

ИУ7-54Б, 16_KOZ, Рунов Константин

СОДЕРЖАНИЕ

| | |
|--|----------|
| ОПРЕДЕЛЕНИЯ | 3 |
| ВВЕДЕНИЕ | 4 |
| 1 Анализ предметной области | 6 |
| 1.1 Области применения балансировки нагрузки | 6 |
| 2 Описание существующих решений | 6 |
| 2.1 Методы на основе хеширования | 6 |
| 2.1.1 Хеширование на основе IP-адреса | 6 |
| 2.1.2 Хеширование на основе URL-адреса | 6 |
| 2.2 Метод фиксированных весов | 6 |
| 3 Классификация существующих решений | 6 |
| 3.1 Иерархия методов | 6 |
| 3.2 Оценка и сравнение | 6 |
| 3.3 Вывод | 6 |
| ЗАКЛЮЧЕНИЕ | 7 |
| СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ | 8 |

ОПРЕДЕЛЕНИЯ

Ниже представлены термины в контексте данной работы и их определения.

Клиент — это программное или аппаратное обеспечение, которое инициирует запросы к серверу для получения определенных услуг, ресурсов или информации.

Сервер — это программное или аппаратное обеспечение, которое предоставляет услуги, ресурсы или информацию клиентам. Сервер обрабатывает запросы, поступающие от клиентов, и предоставляет им необходимую функциональность или данные.

Нагрузка — это объем данных, поступающий на сервер в определенный период времени.

Запрос — это данные, отправляемые клиентом серверу, с целью получения определенных услуг, ресурсов или информации. Запрос содержит необходимую информацию для предоставления сервером функциональности или данных, запрашиваемых клиентом. Запросы составляют нагрузку.

Группа серверов — это совокупность двух или более серверов, объединенных для совместной работы с целью обеспечения более высокой производительности, отказоустойчивости или распределения нагрузки.

ВВЕДЕНИЕ

Для обеспечения бесперебойной работы своих интернет-ресурсов, компаниям необходимо обеспечить бесперебойную работу своих серверов. Сервера могут выходить из строя по многим причинам: аппаратные сбои, проблемы с электропитанием, человеческие ошибки, перегрузка; в связи с чем, у компаний возникает естественная потребность — минимизировать вероятность выхода из строя своих серверов. От аппаратных сбоев и человеческого фактора защититься бывает крайне сложно, но если грамотно распределять между серверами поступающий объём данных, можно уменьшить вероятность выхода серверов из строя в следствие перегрузки.

Этим и занимаются так называемые «балансировщики нагрузки». Балансировщик нагрузки выполняет роль «регулирующего», стоящего перед серверами и направляющего запросы клиентов на все серверы, способные выполнить эти запросы таким образом, чтобы максимально увеличить скорость и загрузку мощностей и не допустить перегрузки одного сервера, что может привести к снижению производительности. Если один из серверов выходит из строя, балансировщик нагрузки перенаправляет трафик на оставшиеся работающие серверы. При добавлении нового сервера в группу серверов балансировщик нагрузки автоматически начинает направлять на него запросы. [1]

Сервер, на который следует направить клиентский запрос, балансировщики нагрузки выбирает в соответствии с различными алгоритмами и методами балансировки нагрузки.

Целью данной работы является анализ и классификация методов балансировки нагрузки высоконагруженных систем.

Для достижения поставленной цели необходимо решить следующие задачи:

- рассмотреть основные подходы к балансировке нагрузки;
- описать методы балансировки нагрузки, относящиеся к одному из подходов;
- предложить и обосновать критерии качества описанных методов;
- сравнить методы по предложенным критериям оценки;

- выделить методы, показывающие лучшие результаты по одному или нескольким критериям.

1 Анализ предметной области

1.1 Области применения балансировки нагрузки

Статическая балансировка

Динамическая балансировка

2 Описание существующих решений

2.1 Методы на основе хеширования

2.1.1 Хеширование на основе IP-адреса

2.1.2 Хеширование на основе URL-адреса

2.2 Метод фиксированных весов

3 Классификация существующих решений

3.1 Иерархия методов

3.2 Оценка и сравнение

3.3 Вывод

ЗАКЛЮЧЕНИЕ

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. What Is Load Balancing? [Электронный ресурс]. — Режим доступа: <https://www.nginx.com/resources/glossary/load-balancing> (дата обращения: 12.11.2023).