Network Working Group
Request for Comments: 2914
BCP: 41
Category: Best Current Practice

S. Floyd
ACIRI

September 2000

Принципы контроля перегрузок

Congestion Control Principles

Статус документа

Этот документ относится к категории Internet Best Current Practices (обмен опытом) для сообщества Internet и служит приглашением к дискуссии в целях дальнейшего развития. Документ может распространяться свободно.

Аннотация

Целью этого документа является разъяснение необходимости механизмов контроля перегрузок в Internet и обсуждение пригодных способов такого контроля. Одной из конкретных целей является разъяснение опасности пренебрежения контролем насыщения. Другой целью является обсуждение роли IETF в стандартизации новых протоколов контроля перегрузок.

1. Введение

Этот документ основан на более ранних RFC и в некоторых случаях просто включает значительные фрагменты текста из [RFC2309, RFC2357]. Много заимствований сделано также из ранних обсуждений необходимости сквозного контроля перегрузок [FF99].

2. Существующие стандарты контроля перегрузок

Стандарты IETF в части сквозного контроля перегрузок сфокусированы на конкретных протоколах (например, TCP [RFC2581], протоколы с гарантированной доставкой и поддержкой групповой адресации [RFC2357]), синтаксисе и семантике обмена между оконечными узлами и маршрутизаторами информацией о насыщении в сети (например, ECN[1] [RFC2481]) или желаемом качестве обслуживания (diff-serv). Роль сквозного контроля перегрузок рассматривается также в информационном RFC «Recommendations on Queue Management and Congestion Avoidance in the Internet» [RFC2309]. Этот документ рекомендует развёртывание механизмов активного управления очередями в маршрутизаторах и продолжение разработки для маршрутизаторов механизмов работы с потоками, безотносительно к уведомлениям о перегрузках. Мы заимствовали из RFC 2309 некоторые аспекты общего обсуждения сквозного контроля перегрузок.

В отличие от перечисленных выше RFC, этот документ является более общим в части принципов контроля перегрузок. Одним из важных факторов обеспечения работы Internet являются механизмы предотвращения перегрузок TCP. Хотя протокол TCP продолжает доминировать на транспортном уровне Internet, он не применяется повсеместно и все больше приложений по той или иной причине отказывается от TCP. Это относится не только к групповому (multicast), но и к индивидуальному трафику типа потоков multimedia, которым не требуются гарантии доставки, а также трафику DNS и протоколов маршрутизации, который представляет собой короткие передачи, играющие важнейшую роль в работе сети. Значительная часть такого трафика не использует никаких форм резервирования пропускной способности или сквозного контроля насыщения. Продолжающееся использование сквозного контроля насыщения обычным (best-effort) трафиком крайне важно для поддержки стабильности Internet.

В этом документе также рассмотрена общая роль IETF в стандартизации новых протоколов контроля перегрузок.

Обсуждение принципов контроля насыщения для дифференцированных и интегрированных услуг не включено в этот документ. Некоторые категории таких услуг включают сквозные гарантии пропускной способности от сети и по этой причине им не нужны механизмы сквозного контроля насыщения.

3. Разработка сквозного контроля перегрузок

3.1. Предотвращение коллапса насыщения

Архитектура протоколов Internet основана на сквозном обслуживании пакетов без организации явных соединений с использованием протокола IP. Преимущества работы без организации явных соединений в форме гибкости и отказоустойчивости продемонстрированы наглядно. Однако эти преимущества не достаются даром - требуется серьёзная работа по проектированию сети для обеспечения высокого качества обслуживания при значительных нагрузках. Фактически, недостаточное внимание к динамической пересылке пакетов может приводить к деградации сервиса или «обвалу» Internet. Это явление впервые наблюдалось на ранней стадии расширения сети Internet в середине 1980-х [RFC896] и получило название «коллапс насыщения» (congestion collapse).

Исходная спецификация TCP [RFC793] включает механизм управления потоком на основе окна, позволяющий получателю влиять на объем данных, передаваемых отправителем. Управление потоком использовалось для предотвращения переполнения буферов данных на приёмной стороне соединения. В [RFC793] отмечено, что сегменты могут теряться в результате ошибок или перегрузки в сети, но не было предложено динамического изменения размера окна управления потоком данных в ответ на возникновение перегрузки.

[1]Explicit Congestion Notification - явное уведомление о насыщении (перегрузке).

Исходное решение проблемы «обвала» Internet предложил Van Jacobson. Начиная с 1986, Jacobson разработал механизмы предотвращения перегрузок, которые стали обязательными для реализаций TCP [Jacobson88, RFC 2581]. Эти механизмы работают на хостах, заставляя соединения TCP «отступать» (back off) во время перегрузки. Мы говорим, что поток TCP «отвечает» за передачу сигналов о перегрузке из сети (например, путём отбрасывания пакетов). Именно эти алгоритмы предотвращения перегрузок TCP не позволяют коллапсировать современной сети Internet.

Однако на этом история не заканчивается. С 1988 года были выполнены важные исследования динамики роста Internet и этот рост продолжается до настоящего времени. Стало ясно, что механизмы предотвращения перегрузки TCP [RFC2581] необходимы и весьма мощны, но их не достаточно для обеспечения качественного сервиса в любых обстоятельствах. В дополнение к разработке новых механизмов контроля насыщения для конечных точек [RFC2357], создаются основанные на маршрутизации механизмы контроля перегрузок.

Важнейшей и до сих пор полностью не решённой проблемой является возможность коллапса Internet в будущем по причине того, что для потоков не используется сквозного контроля перегрузок. В RFC 896 [RFC896] ещё в 1984 году предложено шлюзам детектировать и «подавлять» некорректно работающие хосты: «Отказ откликаться на сообщения ICMP Source Quench следует считать основанием для шлюза «отключить» соответствующий хост. Детектирование таких отказов является непростой задачей и представляет область для дальнейшего исследования». В современных публикациях все ещё встречаются представления о том, что маршрутизаторы детектируют и «наказывают» потоки, для которых не реализован подходящий механизм сквозного контроля насыщения [FF99].

3.2. Беспристрастность

Кроме озабоченности коллапсом насыщения возникает беспокойство по части беспристрастности применительно к трафику best-effort. Поскольку TCP «схлопывается» во время перегрузок, множество соединений TCP могут проходить через общий сильно загруженный канал так, что пропускная способность будет делиться примерно поровну между «близко расположенными» потоками. Равномерность распределения полосы между потоками зависит от использования всеми потоками общего механизма контроля перегрузок. Для протокола TCP это означает соответствие алгоритмов контроля насыщения действующим спецификациям TCP [RFC793, RFC1122, RFC2581].

Важность вопроса беспристрастности по отношению к одновременным потокам растёт по нескольким причинам. Во-первых, за счёт масштабирования окна [RFC1323] отдельные потоки TCP могут получить высокую пропускную способность даже на путях со значительными задержками. Во-вторых, с ростом числа web-приложений у пользователей Internet возросли потребности в пропускной способности с малыми задержками, по сравнению с потребностями в сравнительно медленной передаче больших файлов в фоновом режиме. Рост трафика best-effort, для которого не применяется протокол TCP, дополнительно усиливает озабоченность беспристрастным распределением ресурсов между одновременными потоками обычного (best-effort) трафика в периоды перегрузок.

Популярность Internet вызвала рост числа реализаций TCP. Некоторые из них могут сталкиваться с отказами при реализации механизмов контроля перегрузок TCP в результате неполного соответствия [RFC2525]. В других может сознательно применяться более агрессивный контроль насыщения в части использования пропускной способности по сравнению с другими реализациями TCP - это позволяет разработчикам заявлять, что их TCP «быстрее других». Ожидаемым последствием появления таких реализаций является спираль все более агрессивных реализаций TCP или возрастающей агрессивности транспортных протоколов, что в конце концов приведёт к возврату во времена, когда протоколов контроля перегрузок ещё не было и хроническое насыщение было естественным состоянием Internet.

Есть общеизвестный способ повышения агрессивности даже без изменения транспортного протокола - просто сменить уровень дискретности - открывается множество «параллельных» соединений с одним узлом, как это делали в прошлом некоторые Web-браузеры. В результате вместо спирали роста агрессивности транспортных протоколов возникает спиральный рост агрессивности браузеров или других приложений.

Это поднимает вопрос о подходящей детализации потока, где термином «поток» обозначается уровень детализации, приемлемый для приложений контроля перегрузок и беспристрастного распределения ресурсов. В RFC 2309 сказано: «Есть несколько «естественных» ответов - 1) соединение TCP или UDP (адрес и порт отправителя - адрес и порт получателя), 2) пара «отправитель-получатель» или 3) данный порт отправителя и данный порт получателя. Мы предполагаем, что пара «отправитель-получатель» обеспечивает во многих случаях подходящий уровень детализации. Детализация потоков для контроля перегрузок, по крайней мере частично, является вопросом политики, который должен решаться в более широком сообществе IETF.»

Снова заимствуя из RFC 2309, мы используем термин TCP-совместимый для потока, который в условиях насыщения ведёт себя, подобно потоку TCP. Такой поток отвечает за уведомление о перегрузке и в установившемся состоянии не использует более широкой полосы по сравнению с потоком TCP в сравнимых условиях (потери, RTT, MTU и т. п.)

Потоки удобно делить на три класса: (1) TCP-совместимые, (2) «безответственные», которые не снижают скорости при возникновении перегрузки и (3) потоки, которые реагируют на перегрузку иначе, нежели TCP. Два последних класса включают агрессивные потоки, которые создают значительные угрозы работе Internet, как описано ниже.

В дополнение к беспристрастности для установившихся состояний представляет интерес и беспристрастность на начальном этапе замедленного старта. Одной из проблем является побочное влияние на другие потоки со стороны потока с чрезмерно агрессивной процедурой slow-start. Производительность замедленного старта очень важна для короткоживущих потоков, которые передают лишь незначительный объем данных.

3.3. Оптимизация пропускной способности, задержки и потерь

В дополнение к предотвращению коллапса насыщения и решению проблемы беспристрастности использование сквозного контроля насыщения в потоке позволяет оптимизировать производительность этого потока в части пропускной способности, задержки и потерь. В некоторых случаях (например, в средах со статистическим мультиплексированием) задержка и потери в потоке в значительной степени независимы от скорости этого потока. Таким образом, поток может использовать свой контроль перегрузок для снижения задержки и числа теряемых пакетов. Отметим однако, что в средах типа современной сети Internet с обслуживанием класса best-effort, вопросы коллапса насыщения и беспристрастности для одновременных потоков ограничивают доступные потоку возможности контроля перегрузки.

4. Роль процесса стандартизации

При стандартизации транспортного протокола принимаются во внимание не только аспекты совместимости (например, информационный обмен между оконечными узлами), но и механизмы, влияющие на производительность работы (например, снижение размера окна насыщения TCP в ответ на отбрасывание пакета). В то же время, конкретные детали реализации и другие аспекты транспортного протокола, не влияющие на взаимодействие и не нарушающие существенно производительность, не требуют стандартизации. Аспекты TCP, не требующие стандартизации, включают детали процедуры Fast Recovery после Fast Retransmit [RFC2582]. В приложении используются примеры TCP для более подробного рассмотрения роли процесса стандартизации в разработке контроля насыщения.

4.1. Разработка новых транспортных протоколов

В дополнение к вопросам предотвращения коллапса насыщения при стандартизации новых транспортных протоколов должна приниматься во внимание возможность «перетягивания одеяла» между конкурирующими протоколами. Например, в RFC 2357 [RFC2357] руководитель направления TSV и его аппарат описывают критерии публикации проектов Internet-Draft в качестве RFC для групповых протоколов с гарантированной доставкой. В документе [RFC2357] сказано: «Особое беспокойство IETF вызывает вопрос влияния гарантированной доставки группового трафика на остальной трафик Internet во время перегрузок и, в частности, конкуренция гарантированной доставки multicast-трафика с трафиком TCP... Задачей IETF является стимулирование исследований и внедрения гарантированной доставки группового трафика для обеспечения максимально быстрого удовлетворения потребности приложений в гарантированной доставке группового трафика, обеспечивая при этом защиту Internet от катастроф и коллапсов насыщения, которые могут быть связаны с широким распространением неподходящих механизмов гарантированной доставки группового трафика.»

Список технических вопросов, которые должны быть решены в RFC для новых транспортных протоколов с гарантированной доставкой достаточно большой. Имеются ли механизмы контроля перегрузок? Насколько хорошо они работают? Когда эти механизмы могут отказывать? Отметим, что механизмам контроля перегрузок с более агрессивным, нежели в TCP, поведением нужно будет приложить значительные усилия для доказательства их безопасности в плане стабильности сети.

Разумно считать, что приведённые выше опасения, связанные с новыми транспортными протоколами, относятся не только к протоколам гарантированной доставки группового трафика, но и к индивидуальному трафику с гарантиями доставки или без них, а также к групповому трафику без гарантий доставки.

4.2. Прикладные вопросы, влияющие на контроль перегрузок

Конкретная проблема браузеров, открывающих множество соединений с одним адресатом, была рассмотрена в RFC 2616 [RFC2616], где в параграфе 8.1.4 сказано: «Клиентам, использующим постоянные соединения, **следует** ограничивать количество одновременных соединений, которые могут поддерживаться для данного сервера. Однопользовательским клиентам **не следует** поддерживать более 2 одновременных соединений с сервером или прокси.»

4.3. Стандартизируемые новинки

Наиболее очевидными разработками в рамках IETF, которые будут влиять на развитие контроля перегрузок, являются дифференцированные и интегрированные услуги [RFC2212, RFC2475], а также явные уведомления о перегрузке (ECN) [RFC2481]. Однако некоторые менее значимые разработки также будут оказывать влияние на контроль перегрузок.

Одним из таких направлений является разработка контроля перегрузок в оконечных точках (ECM[1]) [BS00] для обеспечения общего состояния контроля насыщения для множества потоков от одного отправителя к одному получателю. Позволяя множеству одновременных соединений с одним адресатом функционировать, подобно одному потоку, менеджер перегрузок (Congestion Manager) может разрешить для некоторых соединений замедленный старт, что позволяет воспользоваться преимуществами наличия сквозной информации о состоянии насыщения на пути. Кроме того, использование менеджера перегрузок может устранить опасность открытия множества «сессий» контроля перегрузок для одной пары «отправитель-получатель» и может позволить браузерам создавать множество соединений с одним адресатом.

5. Описание коллапса насыщения

В этом разделе рассматриваются некоторые детали коллапса насыщения, вызываемого недоставленными пакетами, и показано как не контролирующие перегрузку потоки могут влиять на коллапс насыщения в Internet. Раздел в значительной мере базируется на работе [FF99].

Неформально коллапс насыщения возникает, когда рост нагрузки в сети приводит к уменьшению объёма полезной работы, которую сеть способная выполнить. Как описано в разделе 3, коллапс насыщения был впервые отмечен в середине 1980-х годов [RFC896] и был обусловлен неоправданными повторами передачи пакетов TCP, которые ещё находились в пути или уже были приняты получателем. Такой коллапс насыщения мы называем классическим. Этот коллапс стабилен и может приводить к многократному снижению пропускной способности сети [RFC896]. Проблема классического коллапса насыщения была в основном решена за счёт улучшения таймеров и добавления механизмов контроля перегрузки в современных реализациях TCP [Jacobson88].

Другой возможный вариант коллапса насыщения связан с недоставленными пакетами. Такой коллапс возникает в тех случаях, когда пропускная способность сети впустую тратится на передачу пакетов, которые будут отброшены на пути к конечному получателю. Вероятно, это наиболее серьёзная из нерешенных проблем, связанных с коллапсом насыщения, в современной сети Internet. Разные сценарии могут приводить к различному уровню коллапса в части доли пропускной способности сети, остающейся для продуктивной работы. Опасность этого вида коллапса обусловлена ростом числа приложений с открытым контуром (open-loop), не использующих сквозного контроля перегрузок. Ещё большую опасность представляют обычные приложения, которые «увеличивают» скорость передачи в ответ на рост потерь (например, автоматически повышая уровень FEC[2]).

[1]Endpoint Congestion Management.
[2]Упреждающий контроль ошибок.

В таблице 1 приведены результаты сценария с возникновением коллапса насыщения из-за недоставленных пакетов, когда дефицитная пропускная способность расходуется на передачу пакетов, не доходящих до адресата. Для моделирования использовался сценарий с тремя потоками TCP и одним потоком UDP, передаваемыми через один сильно загруженный канал 1,5 Мбит/с. Для подключения всех узлов использовались соединения 10 Мбит/с, за исключением получателя потока UDP, который был подключён по каналу 128 Кбит/с, что составляет лишь 9% от пропускной способности общего канала. Когда скорость отправки пакетов UDP превышает 128 Кбит/с, значительная часть пакетов UDP будет отбрасываться на выходном порту в этот оконечный канал 128 Кбит/с.

Таблица 1. Модель с тремя потоками TCP и одним потоком UDP.

Скорость поступления UDP от отправителя	***Полезная пропускная способность для UDP***	***Полезная пропускная способность для TCP***	***Общая полезная пропускная способность***
0,7	0,7	98,5	99,2
1,8	1,7	97,3	99,1
2,6	2,6	96	98,6
5,3	5,2	92,7	97,9
8,8	8,4	87,1	95,5
10,5	8,4	84,8	93,2
13,1	8,4	81,4	89,8
17,5	8,4	77,3	85,7
26,3	8,4	64,5	72,8
52,6	8,4	38,1	46,4
58,4	8,4	32,8	41,2
65,7	8,4	28,5	36,8
75,1	8,4	19,7	28,1
87,6	8,4	11,3	19,7
105,1	8,4	3,4	11,8
131,5	8,4	2,4	10,7

В таблице 1 указана скорость поступления пакетов UDP от источника, полезная полоса UDP (пропускная способность, для реально доставленных данных), полезная полоса TCP (данные, доставленные пользователям TCP) и агрегатная полезная полоса для насыщенного канала 1,5 Мбит/с. В каждой строке указана доля от общей пропускной способности загруженного канала. По мере роста скорости источника UDP полезная пропускная способность TCP уменьшается почти линейно, а полезная пропускная способность UDP остаётся почти постоянной. Таким образом, рост потока UDP приводит лишь к ухудшению пропускной способности для TCP и падению суммарной полезной пропускной способности. На перегруженном канале поток UDP в конечном счёте отнимает полосу, использовавшуюся для потоков TCP, и снижает пропускную способность сети в целом до весьма незначительной доли пропускной способности канала.

Таблица 1 показывает как отсутствие беспристрастности, так и коллапс насыщения. Как рассмотрено в [FF99], контроль перегрузок не является единственным способом обеспечения беспристрастности, планирование на уровне потоков в загруженных маршрутизаторах обеспечивает другой механизм гарантии беспристрастности. Однако, как было отмечено в [FF99], такое планирование не помогает предотвратить коллапс насыщения.

Есть лишь два способа предотвращения опасности, связанной с коллапсом насыщения в результате недоставки пакетов. В первом варианте используется эффективный сквозной контроль перегрузок на оконечных узлах. Более конкретное требование состоит в предотвращении для потока существенных потерь на каналах, нисходящих по отношению к первому перегруженному каналу на пути (здесь мы называем перегруженным канал в том случае, когда какой-либо поток получает пропускную способность за счёт «ущемления» других потоков в этом канале). С учётом того, что оконечные узлы обычно не способны определить число перегруженных каналов на пути (один или множество), более надёжным способом является предотвращение существенных потерь на перегруженном нисходящем канале за счёт сквозного контроля перегрузок и снижения скорости передачи при возникновении потерь.

Вторым вариантом предотвращения коллапса по причине недоставки пакетов может быть гарантия со стороны сети доставки воспринятых перегруженным узлом пакетов конечному получателю [RFC2212, RFC2475]. Отметим, что выбор одного из этих вариантов не препятствует одновременному использованию другого варианта. Например, для части трафика может применяться сквозной контроль насыщения, а для остального - сетевые гарантии доставки.

6. Формы сквозного контроля перегрузок

В этом документе рассмотрены вопросы, связанные с коллапсом насыщения и беспристрастностью новых форм контроля перегрузок по отношению к TCP. Однако это совсем не означает, что озабоченность такими вопросами требует для обычного трафика TCP (best-effort) развёртывать систему контроля перегрузок на основе алгоритма AIMD[1] со снижением скорости передачи вдвое при каждой потере пакета. В этом разделе специально рассматривается влияние концепций коллапса перегрузок и беспристрастности TCP.

6.1. Сквозной контроль для предотвращения коллапса насыщения

Для предотвращения коллапсов перегрузки, вызванных недоставленными пакетами, требуется, чтобы потоки не передавали пакеты с чрезмерной скоростью, не было множества перегруженных каналов и частого отбрасывания пакетов в нисходящем канале. Поскольку в коллапсе насыщения в результате недоставки пакетов важную роль играют пакеты, занимающие канал и в последствии отбрасываемые в нисходящем канале, такой вариант коллапса перегрузки не может возникать в средах, где каждый поток проходит лишь через один перегруженный канал или в нисходящем канале после первого перегруженного соединения отбрасывается небольшое число пакетов. Таким образом, любой вариант контроля перегрузок, позволяющий избежать чрезмерной скорости передачи при возникновении достаточно больших потерь, сможет предотвратить коллапс насыщения, вызываемый недоставленными пакетами.

Хотелось бы отметить, что добавление явных уведомлений о перегрузке (ECN) в архитектуру IP само по себе не устраняет проблему коллапса насыщения для трафика best-effort. ECN позволяет маршрутизаторам устанавливать битовый флаг в заголовках пакетов для индикации перегрузки оконечным узлам вместо использования в качестве такого индикатора фактов отбрасывания пакетов. Однако маркировку ECN можно использовать вместо отбрасывания

[1]Additive-Increase Multiplicative-Decrease - адаптивный рост, мультипликативное снижение.

пакетов при возникновении незначительной перегрузки. При существенной перегрузке, когда буферное пространство маршрутизатора полностью заполнено, у него не остаётся иного выхода, кроме отбрасывания пакетов.

6.2. Сквозной контроль насыщения для беспристрастности с TCP.

Высказанные в [RFC2357] опасения относительно беспристрастности к TCP накладывают существенные ограничения на выбор решений для сквозного контроля перегрузок применительно к трафику best-effort. Среда с планированием на уровне отдельного потока изолирует потоки друг от друга и снимает требование совместимости механизма контроля насыщения с TCP. Среда с дифференцированными услугами, где потоки маркируются по неким классам diff-serv, будет выполнять планирование независимо от трафика best-effort и это может приводить к тому, что совместимость с TCP не будет требоваться для целого класса diff-serv. Аналогично, управляемая по ценам среда или класс diff-serv со своей ценовой парадигмой могут забыть о вопросе беспристрастности для TCP. Однако в современной среде Internet, где другой трафик best-effort может полностью занимать очереди FIFO пакетами TCP, отсутствие беспристрастности для TCP может приводить к тому, что один поток будет отнимать ресурсы остальных в моменты высокой загрузки, как было показано выше с таблице 1.

Однако список совместимых с TCP процедур контроля перегрузок не ограничивается AIMD с такими же параметрами роста и снижения, как в TCP. Совместимые с TCP процедуры контроля насыщения включают основанные на скорости передачи варианты AIMD с разными наборами параметров роста-снижения, обеспечивающие такое же поведение в стабильном состоянии, основанный на выравнивании контроль насыщения, где отправитель подстраивает скорость передачи в соответствии с информацией о средней частоте отбрасывания пакетов за достаточно большое время, многоуровневая групповая передача, где подписчики могут включать и исключать себя из multicast-группы разных уровней, а также иные формы, которые мы ещё не рассматривали.

7. Благодарности

Значительная часть этого документа заимствована из предшествующих RFC, посвящённых сквозному контролю перегрузок. Этот документ является попыткой обобщения идей, обсуждаемых в течение ряда лет с участием множества людей. В частности, следует отметить участников групп End-to-End Research и Reliable Multicast Research, а также направления Transport Area. Этот документ был существенно развит, благодаря дискуссиям и откликам в рабочей группе Transport Area. Отдельная благодарность Mark Allman за отклики на ранние версии этого документа.

8. Литература

[BS00] Balakrishnan H. and S. Seshan, "The Congestion Manager", Work in Progress[1].

[DMKM00] Dawkins, S., Montenegro, G., Kojo, M. and V. Magret, "End-to-end Performance Implications of Slow Links", Work in Progress[2].

[FF99] Floyd, S. and K. Fall, "Promoting the Use of End-to-End Congestion Control in the Internet", IEEE/ACM Transactions on Networking, August 1999. URL http://www.aciri.org/floyd/end2end-paper.html

[HPF00] Handley, M., Padhye, J. and S. Floyd, "TCP Congestion Window Validation", RFC 2861, June 2000.

[Jacobson88] V. Jacobson, Congestion Avoidance and Control, ACM SIGCOMM '88, August 1988.

[RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC896] Nagle, J., "Congestion Control in IP/TCP", RFC 896, January 1984.

[RFC1122] Braden, R., Ed., "Requirements for Internet Hosts – Communication Layers", STD 3, RFC 1122, October 1989.

[RFC1323] Jacobson, V., Braden, R. and D. Borman, "TCP Extensions for High Performance", RFC 1323, May 1992.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2212] Shenker, S., Partridge, C. and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.

[RFC2309] Braden, R., Clark, D., Crowcroft, J., Davie, B., Deering, S., Estrin, D., Floyd, S., Jacobson, V., Minshall, G., Partridge, C., Peterson, L., Ramakrishnan, K.K., Shenker, S., Wroclawski, J., and L. Zhang, "Recommendations on Queue Management and Congestion Avoidance in the Internet", RFC 2309, April 1998.

[RFC2357] Mankin, A., Romanow, A., Bradner, S. and V. Paxson, "IETF Criteria for Evaluating Reliable Multicast Transport and Application Protocols", RFC 2357, June 1998.

[RFC2414] Allman, M., Floyd, S. and C. Partridge, "Increasing TCP's Initial Window", RFC 2414, September 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.

[RFC2481] Ramakrishnan K. and S. Floyd, "A Proposal to add Explicit Congestion Notification (ECN) to IP", RFC 2481, January 1999.

[RFC2525] Paxson, V., Allman, M., Dawson, S., Fenner, W., Griner, J., Heavens, I., Lahey, K., Semke, J. and B. Volz, "Known TCP Implementation Problems", RFC 2525, March 1999.

[RFC2581] Allman, M., Paxson, V. and W. Stevens, "TCP Congestion Control", RFC 2581, April 1999.

[RFC2582] Floyd, S. and T. Henderson, "The NewReno Modification to TCP's Fast Recovery Algorithm", RFC 2582, April 1999.

[RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.

[1]Работа завершена и опубликована в RFC 3124. *Прим. перев.*
[2]Работа завершена и опубликована в RFC 3150. *Прим. перев.*

[SCWA99] S. Savage, N. Cardwell, D. Wetherall, and T. Anderson, TCP Congestion Control with a Misbehaving Receiver, ACM Computer Communications Review, October 1999.

[TCPB98] Hari Balakrishnan, Venkata N. Padmanabhan, Srinivasan Seshan, Mark Stemm, and Randy H. Katz, TCP Behavior of a Busy Internet Server: Analysis and Improvements, IEEE Infocom, March 1998. Доступен по ссылке http://www.cs.berkeley.edu/~hari/papers/infocom98.ps.gz.

[TCPF98] Dong Lin and H.T. Kung, TCP Fast Recovery Strategies: Analysis and Improvements, IEEE Infocom, March 1998. Доступен по ссылке http://www.eecs.harvard.edu/networking/papers/infocom-tcp-final-198.pdf.

9. Связанные с TCP вопросы

В этом разделе рассматриваются некоторые частные аспекты контроля перегрузок в TCP для иллюстрации принципов реализации контроля насыщения, включая некоторые детали, связанные со встраиванием такого контроля в транспортный протокол.

9.1. Замедленный старт

Отправитель TCP не может создать новое соединение путём передачи большого объёма данных (например, размером в анонсированное получателем окно) сразу. Отправитель TCP ограничен небольшим начальным значением окна насыщения. В течение замедленного старта (slow-start) отправитель TCP может увеличивать свою скорость передачи не более, чем вдвое за каждый период кругового обхода. Процедура замедленного старта завершается при обнаружении перегрузки или в случае, когда размер окна насыщения на стороне отправителя становится больше порога замедленного старта ssthresh.

Проблема, способная повлиять на глобальный контроль перегрузок, и по этой причине явно рассматриваемая в процессе стандартизации, включает увеличение размера начального окна [RFC2414,RFC2581].

К проблемам, которые не будут решаться в процессе стандартизации и в общем случае отнесены к не требующим стандартизации, относятся такие вопросы, как использование (или отказ от него) задания темпа на основе скорости или механизмы ускоренного завершения процедуры slow-start до того, как размер окна достигнет ssthresh. Такие механизмы приводят к поведению процедуры slow-start, столь же или более консервативному по сравнению со стандартным TCP.

9.2. Аддитивный рост, мультипликативное снижение

При отсутствии перегрузки отправитель TCP увеличивает своё окно насыщения не более чем на один пакет за период кругового обхода. При отклике на индикацию перегрузки отправитель TCP уменьшает размер своего окна насыщения вдвое (точнее, размер нового окна насыщения составляет половину от меньшего из значений размера окна насыщения и анонсируемого получателем окна).

Вопрос, который может влиять на глобальный контроль перегрузок и, следовательно, будет в конце концов явно решён в процессе стандартизации, включает предлагаемое добавление контроля насыщения для возврата потока «чистых подтверждений».

В процессе стандартизации не был рассмотрен (и обычно считается не требующим стандартизации) вопрос применения окна перегрузки к верхней границе числа байтов, которые предполагаются находящимися в именованном канале (pipe), вместо его использования в качестве скользящего окна, начинающегося с кумулятивного подтверждения. Ясно, что анонсируемое получателем окно применяется как скользящее окно, начиная с поля кумулятивного подтверждения, поскольку пакеты, полученные «над полем» кумулятивного подтверждения, сохраняются в приёмном буфере TCP и не доставляются приложению. Однако окно насыщения, применённое к числу пакетов, остающихся в именованном канале (pipe), не обязательно включает пакеты, принятые получателем TCP с нарушением порядка.

9.3. Таймеры повтора передачи

Отправитель TCP устанавливает таймер повторной передачи, чтобы определять отбрасывание пакетов в сети. Когда отсчёт этого таймера завершается, отправитель делает вывод о потере пакета, устанавливает для ssthresh значение в половину размера текущего окна и заново передаёт потерянный пакет. Если отсчёт таймера завершается до того, как придёт подтверждение доставки повторно переданного пакета, значение этого таймера для следующего отсчёта увеличивается вдвое.

Вопрос, который может оказать влияние на глобальный контроль перегрузок и по этой причине будет явно решаться в процессе стандартизации, состоит в разработке изменённого механизма установки значения таймера повторной передачи. Изменение механизма может приводить к значительному росту числа повторов передачи в результате завершения отсчёта до прибытия отправителю подтверждения даже без потери пакетов. Это следует принимать во внимание в процессах стандартизации Internet, поскольку преждевременное завершение отсчёта таймеров повтора передачи может приводить к неоправданному росту числа повторно переданных пакетов на загруженных каналах.

9.4. Ускоренный повтор и быстрое восстановление

Увидев три дубликата подтверждений, отправитель TCP делает вывод о потере пакета. В этом случае он устанавливает для ssthresh значение в половину размера текущего окна, делает размер окна насыщения не больше половины от предыдущего окна и заново передаёт потерянный пакет.

Вопрос, который может оказать влияние на глобальный контроль перегрузок и по этой причине будет явно решаться в процессе стандартизации, состоит в предложении (если оно есть) предполагать потерю пакета после одного или двух дубликатов подтверждений. Непродуманная реализация такого предложения может привести к росту числа пакетов, без необходимости передаваемых повторно по загруженному пути.

Вопрос, который не был решён в процессе стандартизации и не считается требующим стандартизации, связан с предложением передавать «новый» или предположительно потерянный пакет в ответ на дубликат подтверждения или частичное подтверждение, если размер окна насыщения это позволяет. Примером этого может служить передача нового пакета в ответ на получение одного дубликата подтверждения, чтобы сохранить «часы подтверждений», если

дальнейших подтверждений не будет. Такое предложение является примером благоприятного изменения, не затрагивающего взаимодействие и не влияющего на глобальный контроль насыщения, которое, следовательно, может быть реализовано производителями без вмешательства процесса стандартизации IETF (этот вопрос был фактически решён в [DMKM00], где сказано: «исследователи могут пожелать провести эксперименты с отправкой нового трафика в сеть при получении дубликатов подтверждений, как описано в [TCPB98] и [TCPF98]»).

9.5. Другие аспекты контроля перегрузок TCP

К другим вопросам контроля насыщения в TCP, не затронутым выше, относится восстановление после простоя или ограниченного приложением периода [HPF00].

10. Вопросы безопасности

Этот документ посвящён рискам, связанным с контролем перегрузок или его отсутствием. В параграфе 3.2 рассмотрена возможная утрата беспристрастности при отсутствии в одновременных потоках совместимых методов контроля насыщения, а раздел 5 посвящён опасностям, связанным с коллапсом насыщения в тех случаях, когда потоки не используют сквозного контроля перегрузок.

Поскольку в этом документе не предлагается каких-либо конкретных механизмов контроля насыщения, в него не требуется включать какие-то конкретные меры, связанные с контролем перегрузок. Однако следует отметить наличие широкого спектра вопросов безопасности, связанных с контролем перегрузок, которые нужно рассмотреть в документах IETF.

Например, конкретным механизмам контроля насыщения следует быть максимально устойчивыми к попыткам отдельных конечных узлов обойти сквозной контроль перегрузок [SCWA99]. Это вызывает особую озабоченность контролем перегрузок при групповой адресации по причине широкого распространения группового трафика и больших возможностей для отдельных получателей отказаться от информирования о перегрузках.

В RFC 2309 рассмотрены возможные опасности для Internet от «безответственных» потоков, которые не снижают скорость передачи при возникновении перегрузок, и описана потребность в механизмах воздействия сети на потоки, не реагирующие на индикацию перегрузки. Отметим, что этот вопрос продолжает требовать дальнейшего исследования, разработки, оценки и внедрения.

Поскольку сеть Internet объединяет в себе огромное число потоков, риск воздействия на инфраструктуру в целом путём обхода контроля перегрузок небольшим числом пакетов невелик. Гораздо опасней внедрение и широкое распространение многочисленных конечных узлов, нарушающих сквозной контроль перегрузок.

Адрес автора

Sally Floyd

AT&T Center for Internet Research at ICSI (ACIRI)

Phone: +1 (510) 642-4274 x189

EMail: floyd@aciri.org

URL: http://www.aciri.org/floyd/

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Полное заявление авторских прав

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.