

МОДЕЛЬ ЧЕРГ КАФКА ОБРОБКИ ПОДІЙ У РЕАЛЬНОМУ ЧАСІ

Маньківський В.Б., Гриценко Д.Г.

Навчально-науковий інститут телекомунікаційних систем

КПІ ім. Ігоря Сікорського, Україна

E-mail: v.b.mankivskiy@gmail.com, grytsenko.danyl@iit.kpi.ua

KAFFKA QUEUES MODEL FOR REAL TIME EVENT PROCESSING

This paper considers the problem of efficient message queue management in the Apache Kafka system, which is used for processing streaming data in distributed environments. A mathematical model based on the theory of M/M/c/K queues is proposed, which allows estimating the probability of message loss (p_{loss}) depending on the size of the Kafka buffer (B). The impact of this parameter on system performance is analyzed. The simulation results demonstrate that increasing the buffer size leads to an exponential decrease in the probability of loss, but after a certain threshold, a saturation effect is observed.

Apache Kafka широко використовується для обробки подій у реальному часі в таких галузях, як фінансові технології, IoT та телекомунікації. Однак продуктивність системи значною мірою залежить від правильної конфігурації параметрів черг. У даній роботі ми дослідили вплив розміру буфера B на ймовірність втрат повідомлень та запропонували оптимізаційні підходи.

Проблематика дослідження. У багатьох реальних сценаріях використання Kafka (наприклад, фінансові транзакції, моніторинг IoT-пристроїв, обробка потоків даних у реальному часі) надходження повідомлень є нерівномірним і може значно перевищувати можливості їх обробки. Це може призводити до переповнення черг, втрати критично важливих даних та зниження продуктивності системи.

Основними проблемами, які досліджуються у цій роботі, є:

- Визначення оптимального розміру буфера (B), який мінімізує втрати повідомлень, не перевантажуючи пам'ять системи.
- Дослідження впливу різних параметрів Kafka (кількість консьюмерів, швидкість обробки) на ефективність обробки повідомлень в черзі.
- Аналіз впливу ефекту насичення на продуктивність Kafka.

Для аналізу вище зазначених проблем, перше за все, необхідно визначитися з моделлю, яка би змогла описати поставлені проблеми і підштовхнула до їх вирішення.

Математична модель черги у Kafka. Система чергування у Kafka може бути змодельована як система масового обслуговування M/M/c/K, де:

- λ - інтенсивність надходження повідомлень (подій за секунду).
- μ - швидкість обробки повідомлень одним консьюмером.
- c - кількість консьюмерів у Consumer Group.
- B - розмір буфера Kafka (у кількості повідомлень).
- P_{loss} - ймовірність втрати повідомлення через переповнення черги.

Модель M/M/c/K є класичною в теорії масового обслуговування і найкраще підходить для опису черг у Kafka через такі особливості:

1. Poisson-розподіл надходження повідомлень: Потоки даних у Kafka часто мають випадковий характер, тому припущення про випадкові часові інтервали між подіями відповідає реальним сценаріям використання.

2. Експоненційний розподіл часу обробки: Консьюмери можуть обробляти події з випадковою затримкою, що характерно для багатьох потокових сервісів.

3. Обмежена довжина черги (B): Kafka має обмежений розмір буфера, після чого нові повідомлення втрачаються. Ймовірність втрат розраховується за формулою Ерланга (1):

$$P_{loss} = \frac{\left(\frac{\lambda}{\mu}\right)^B / B!}{\sum_{n=0}^B \left(\frac{\lambda}{\mu}\right)^n / n!} \quad (1)$$

Умова неперервності полів на границі резонатор – вакуум при означає рівність тангенціальних складових відповідних поперечних компонент магнітного та електричного полів. В результаті було отримане наступне характеристичне рівняння, яке використовувалось для побудови залежності, показаної на рис. 2:

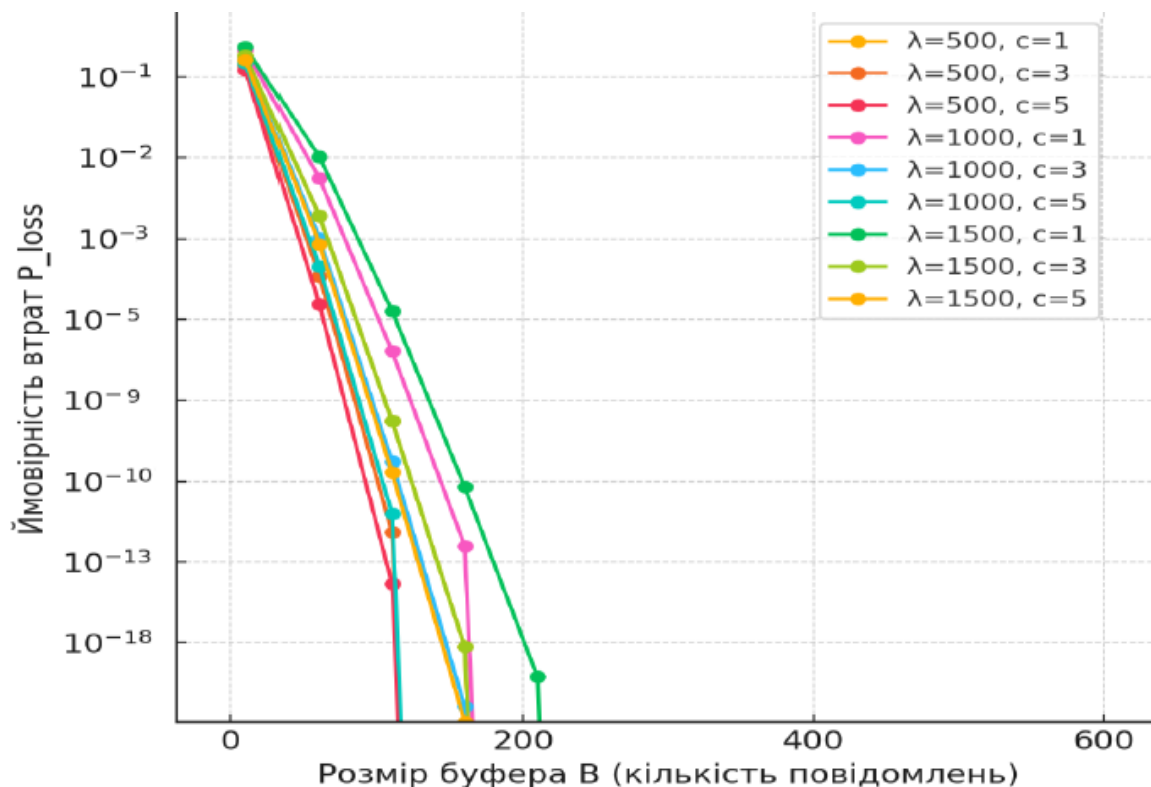


Рис.1. Залежність ймовірності втрат повідомлень (P_{loss}) від розміру буфера Kafka (B).

Рисунок 1 демонструє залежність ймовірності втрат повідомлень P_{loss} від розміру буфера Kafka B для різних значень інтенсивності надходження повідомлень (λ) та кількості консьюмерів (c). Як видно, збільшення B

суттєво знижує втрати, проте після певного рівня ефект насичення стає помітним.

Графіки показали наступні тенденції:

- При малих значеннях B (10-50) втрати повідомлень перевищують 15%, особливо при високих значеннях λ .
- Збільшення B до 200 суттєво зменшує втрати, досягаючи значень $<1\%$.
- Подальше збільшення B понад 500 майже не впливає на втрати, але може підвищувати середній час затримки.
- Зі збільшенням кількості консьюмерів (c) затримка зменшується, але при $c > 5$ ефект покращення продуктивності стає незначним.

Окреслені перспективні напрямки досліджень Kafka черг через математичну модель. Розраховані параметри дають змогу зробити висновок, що збільшення буфера (B) суттєво знижує ймовірність втрат повідомлень, однак надмірне його збільшення не дає значного покращення після певного порогу. Оптимізація розміру буфера повинна враховувати інтенсивність надходження подій (λ) і продуктивність консьюмерів (μ). Використання математичної моделі $M/M/c/K$ дає змогу більш точно прогнозувати продуктивність Kafka та уникати надмірних втрат повідомлень у реальному часі. Отримані результати можуть бути використані для розробки ефективних алгоритмів керування чергами у реальних розподілених системах.

Література

1. Jun Rao Jay Kreps, Neha Narkhede. Kafka: A distributed messaging system for log processing. In Proceedings of 6th International Workshop on Networking Meets Databases (NetDB), Athens, Greece, pages 1–7, 2011.
2. Kafka Inside Keystone Pipeline. <https://medium.com/netflix-techblog/kafka-inside-keystone-pipeline-dd5aeabaf6bb/>. [Online; accessed 25-Jun-2018].
3. Romanov, O., Nesterenko, M., Mankivskyi, V., Zhuk, O. Principles of Building Modular Control Plane in Software-Defined Network//Lecture Notes in Networks and Systems, 2023, 548, страницы 333–355. https://link.springer.com/chapter/10.1007/978-3-031-16368-5_17
4. G. Wang, J. Koshy, S. Subramanian, K. Paramasivam, M. Zadeh, N. Narkhede, et al., "Building a replicated logging system with apache kafka", Proc. VLDB Endow., vol. 8, no. 12, pp. 1654-1655, Aug. 2015, [online] Available: <http://dx.doi.org/10.14778/2824032.2824063>.
5. P. Helland, "Immutability changes everything", Queue, vol. 13, no. 9, pp. 40:101-40:125, Nov. 2015, [online] Available: <http://doi.acm.org/10.1145/2857274.2884038>.
6. O. Romanov, M. Nesterenko, and V. Mankivskyi, "The usage of regress model coefficient utilization of channels for creating the load distribution plan in network" in visnyk ntuu kpi seriia-radiotekhnika radioaparotobuduvannia, 2016, pp. 34-42.