

Влияние современных способов оптимизации загрузки интернет-страниц в браузерах на параметры качества обслуживания http-трафика

В соответствии с концепцией предоставления сервисов WEB2.0 интернет-браузеры (Firefox, IE, Google Chrome и др.) широко применяют методы оптимизации загрузки web-страниц такие как: параллельные HTTP-соединения и HTTP-конвейер. Изменение алгоритмов загрузки страниц приводит к изменению параметров HTTP-трафика, поэтому важно исследовать влияние современных методов оптимизации загрузки интернет-страниц в браузерах на качества обслуживания HTTP трафика. В данной работе было исследовано влияние количества параллельных HTTP-соединений на параметры качества Web-трафика, передающегося в сети. Исследование проводилось с по-мощью имитационного моделирования в программном комплексе NS-2 с подключенным модулем PackMIME. Проведено моделирование процесса обработки IP-пакетов в сети провайдера между опорным и пограничным маршрутизаторами. Результаты исследования являются зависимости основных параметров качества обслуживания, таких как средняя задержка и потери пакетов от количества параллельных HTTP-соединений. На основании полученных результатов можно судить об эффективности современных механизмов оптимизации, применяемых в интернет-браузерах, а также оценить время загрузки веб-страницы.

Ключевые слова: WEB2.0, HTTP/1.1, Web-браузер, моделирование, NS-2, PackMIME

Деарт В.Ю.,
к.т.н., доцент кафедры АЭС МТУСИ,
vdeart@mail.ru,

Кожухов И.С.,
аспирант кафедры АЭС МТУСИ,
ivankozhuhov@gmail.com.

Введение

В современных телекоммуникационных сетях передачи данных широко применяется концепция предоставления сервисов WEB2.0. Она помогает пользователям получать более быстрый доступ к информации, а также делает веб-ресурсы более интерактивными, позволяя самим пользователям наполнять их информацией. В соответствии с вышеупомянутой концепцией появилась возможность создавать целые порталы, состоящие из динамических веб-страниц или страниц агрегирующих информацию с различных сайтов. Таким образом, в соответствии с концепцией WEB2.0 меняются не только взаимоотношения между пользователями и веб-ресурсами, но и структура HTTP-трафика в целом [1, 2].

В тоже время, новые инструменты оптимизации в веб-браузерах также ведут к изменениям в структуре HTTP-трафика: растёт объем передаваемого трафика и ускоряется загрузка данных за счет введения параллельных сессий. Все перечисленные факторы радикально меняют параметры HTTP-трафика, которые мы хотели бы оценить в нашем исследовании.

1. Анализ принципов работы современных браузеров

В последнее время мы наблюдаем быструю модернизацию различных браузеров. Можно легко найти последние версии Internet Explorer, Mozilla Firefox, Opera или Google Chrome. Современный Интернет-обозреватель активно обрывает всевозможными механизмами оптимизации скорости загрузки интернет-страниц:

- Cookie;
- HTTP кэширование (HTTP Caching);

- HTTP-сжатие (HTTP-compression);
- Предвыборка DNS (DNS pre-resolution);
- Режим постоянного соединения (Keep-Alive);
- Технология HTTP конвейера (HTTP pipelining);
- Параллельные HTTP соединения (Smart parallel downloads);
- и др.

Методы HTTP конвейера и параллельных HTTP соединений являются наиболее популярными в последних версиях браузеров. В таблице 1 мы видим, как количество параллельных TCP-сессий изменяется с каждой новой версией браузеров [3].

Таблица 1

Таблица соответствия браузера, его версии и используемому количеству одновременных соединений

Браузер	Версия	Максимальное кол-во параллельных соединений к одному хосту	Максимальное кол-во поддерживаемых соединений
Internet Explorer	6	2	34
	7	2	55
	8	6	35
	9	6	35
Mozilla FireFox	4	6	30
	5	6	30
Safari	5	6	30
	6	5	5
Opera	9	4	20
	10	8	30
	11	8	32
	Mini 5	11	30
Google Chrome	9	6	35
	10	6	35
	11	6	35

Как видно из таблицы 1, количество сессий увеличивается от версии к версии для большинства браузеров. Также следует принять во внимание влияние HTTP конвейера. Таким образом, оба фактора могут привести к увеличению количества всплесков (bursts) в структуре HTTP-поток.

2. Имитационное моделирование

Для полного исследования HTTP-трафика необходимо воссоздать изменяющиеся условия в сети и наблюдать изменения – построить динамическую картину поведения HTTP-трафика в сети. Инструментом, который позволит это осуществить, является имитационное моделирование. Имитационное моделирование воссоздаёт взаимодействия различных протоколов в условиях реальной сети за счёт аппаратно-программного комплекса на базе ЭВМ.

Модель HTTP-трафика для исследования строилась с помощью комплекса имитационного моделирования NS2 (Network Simulator ver.2 – эмулятор сети), который пользуется большой популярностью среди исследователей за свою гибкость и доступность [4]. На основе данного симулятора с подключенным модулем PackMIME была создана базовая имитационная модель HTTP-трафика, учитывающая зависимую генерацию вторичных TCP-сессий, реализованную в современных обозревателях. При этом в исходном виде модуль PackMIME не позволяет организовывать вторичные параллельные сессии, поэтому основные исходные файлы модуля были изменены с целью обеспечения возможности задания зависимого потока имитирующего работу веб-браузера с механизмом оптимизации с помощью параллельных HTTP соединений.

За исходные данные базовой имитационной модели были приняты параметры сети и характеристики HTTP-трафика, полученные в работе [5]. Топология моделируемой сети представлена на рис. 1. В левой части располагается сеть доступа, справа сеть Интернет, представляющая собой веб-сервисы и веб-серверы. На схеме опорный маршрутизатор агрегирующий поток данных от пользователей обозначен как R1. Пограничный маршрутизатор обеспечивающий выход в сеть Интернет обозначен как R2.

Анализ пропускной способности при передаче веб-трафика показал, что максимальное влияние на качество обслуживания оказывает пропускная способность канала 2 между опорным (R1) и пограничным (R2) роутерами. Участок сети доступа (канал 1) между локальной сетью и пограничным маршрутизатором организован с применением активной оптической технологии Ethernet со скоростью 1 Гбит/с. Участок сети (канал 3) между опорным маршрутизатором и сетью серверов также имеет пропускную способность 1 Гбит/с. Тогда как в базовой имитационной модели пропускная способность канала 2 является 100 Мбит/с. Таким образом, определяющим характер HTTP-трафика является участок моделируемой сети между двумя маршрутизаторами, поскольку он обладает меньшей пропускной способностью, но при этом используется всеми пользователями сети для доступа в Интернет.

HTTP-трафик имеет асимметричный характер. В восходящем направлении пользователи веб-сервисов в качестве полезной нагрузки передают только HTTP-запросы GET

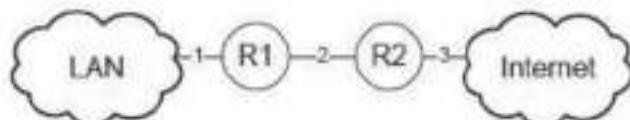


Рис. 1. Топология моделируемой сети

и подтверждения принятых данных, средним размером 650-700 байт [6], задержка или потеря которых оказывает влияние на комплексное качество сервиса предоставляемого пользователю (QoE – Quality of Experience) [5]. В нисходящем направлении основную часть передаваемых данных составляют полезные данные (HTTP-ответы), размером 8-10 Кбайт [5], задержка или потеря которых напрямую влияет на отображение веб-страницы пользователем, а также на комплексное качество предоставления веб-сервиса. Это означает, что трафик от веб-серверов к пользователям примерно в 10 раз больше чем в обратном направлении. На данный момент методики оценки комплексного качества предоставления веб-сервиса находятся на стадии разработки, поэтому для данной работы в качестве оцениваемых параметров качества обслуживания были выбраны задержка и потери пакетов, как объективные показатели, позволяющие производить качественное сравнение различных условий передачи данных веб-приложений.

Также мы сосредоточили внимание на вопросе начечности (burst arrival) современного интернет трафика. Как это было показано в работе [6], интенсивность организации TCP-сессий при передаче HTTP-трафика описывается двумя независимыми распределениями: экспоненциальным и равномерным. Наличие экспоненциального распределения объясняется обобщением трафика от большого числа пользователей, тогда как, наличие равномерного распределения обусловлено работой веб-браузера.

Моделирование было произведено для нисходящего HTTP потока (downstream) от серверов к пользователям. Размеры буферов в обоих роутерах равняются 100 пакетам, а также выбран механизм FIFO (первый пришел – первый ушел) сброса пакетов при перегрузке.

3. Результаты моделирования

Для исследования влияния параллельных TCP-сессий на параметры качества обслуживания HTTP-трафика мы использовали базовую имитационную модель, описанную в [6].

Исследование зависимости средней задержки пакетов от различных значений интенсивности вторичных сессий (1/1 и 1/2) было проведено ранее [6]. В данном случае интенсивность 1/2 означает, что на каждую вторую первичную открывается одна вторичная TCP-сессия.

В нашей работе мы использовали модифицированный PackMIME модуль для введения различного количества параллельных сессий: 2, 4, 6 и 8. В данном случае две параллельные сессии означают, что на каждую первичную сессию приходится две вторичные.

Мы полагаем, что первичные запросы создаются в момент, когда пользователь кликает мышкой в браузере и имеют экспоненциальное распределение между каждым кликом. В свою очередь каждый клик устанавливает первую TCP-сессию с сервером. После того как браузер получит первый ответ сервера, он с очень маленькими интервалами открывает дополнительные параллельные сессии. Предполагается, что подобная процедура увеличивает всплески в TCP-трафике.

На рисунке 2 представлена зависимость средней задержки пакетов в очереди маршрутизатора в зависимости от использования канала Интернет (канал 2) для различного числа параллельных TCP-соединений.

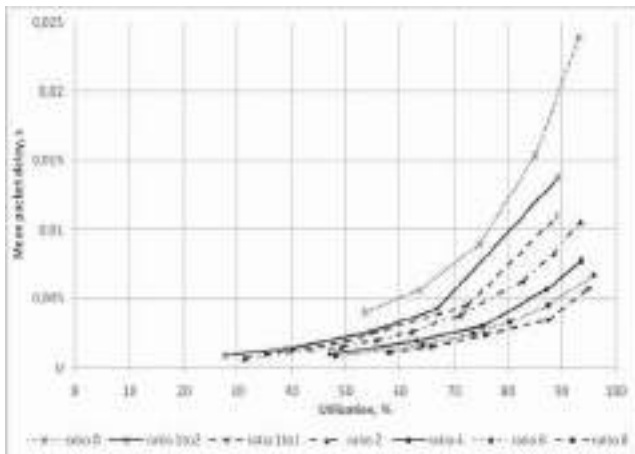


Рис. 2. Средняя задержка пакета в зависимости от числа параллельных TCP-соединений

Как видно из рис. 2 увеличение числа параллельных TCP-сессий приводит к уменьшению средней задержки пакетов.

При последующем исследовании модели выяснилось, что увеличение числа параллельных сессий никак не влияет на такой параметр качества обслуживания как процент потерянных пакетов. То есть характер зависимости процента потерянных пакетов от использования канала практически не изменяется с ростом количества вторичных TCP-сессий.

4. Направление дальнейших исследований

Сегодня задача измерения и контроля параметров качества обслуживания уже достаточно полно и всесторонне исследована, а также стандартизирована. Но в современном мире телекоммуникаций обеспечение заданных параметров качества QoS уже не является достаточным требованием для успешного ведения бизнеса. Поэтому на данный момент остро стоит проблема оценки качества восприятия услуги непосредственно пользователем QoE (Quality of Experience) и соотношения данной оценки QoE с хорошо исследованными параметрами QoS.

В исследовании [7] приводится формула взаимосвязи между оценкой QoE и QoS параметрами, названная IQX гипотеза (IQX hypothesis). График данной зависимости представлен на рисунке 3. Он разбит на три зоны. В зоне 1 при ухудшении параметров QoS пользователь, по началу, не ощущает ухудшения качества предоставляемой услуги. При дальнейшей деградации качества обслуживания QoS (зона 2) ухудшение качества становится всё более заметно и оценка QoE пользователем предоставляемой услуги падает по экспоненте. В конце концов, наступает момент (зона 3), когда происходит отказ пользователем от услуги.

В дальнейшем мы планируем исследовать справедливость вышеупомянутой IQX гипотезы для услуги Web-browsing, приняв за метрику QoE среднее время загрузки HTTP-страницы.

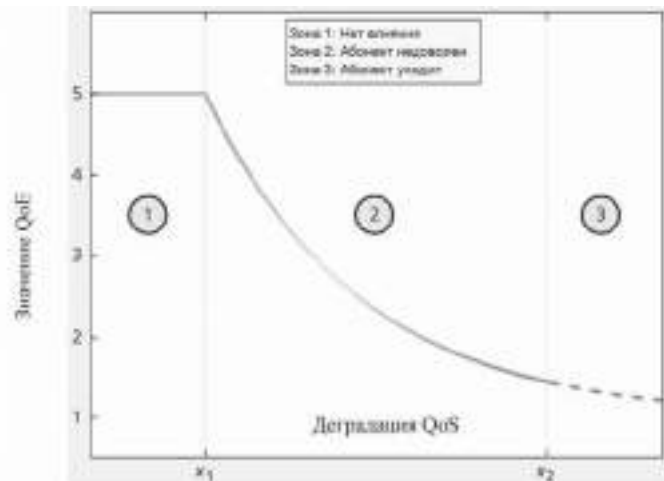


Рис. 3. Зависимость метрик QoE от параметров качества обслуживания QoS

Заключение

Все современные браузеры используют механизмы оптимизации загрузки с помощью параллельных TCP-соединений к одному серверу. На сегодняшний момент число параллельных сессий варьируется от 4 (IE) до 11 (Opera) в зависимости от типа веб-браузера. Параллельные соединения могут иметь огромное влияние на параметры качества обслуживания HTTP-трафика, приводя к увеличению количества всплесков (bursts). Тем не менее, исходя из результатов моделирования, мы видим, что для значений использования канала 2 более 80% ведёт к некоторому снижению средней задержки пакетов.

Литература

1. Padhye J., Firoiu V., Towsley D., Kurose J. Modeling TCP Throughput: A Simple Model and its Empirical Validation // In proceedings ACM SIGCOMM Computer Communication Review, Volume 28, Issue 4, 1998. – P.303-314.
2. Shuai L., Xie G., Yang J. Characterization of HTTP Behavior on Access Networks in WEB2.0 // In proceedings ICT 2008, June, 2008. – P.1-6.
3. <http://www.browserscope.org>, 2011.
4. Wiegler M. PackMime-HTTP: Web Traffic Generation in NS2 // <http://www.isi.edu/nsnam/ns/doc/node552.html>, 2007.
5. Пилиugin А.В., Деарт В.Ю. Модель HTTP-трафика для сети Интернет эпохи WEB2.0 // Труды конгресса по интеллектуальным системам и информационным технологиям "АИС-IT'10". Научное издание в 4-х томах. – М.: Физматлит, 2010. – Т.1. – С.556-562.
6. Пилиugin А.В. Разработка метода оценки параметров качества обслуживания HTTP-трафика в мультисервисных сетях доступа // Диссертация на соискание ученой степени кандидата технических наук. – М.: МТУСИ, 2011.
7. Markus Fiedler, Tobias Hossfeld, Phuoc Tran-Gia. A generic quantitative relationship between Quality of Experience and Quality of Service // IEEE Network, 24(2):36-41, March 2010.