичного оновлення параметрів моделі у процесі роботи мережі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Gao, W., Wang, H., Zhang, X., & Chen, J. (2023). *PDLB: Path diversity-aware load balancing with adaptive flowcell granularity under asymmetric topology. Journal of Cloud Computing, 12*(1). https://doi.org/10.1186/s13677-023-00548-x

[2] Liao, X., Bao, B., Cui, W., & Liu, D. (2024). *Load balancing and topology dynamic adjustment strategy for power information system network: A deep reinforcement learning-based approach. Frontiers in Energy Research, 11*, Article 1342854. https://doi.org/10.3389/fenrg.2023.1342854

[3] Waghmode, S. (2024). *Adaptive load balancing in distributed cloud environment. Information and Digital Technologies Review.* https://doi.org/10.3233/IDT-240672

[4] Toka, L., Rossi, D., & Colle, D. (2022). *A survey on adaptive load balancing in data centers: Algorithms, architectures, and future trends. IEEE Communications Surveys & Tutorials, 24*(3), 1587–1614. https://doi.org/10.1109/COMST.2022.3148159

[5] Wang, Q., Lu, Y., & Jiang, Z. (2021). *Multi-objective optimization of load balancing in software-defined data center networks. Computer Networks, 197,* Article 108309. https://doi.org/10.1016/j.comnet.2021.108309

[6] Поперешняк, С. В., Панченко, С. В., Федорченко, О. Д., & Ільїн, С. А. (2024). *Особливості обробки та збереження даних за допомогою віртуальної файлової системи. Проблеми програмування, (2–3)*, 334–342. https://doi.org/10.15407/pp2024.02-03.334

Ilin S.A.

Multi-criterion load distribution model between data centers and intranet networks

Abstract. *The rapid growth of data volumes and the evolution of distributed computing highlight the need for efficient load balancing mechanisms across data centers within intranet networks. Dynamic topology changes, fluctuating traffic, and node failures complicate the maintenance of stable system performance. Traditional algorithms such as Round Robin and Least Connections do not account for network dynamics, resulting in inefficient resource utilization. This paper proposes an adaptive multi-objective load balancing model that considers data transmission delay, node utilization, energy consumption, and link reliability. The model dynamically updates optimization parameters in real time, maintaining a balance between performance, energy efficiency, and network resilience. Simulation experiments demonstrated a 30% reduction in average latency and a 15–18% improvement in energy efficiency compared to classical methods. The results confirm the applicability of the proposed model in software-defined networking (SDN) environments and enterprise data centers with variable topology.*

Keywords: *load balancing, data center, intranet network, multi-objective optimization, adaptive model, energy efficiency, software-defined networking.*