УДК 004.724 : 004.05

Черниговский Александр Валерьевич,
аспирант кафедры «Вычислительные машины, комплексы, системы и сети», ФГБОУ ВО «Ангарский государственный технический университет», тел. 89041112385, e-mail: chernigovsky.alex@gmail.com
Кривов Максим Викторович,
к.т.н., доцент, зав. кафедрой «Вычислительные машины, комплексы, системы и сети», ФГБОУ ВО «Ангарский государственный технический университет», тел. 89025614935, e-mail: vmk@angtu.ru

АНАЛИЗ МЕТОДОВ РАСПРЕДЕЛЕНИЯ СЕТЕВОГО ТРАФИКА МЕЖДУ ПОЛЬЗОВАТЕЛЯМИ СЕТИ

Chernigovskiy A.V., Krivov M.V.

METHODS OF NETWORK TRAFFIC DISTRIBUTION BETWEEN USERS

Аннотация. *В статье рассмотрены основные показатели качества работы сети, приведена их взаимосвязь с различными типами трафика, а также представлены основные алгоритмы управления работой сети.*

Ключевые слова: *сетевой трафик, компьютерные сети, качество обслуживания.*

Abstract. *In this paper based parameters of network quality of service were considered. Their relation with different traffic types was considered. Basic network control algorithms were presented.*

Keywords: *network traffic, computer networks, Quality of Service.*

Современные тенденции развития общественной и корпоративной сферы предъявляют все более высокие требования для всей IT-индустрии в целом. Отдельное внимание при этом уделяется различным сетевым технологиям, т.к. именно коммутируемые компьютерные сети обеспечивают взаимосвязанность всех элементов информационных систем.

Особой проблемой в данной области является выбор методов интеллектуального управления трафиком. Прежде всего, это связано с тем, что различные виды сетевого трафика могут иметь различные характеристики, а, следовательно, их математическое описание может существенно различаться. Поэтому одной из первичных задач было установление единой математической модели, которая бы с достаточной точностью описывала большинство видов трафика [1-4].

Другим направлением развития сетевых технологий стала разработка программных и аппаратных средств контроля и управления передачей данных.

В данной области одной из основных является задача распределения пропускной способности канала связи между пользователями сети. Главная трудность здесь заключается в ограниченности ширины канала связи, в то время, как объем нагрузки на сеть не является постоянным во времени и может значительно превышать имеющуюся пропускную способность канала.

Как показала практика [например, 5], наращиванием пропускной способности канала невозможно решить проблему перегрузки сети. Поэтому в данной области все большую значимость обретают системы распределения сетевого трафика.

Выбор способа распределения трафика прежде всего зависит от характеристик потока данных и параметров качества работы сети [6].

Основной характеристикой потока данных является тип сетевой активности. Он может быть потоковым и пульсирующим. Их различие состоит в том, что потоковый трафик (рисунок 1 а) характеризуется постоянной битовой скоростью (количеством пакетов, передаваемых за одинаковые отрезки времени), в то время как для пульсирующего трафика данная скорость может меняться по времени (рисунок 1 б).

К характеристикам работы сети прежде всего стоит отнести величину потерь сетевых пакетов данных. Потери могут быть вызваны аппаратными проблемами (например, повреждением физического кабеля), перегрузкой буфера коммутируемого оборудования (т.е. перегрузкой сети), битовыми ошибками или правилами ACL.

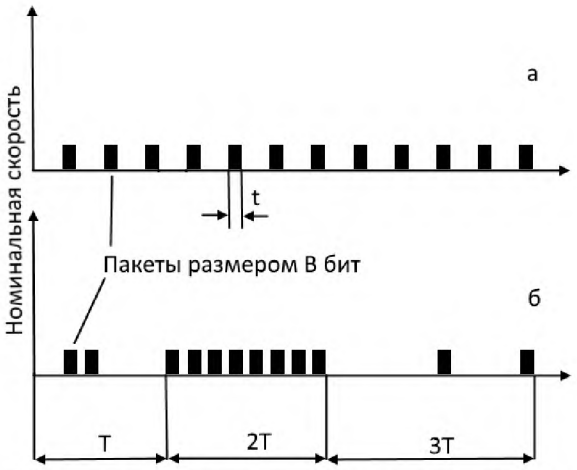

Рисунок 1 – Потоковый (а) и пульсирующий (б) трафик

Для улучшения работы сети величина потерь должна быть сведена к минимуму, однако в зависимости от типа трафика она может различаться [7]. Так, для особо чувствительных приложений критично появление хотя бы одного «битого» или искаженного пакета данных. В свою очередь, для условно устойчивых приложений (например, различные мультимедиа, а также другие данные, передаваемые по протоколу UDP) потеря одного или нескольких пакетов не является критичной вследствие возможности самих приложений аппроксимировать соседние пакеты. Но даже в таком случае порог потерь не должен превышать 1 %, чтобы избежать проблем воспроизведения данных.

Еще одним важным критерием работы сети является задержка передачи пакетов данных (рисунок 2). По сути своей она представляет собой интервал времени, за который данные будут доставлены от отправителя получателю, и состоит из следующих частей [7-8]:

- время на передачу сигнала в физической среде;
- время, затрачиваемое маршрутизатором на разбивку пакетов данных для последующей передачи по каналу связи (задержка сериализации);
- время нахождения пакетов данных в очереди для обработки маршрутизатором и ожидания передачи вследствие ограничения скорости.

По чувствительности к задержкам передачи пакетов характеристики трафика также различаются [7].

Для асинхронных пакетов данных время задержки практически не влияет на их качество (например, при работе электронной почты). Пакеты данных интерактивных приложений более чувствительны к задержкам, чем асинхронные, однако при этом не происходит значительного снижения качества передаваемых данных (например, при задержках передачи потокового видео во время просмотра в нем можно наблюдать незначительные артефакты).

Наибольшую чувствительность к задержкам пакетов данных имеют такие виды трафика, где требуется высокая точность и строгая взаимная синхронизация передаваемых данных. В эту категорию попадает изохронный трафик (например, VoIP), для которого в случае превышения некоторого порога задержки возникает существенное ухудшение качества передаваемых данных (например, искажение голоса и простаивание линии связи). А сверхчувствительным к задержкам является трафик систем управления, работающих в режиме реального времени – брокерских терминалов, медицинских систем, промышленных объектов управления и т.д.

Отдельного внимания заслуживает такая характеристика сети, как джиттер (рисунок 2). Это отклонения от среднего времени передачи последовательности пакетов данных, которые носят колебательный или пульсирующий характер.

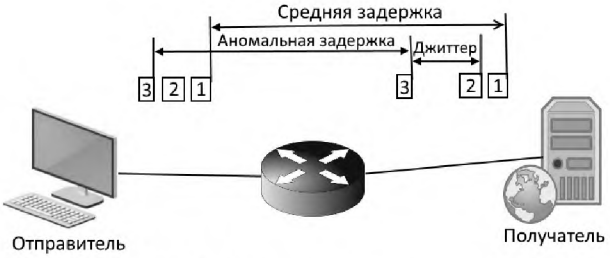

Рисунок 2 – Основные характеристики сети

Все эти особенности трафика учитываются при реализации системы качества обслуживания (Quality of service, QoS), которая по сути определяет возможность обеспечения максимально эффективной передачи данных в заданных условиях системы.

Если подходить к вопросу о распределении трафика между пользователями сети с точки зрения системы качества обслуживания, то можно выделить три основных модели QoS [8].

Первой моделью, которая была применена в IP-сетях, стал Best Effort Service (BE,

сервис негарантированной доставки). Особенность его в том, что весь трафик имеет одинаковый приоритет, т.е. по сути равнозначен между собой, следовательно, не требует дополнительных регулирующих механизмов. К его достоинствам можно отнести простоту и использование всей пропускной способности канала, что позволяет активно применять его в трансконтинентальных каналах связи, в ряде центров обработки данных, а также в пользовательских запросах, отправленных в Интернет.

К недостаткам данного метода можно отнести высокую вероятность возникновения задержек и потерю части передаваемых данных при значительной загрузке канала.

При стремительном росте объемов Интернет-трафика появилась вторая модель – Integrated Service (IS, интегрированная модель обслуживания). Идея этой модели заключалась в резервировании полос пропускания по протоколу RSVP для каждого потока данных по мере прохождения запроса. Механизм работы заключается в том [9], что до начала передачи на узлы прохождения передаются запросы RSVP Path с желаемой полосой пропускания. Если запрос возвращается в виде RSVP Resv, то устанавливается TCP-сессия. Если RSVP возвращает ошибку, механизм IS невозможен, и передача будет осуществляться по BE. Технология IS не оправдала возложенные на нее ожидания из-за повышенной потребности в CPU, памяти и потребности в согласовании выделяемой полосы.

Третий метод – Differentiated Service (DS, модель дифференцированного обслуживания) – является более сложным, но? в то же время? обеспечивает наиболее гибкое распределение трафика между пользователями. В основе алгоритмов реализации данного метода лежит принцип приоритезации трафика. Он означает, что для более эффективной работы сети необходимо четко классифицировать трафик (зачастую по нескольким уровням), а также использовать различные механизмы управления сетевыми ресурсами. В DS отсутствует сигнализация RSVP, акцент делается не на потоки, а на то, что каждый узел сети, по-разному обрабатывает проходящие через него пакеты данных, исходя из их заголовков, тем самым образом обеспечивая данным нужный класс.

Базовые задачи QoS заключаются в обеспечении необходимых условий и параметров передачи пакетов данных должного качества. К этим задачам относятся классификация, маркировка, управление перегрузками, их предотвращение и ограничение скорости [10].

Классификация пакетов – это механизм, определяющий принадлежность пакета к конкретному классу трафика.

В отличие от него маркировка пакетов выполняется для установления соответствующего приоритета в виде метки по типу обслуживания (Type of Service, TOS) (рисунок 3).

Управление перегрузками предполагает организацию работы сети в моменты ее пиковой загрузки и включает следующие алгоритмы [8]:

1. Алгоритм обработки очереди FIFO (First In First Out) предполагает, что запросы не имеют какого-либо приоритетного обслуживания и обрабатываются по мере поступления. Если память буфера оказывается заполнена, все новые заявки будут отбрасываться до тех пор, пока не обработаются находящиеся в буфере заявки или пока в нем не высвободится требуемое количество памяти.

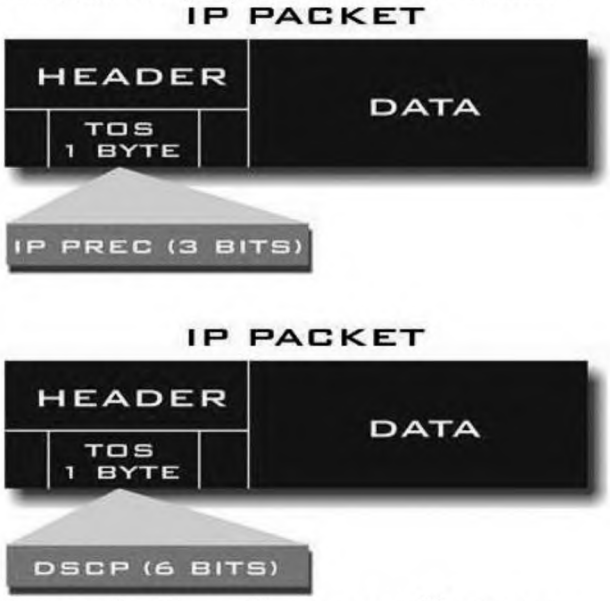

Рисунок 3 – Типы обслуживания IP-пакетов

2. Алгоритм очереди приоритетного обслуживания (Priority Queuing, PQ, классификация пакетов по признакам) основан на распределении трафика в соответствии с четырьмя уровнями приоритета (высокий, средний, нормальный и низкий). Данный алгоритм работает с сетевыми интерфейсами, портами, локальными и удалёнными IP-адресами. Обработка очередей пакетов про-

исходит от высшего приоритета к низшему. При этом пока не закончится обработка запросов с высоким приоритетом, низкоприоритетные запросы обрабатываться не будут, а при больших сетевых нагрузках они вообще могут быть заблокированы.

3. Алгоритм настраиваемых очередей (Custom Queuing, CQ) используется для исключения недостатков системы очередей PQ. Он может формировать от 0 до 16 очередей трафика, где наивысший приоритет (0) имеет системная очередь (используется системой для передачи пакетов управления). Очереди CQ обрабатываются в циклическом режиме, при этом все очереди делятся на группы с заданной пропускной способностью. В таком случае каждому виду трафика (даже низкоприоритетному) гарантированно достается своя полоса пропускания.

4. Алгоритм взвешенного равномерного обслуживания (Weighted Fair Queuing) поддерживает 8 уровней приоритета, но в отличие от предыдущей модели, наивысшим приоритетом характеризуется восьмой уровень. Данный алгоритм также работает в циклическом режиме и используется для приоритезации потоков с меньшей сетевой интенсивностью

5. Алгоритм с низкой задержкой (Low Latency Queuing, LLQ) используется в тех случаях, когда необходимость низкой задержки более важна, чем ширина полосы пропускания. По сути своей он является комбинацией LLQ- и PQ-алгоритмов. Пакетам, для передачи которых важна низкая задержка, присваивается приоритет самого высокого уровня, а пакеты с более низким уровнем приоритета будут ждать освобождения памяти буфера. LLQ дает хорошие показатели в VoIP и видеоконференциях.

Механизмы предотвращения перегрузок сети необходимы в том случае, когда очередь в маршрутизаторе оказывается переполнена.

Одним из самых простых методов предотвращения перегрузки сети является Tail Drop, суть которого сводится к отбрасыванию всех новых поступивших пакетов данных, не помещающихся в буфер (рисунок 4 а).

Не менее простой метод Head Drop позволяет отбрасывать пакеты данных, которые стоят в очереди очень давно (рисунок 4 б).

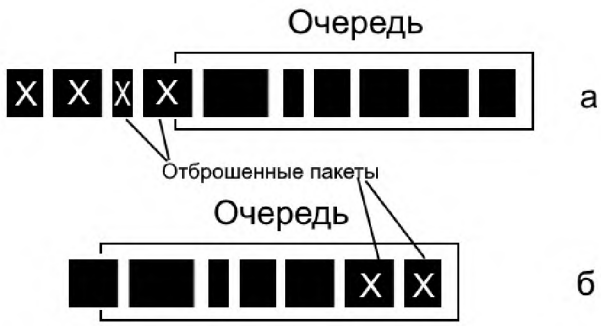

Рисунок 4 – Методы Tail Drop (а) и Head Drop (б)

Оба эти метода могут работать как по отдельности, так и одновременно (рисунок 5).

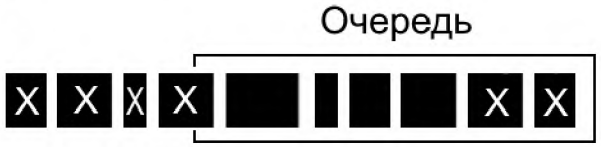

Рисунок 5 – Одновременная работа Tail Drop и Head Drop

К более сложным методам относят интеллектуальные методы предотвращения перезагрузок: произвольного раннего обнаружения (Random Early Detection, RED) и взвешенного произвольного раннего обнаружения (Weighted Random Early Detection, WRED) [11].

При работе метода RED при заданном уровне (например, 80 %) заполнения буфера случайные пакеты данных начинают отбрасываться.

Метод WRED предполагает, что у очередей пакетов с разным приоритетом будет различный отброс пакетов, причем отбрасывание происходит по методу Tail Drop. Например, в первой очереди (рисунок 6, кривая 1) отбрасывание пакетов данных начинается при заполнении буфера маршрутизатора на 20 %. При увеличении объема буфера с 20 до 40 % вероятность отбрасывания составляет около 20 %.

Во второй очереди (рисунок 6, кривая 2) отбрасывание пакетов данных происходит при 30 %-ной загрузке буфера, а к 50 %-ной вероятность отбрасывания пакетов достигает 10 %. Для третьей очереди (рисунок 6, кривая 3) нижний и верхний уровни загрузки буфера составляют 50 и 100 % соответственно, а вероятность отбрасывания пакетов снижается до 5 %.

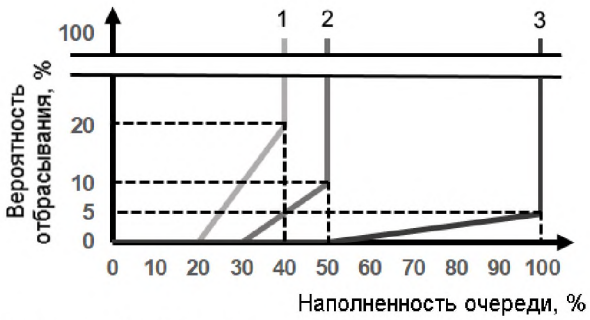

Рисунок 6 – Предотвращение перегрузок сети методом WRED

Ограничение скорости может осуществляться двумя основными методами – полисингом или шейпингом [12]. Полисинг предполагает ограничение ширины канала путем отбрасывания лишнего трафика (рисунок 7).

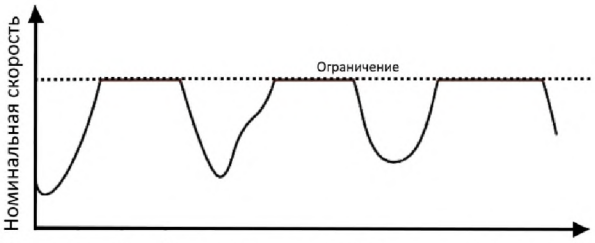

Рисунок 7 – Ограничение скорости передачи данных методом полисинга

В основе шейпинга лежит принцип буферизации трафика (рисунок 8).

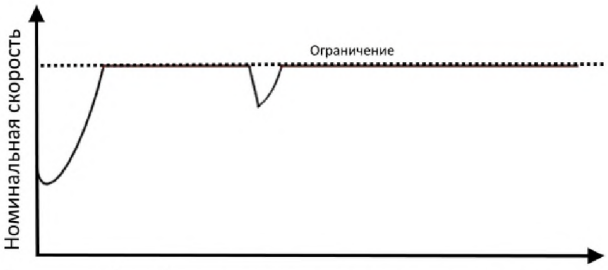

Рисунок 8 – Ограничение скорости передачи данных методом шейпинга

В заключение можно отметить, что проблема низкой эффективности работы сети не имеет однозначного решения. Причины ухудшения ее работы могут быть как на физическом, так и на методологическом уровне. И прежде чем использовать различные методы управления трафиком, необходимо изучить структуру сетевых потоков и самой сети, принципы ее работы и применяемые алгоритмы.

Иногда решением проблемы улучшения качества обслуживания сети может стать расширение полосы пропускания, однако в большинстве случаев оно не является достаточной мерой. Именно поэтому основной акцент делается на выявление слабых мест системы, а также на дальнейший подбор оптимального способа алгоритма функционирования самой сети или отдельных ее элементов, в том числе, на основе нейросетевых технологий [например, 13, 14].

СПИСОК ЛИТЕРАТУРЫ

1. Leland W. E. On the self-similar nature of Ethernet traffic / W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson // Computer commu-nication review. – 1995. – Issue 1. – P. 203.

2. Черниговский, А.В. Основные модели сетевого трафика / А.В. Черниговский, М.В. Кривов // Вестник АнГТУ. – 2017. – № 11. – С. 137-143.

3. Осин, А.В. Регрессионные и фрактальные модели телекоммуникационного трафика / А.В. Осин // Электротехнические и информационные комплексы и системы. – 2006. – Т. 3. – № 1. – С. 28-37.

4. Сидорова, О.И. Математические модели трафика в современных телекоммуникационных системах: автореф. дис. … канд. физ.-мат. наук: 05.13.18 / Сидорова О.И. – Тверь, 2009. – 19 с.

5. Cisco QoS для начинающих / [Электронный ресурс] // Cisco [сайт]. URL: http://net-work.xsp.ru/8_3_11.php (дата обращения сентябрь 2018 г.).

6. Нетес, В.А. Мониторинг параметров работы сетей и временная синхронизация / В.А. Нетес // T-Comm – Телекоммуникации и Транспорт. – 2014. – № 2. – С. 36-37.

7. Олифер В., Олифер Н. Компьютерные сети. Принципы, технологии, протоколы: Учебник для вузов. 5-е изд. – СПб.: Питер, 2017. – 992 с.

8. Со-Мин-Тун. Исследование качества обслуживания в локальных вычислительных сетях / Со-Мин-Тун // Горный информационно-аналитический бюллетень. – 2008. – № 2-2. – С. 202-213.

9. Легков, К.Е. Процедуры и временные характеристики оперативного управле-

ния трафиком в транспортной сети специального назначения пакетной коммутации / К.Е. Легков // T-Comm – Телекоммуникации и Транспорт. – 2012. – № 6. – С. 22-26.

10.Шринивас В. Качество обслуживания в сетях IP.: Пер. с англ. – М.: Издательский дом «Вильямс», 2003. – 368 с.

11.Лемешко, А.В. Модель и метод предотвращения перегрузки с активным управлением очередью на узлах телекоммуникационной сети / А.В. Лемешко, М.В. Семеняка // Проблемы телекоммуникаций. – 2014. – № 2 (14). – С. 91-104.

12.Comparing traffic policing and traffic shaping for bandwidth limiting / [Электронный ресурс] // Cisco [сайт]. URL: https://www.cisco.com/c/en/us/support/docs/quality-of-service-qos/qos-policing/19645-policevsshape.html (дата обращения сентябрь 2018 г.).

13.Auld, T. Bayesian neural networks for Internet traffic classification / T. Auld, A. W. Moore, S. F. Gull // IEEE Transactions on neural networks. – 2007. – Vol. 18. – №. 1. – pp. 223-239.

14.Michael A. K. J., Valla E., Neggatu N. S., Moore A. W. Network traffic classification via neural networks : Technical report. – University of Cambridge. – 2017, September. – 25 p.