

Телекоммуникационные системы и компьютерные сети



Лозинская В.Н., доцент, к.т.н.
Костюк Д.Е., магистрант

ГОУ ВПО «Донецкий Национальный Технический Университет»

АНАЛИЗ МЕТОДОВ И АЛГОРИТМОВ БАЛАНСИРОВКИ НАГРУЗКИ

В современном мире телекоммуникационные системы работают при пиковых нагрузках на узлы связи. Данная проблема является актуальной и требует современных решений для обеспечения необходимого уровня показателей качества обслуживания. Решением данного вопроса является оптимизация процесса работы узлов сети таким образом, чтобы максимально использовались ресурсы каждого, не допуская при этом выхода его из строя.

Целью исследования является анализ существующих методов балансировки для повышения эффективности использования ресурсов сети с обеспечением необходимого уровня качества обслуживания.

Задачи настоящего исследования заключаются в анализе современных алгоритмов распределения нагрузки и выявление недостатков исследуемых алгоритмов и методов.

При исследовании алгоритмов и методов балансировки нагрузки в телекоммуникационных системах и компьютерных сетях следует отметить следующие параметры качества обслуживания:

- доступная пропускная способность (B, bandwidth);
- уровень потери пакетов (Pl, PacketLoss);
- задержка (D, Delay);

- джиттер задержки (J, jitter)[1].

Выбор метода балансировки зависит от конкретных условий, в которых данная балансировка должна осуществляться. Например, для достижения следующих целей.

- предсказуемость: необходимо знать и понимать, как система балансировки будет вести себя в определённых условиях;
- равномерная загрузка ресурсов системы;
- масштабируемость;
- эффективность;
- сокращение времени выполнения запроса;
- сокращение времени отклика.

Для реализации эффективного распределения нагрузки между узлами в сети с высокой пропускной способностью алгоритм должен обладать следующими показателями:

- быстрая обработка данных и передача их в сеть;
- резервирование полосы пропускания;
- реализация алгоритма распределения на основе требований QoS к определенным типам трафика;
- гибкость использования алгоритмов обслуживания очередей на узлах

сети.

Достижимость вышеперечисленного происходит за счет использования основных алгоритмов балансировки: Round Robin; Weighted Round Robin; Least Connections; Sticky Session [5].

В качестве допущения для исследования алгоритмов балансировки примем за основу потоки трафика на выходе из устройства, распределяемые по очередям с обслуживанием WFQ, которое должно удовлетворять современным требованиям QoS. В таком случае принято ещё одно допущение: критерием для балансировки потоков будут выступать протоколы транспортного уровня модели OSI; в качестве максимально приоритетного трафика будут выступать потоки с использованием транспортных протоколов RTP/UDP (используются для передачи голоса и видео при общении в реальном времени), которые будут отнесены к наиболее требовательному классу с установлением необходимых для классификации идентификаторов поля TOS в IP пакете.

Алгоритм балансировки Round Robin [2] представляет собой перебор запросов по круговому циклу: первый запрос — первому узлу, второй запрос — второму узлу и так до достижения последнего. Данный алгоритм балансирует нагрузку независимо от используемых пакетов протокола. Главным критерием является обращение к узлу по доменному имени. К достоинствам алгоритма можно отнести следующее: низкую стоимость; простоту реализации; отсутствие связи между узлами участниками алгоритма; алгоритм работает независимо от нагрузки на сервер. Несмотря на достоинства, у алгоритма также существуют следующие недостатки: отсутствие равнонабора ресурсов каждого узла; отсутствие алгоритма проверки загрузки сервера; отсутствие алгоритма проверки необходимой пропускной способности и отсутствие

возможности управления потоками на основании используемого транспортного протокола.

Усовершенствованный алгоритм Round Robin (Weighted Round Robin [2]), согласно которому каждому узлу-участнику алгоритма присваивается весовой коэффициент на основе его производительности: больший коэффициент более производительному узлу. Данный алгоритм позволяет использовать ресурсы системы более гибко, но также не решает всех проблем с отказоустойчивостью узлов.

Алгоритм Least Connections обладает несомненным преимуществом по отношению к двум рассмотренным ранее – контроль количества подключений к узлу. Данная особенность алгоритма позволяет распределить нагрузку на узлы так, чтобы избежать отказа одного или более узлов из-за перегрузки. Это осуществляется путем передачи запроса узлу с наименьшим количеством активных подключений.

Также существует усовершенствованный вариант данного алгоритма — Weighted Least Connections, который при передаче запроса учитывает не только количество активных подключений, но и весовой коэффициент узла кластера, что позволяет выстроить кластерный сервис с разным набором ресурсов.

Алгоритм Sticky Sessions характеризуется распределением входящих запросов таким образом, чтобы один и тот же узел обслуживал одну группу клиентов. Такой алгоритм используется, например, в веб-серверах. В указанном алгоритме запросы распределяются между узлами на основе IP-адреса клиента, что реализуется методом IP hash.

На сегодняшний день существует множество алгоритмов балансировки нагрузки, которые способны обеспечить необходимый уровень качества обслуживания и отвечают основным требованиям распределения: эффективность; равномерное использование ресурсов системы; предсказуемость; уменьшение времени отклика; уменьшение времени ответа на запрос.

При использовании алгоритма Round Robin нагрузка на узлы сети будет распределиться равномерно, но данный алгоритм работает без учета состояния узлов, что может привести к «отказу в обслуживании» дальнейших потоков.

Использование Least Connections позволяет распределить нагрузку таким образом, чтобы все узлы сети были заняты на основании учёта состояния занятости, что позволяет осуществить передачу запросу «свободному» узлу. Данный алгоритм обладает существенным недостатком — отсутствием анализа приоритета передачи.

Sticky Sessions распределяет нагрузку так, чтобы при повторном обращении клиент попадал на тот же узел. Это позволяет сэкономить ресурсы всей системы за счет группового обслуживания и планирования инфраструктуры под необходимое количество клиентов. Данный алгоритм также обладает недостатком отсутствия анализа состояния занятости и обслуживания по приоритетам.

Несмотря на очевидные достоинства, в рамках исследовательской задачи, обладают серьезным недостатком — использование большого количества

ресурсов системы алгоритмом балансировки, уменьшению влияния которого посвящено дальнейшее исследование.

Список используемых источников:

1. Klampfer Sasa, Influences of Classical and Hybrid Queuing Mechanisms on VoIP's QoS Properties [Электронный ресурс] / Sasa Klampfer, Amor Chowdhury, Joze Mohorko, Zarko Cucej. - Режим доступа: <http://cdn.intechopen.com/pdfs-wm/13380.pdf>
2. Braden R., RFC 2309. Recommendations on queue management and congestion avoidance in the Internet [Электронный ресурс] / D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, and L. Zhang. – Режим доступа: <http://tools.ietf.org/pdf/rfc2309.pdf>