

ПОРІВНЯННЯ МЕТОДІВ БАЛАНСУВАННЯ НАВАНТАЖЕННЯ

Ковріжний О.В.

Науковий керівник – д.т.н., проф. Кривуля Г.Ф.

Харківський національний університет радіоелектроніки

(61166, Харків, пр. Науки, 14, каф. АПОТ, тел. (057) 702-13-26)

e-mail: oleksii.kovrizhnyi@nure.ua, тел. 380979836976

This paper focuses on the problem of load balancing network traffic. Review of existing methods of load balancing, identifying their strength and weaknesses.

З проблемою балансування навантаження зустрічається будь-який веб-проект, бо відмова роботи серверу через велику кількість вхідних запитів може відбутися несподівано та призвести до небажаних матеріальних втрат. Стійкість серверу до навантаження можна вирішувати різними шляхами: будь-то зміна апаратної складової чи оптимізацією використовуваних алгоритмів та програмних кодів. Великі веб-проекти вирішують цю проблему шляхом кластеризації декількох серверів, навантаження між якими розподіляється за допомогою спеціальних методик з використання, як апаратних, так і програмних інструментів. Мета дослідження – визначити можливі переваги та недоліки балансування навантаження на різних рівнях мережевої моделі OSI. Задача – полягає у порівнянні існуючих методів балансування навантаження, можливості їх комбінування та визначення області використання.

Балансування навантаження - це метод для розподілу роботи між декількома обчислювальними ресурсами, такими як комп'ютери, кластери, мережі, центральні процесори та диски (носії інформації). Мета балансування полягає у оптимізації використання ресурсів, максимізація пропускної здатності, мінімізація часу відповіді та запобігання перенавантаження.

Рішення з балансування навантаження часто поділяють на дві категорії: L4 та L7, які відповідають транспортному та прикладному рівням мережевої моделі OSI. Стосовно категорії L4 балансування відбувається шляхом TCP-з'єднання клієнта з балансувальником, який в свою чергу завершає з'єднання, обирає найбільш доступний сервер та встановлює з ним зв'язок. Таким чином всі маніпуляції відбуваються на рівні з'єднання/сеансу протоколів TCP/UDP.

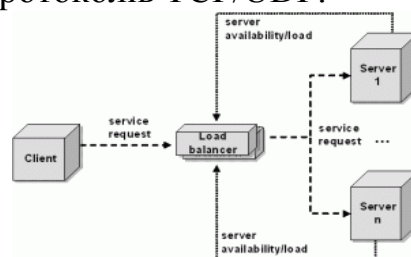


Рисунок 1 – Класична архітектура балансування навантаження

Слід зазначити, що балансувальник на рівні L4 робить одне вихідне TCP-з'єднання для кожного вхідного, приводячи до двох вхідних и двох вихідних з'єднань у ланцюгу клієнт-сервер. Розглянемо ситуацію, коли клієнт А відправляє 1 запит в хвилину, а клієнт Б – 50 запитів у секунду по вже встановленому з'єднанню. В результаті сервер, який був обраний для клієнта А, буде оброблювати приблизно в 3000 разів менше, ніж сервер обраний для клієнта Б. Тому безперервне TCP-з'єднання порушує перш за все мету балансування навантаження і є недоліком.

Балансування категорії L7 працює на прикладному рівні моделі OSI, орієнтовані на роботу з високорівневими протоколами такими як HTTP/HTTPS. В процесі балансування відбувається аналіз запитів клієнта та перенаправлення на різні сервери в залежності від характеру контенту (за такими принципами працює модуль Upstream веб-серверу Nginx). Оскільки L7 балансувальник виконує значно складніший аналіз, перетворення та маршрутизацію трафіку веб-додатку, тому навантаження на апаратну складову значно збільшується, що приводить до зменшення продуктивності обробки навантаження (яка вимірюється у оброблених пакетах за секунду), порівняно з оптимізованими L4 балансувальниками. Слід зазначити про значну складність розробки алгоритмів аналізу трафіку, серед яких вірогідність програмної помилки значно більша.

Підводячи підсумки можна зазначити, що важливу роль для сучасних мережевих протоколів відіграє балансувальники категорії L7. Хоча порівнюючи функціонально категорія L7 може повністю замінити L4, але кожний метод балансування виконує свою задачу і займає відповідне місце. Балансування категорії L4 використовується не менше, бо майже всі розподілені архітектури використовують дворівневу систему балансування навантаження L4/L7 для інтернет-трафіку. Тому для найбільш продуктивного балансування навантаження на сервери слід поєднувати можливості різних методів.

Список джерел:

1. Bourke T. Server Load Balancing // O'Reilly & Associates. – 2001. – V.126/ – N 2. – P. 182.
2. Introduction to modern network load balancing and proxying [Електронний ресурс] – Режим доступу: [www / URL: https://blog.envoyproxy.io/introduction-to-modern-network-load-balancing-and-proxying-a57f6ff80236](http://www.blog.envoyproxy.io/introduction-to-modern-network-load-balancing-and-proxying-a57f6ff80236).