

КОМПЛЕКСНА СИСТЕМА БАЛАНСУВАННЯ ДОСТУПУ ДО WEB-РЕСУРСІВ

Захарченко Сергій, канд. техн. наук, доцент кафедри обчислювальної техніки,
Яланський Дмитро, студент групи ІКС-15мн,
Вінницький національний технічний університет, Україна

Ще на ранній стадії розвитку будь-якого веб-проекту необхідно надати особливе значення плануванню розподілу навантаження. Перевантаження серверу може призвести до виходу з ладу (падіння), що завжди відбувається несподівано і в непотрібний момент, загрожує досить серйозними наслідками — як моральними, так і матеріальними. В деякій мірі проблеми недостатньої продуктивності серверу, які виникають в результаті зростання навантаження можна вирішувати шляхом збільшення рівня потужності серверу, або ж оптимізацією використовуваних алгоритмів, програмних кодів і т.п. Але рано чи пізно настає момент, коли і ці заходи виявляються недостатніми.

Для вирішення проблеми зростання навантаження доводиться використовувати різні програмно-апаратні методи. Дієвим методом, що допомагає не тільки у вирішенні проблеми високих навантажень, а й забезпечує резервне збереження даних, є кластеризація (об'єднання кількох серверів в одну цілісну систему, кластер). Навантаження між вузлами одного кластеру розподіляється за допомогою комплексу спеціальних методів, які називають методами балансування. Ефективність кластеризації безпосередньо залежить від того, як розподіляється (балансується) навантаження між елементами кластера [1].

Процедура балансування здійснюється за допомогою алгоритмів і методів, що відповідають таким рівням моделі OSI:

- мережевому.
- транспортному;
- прикладному.

При використанні балансування на мережевому рівні необхідно реалізувати можливість обробки запитів одного і того ж доменного імені різними фізичними серверами. Така можливість реалізується кількома способами:

- DNS-балансування — одному доменному імені виділяється кілька IP-адрес на авторитетному DNS-сервері. Вибір IP адреси серверу, на який буде направлений клієнтський запит, зазвичай визначається за допомогою спеціального алгоритму. Також цей спосіб надає можливість балансування за територіальною ознакою, яку можна реалізувати шляхом розміщення серверів з однаковими ресурсами в територіально різних регіонах глобальної мережі. Це дозволить направляти запити клієнтів до територіально найближчого серверу [2].
- Побудова NLB-кластера — сервери об'єднуються в кластер, що

складається з вхідних і обчислювальних вузлів. Розподіл навантаження здійснюється за допомогою спеціальних алгоритмів, що сприяють не тільки розподілу навантаження, а й централізованій фільтрації доступу до web-ресурсів і надійному збереженню даних.

Балансування на транспортному рівні передбачає те, що клієнтський запит буде спочатку отриманий балансувальником, а потім направлений на один із серверів, який і буде його обробляти. Вибір сервера, на якому буде оброблятися запит, здійснюється шляхом простого кругового перебору (round robin), шляхом вибору найменш завантаженого сервера з пулу тощо.

При балансуванні на прикладному рівні балансувальник аналізує клієнтські запити і перенаправляє їх на різні сервери залежно від характеру запитуваного контенту.

Для досягнення максимальної ефективності необхідно вдаватись до поєднання різних методів балансування. Результатом поєднання може бути багаторівнева комплексна система балансування доступу до web-ресурсів, реалізована шляхом використання окремого DNS-сервера та кількох NLB-кластерів, що розміщені в територіально різних частинах глобальної мережі. Корисність даної системи в тому, що можна отримати стійку до навантажень систему, використовуючи кілька рівнів балансування. DNS-сервер з модулем GeoDNS зможе виконувати направлення запитів до найближчого відносно географічного розміщення клієнта NLB-кластера, а в зібраному кластері реалізуватиметься процес балансування доступу і стабілізації роботи шляхом розподілу запитів між серверами в кластері та фільтрації трафіку. Структурну схему комплексної системи балансування доступу до web-ресурсів зображено на рис. 1.

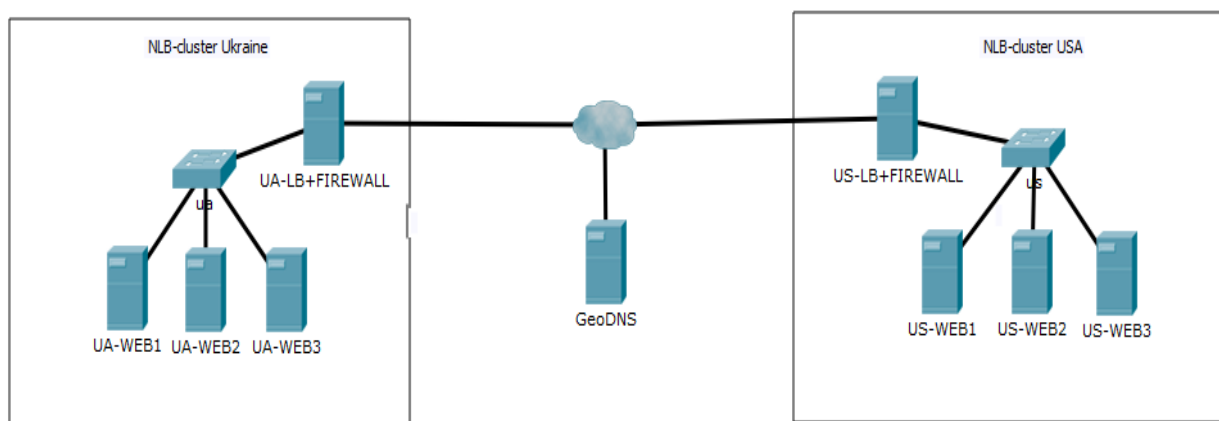


Рисунок 1 – Комплексна система балансування доступу до web-ресурсів

Список використаної літератури

1. Gillam, Lee. Cloud Computing: Principles, Systems and Applications / Nick Antonopoulos, Lee Gillam. — L.: Springer, 2010. — 379 с.
2. Крікет Лі. DNS BIND Керівництво для системних адміністраторів / Крікет Лі, Пол Альбітц. — Символ-Плюс, 2008. — 709 с.