Internet Engineering Task Force (IETF)
Request for Comments: 9097
Category: Standards Track
ISSN: 2070-1721

A. Morton
AT&T Labs
R. Geib
Deutsche Telekom
L. Ciavattone
AT&T Labs

November 2021

Metrics and Methods for One-Way IP Capacity

Показатели и методы измерения пропускной способности IP в одном направлении

Аннотация

Этот документ пересматривает показатели пропускной способности сети (Network Capacity Metrics), заданные в RFC 5136. Документ задаёт более практичное определение показателя максимальной пропускной способности на уровне IP (Maximum IP-Layer Capacity Metric), относящееся к измерениям и описывает соответствующие методы измерений.

Статус документа

Документ относится к категории Internet Standards Track.

Документ является результатом работы IETF[1] и представляет согласованный взгляд сообщества IETF. Документ прошёл открытое обсуждение и был одобрен для публикации IESG[2]. Дополнительную информацию о стандартах Internet можно найти в разделе 2 в RFC 7841.

Информация о текущем статусе документа, найденных ошибках и способах обратной связи доступна по ссылке https://www.rfc-editor.org/info/rfc9097.

Оглавление

[1]Internet Engineering Task Force - комиссия по решению инженерных задач Internet.
[2]Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

1. Введение

Усилия IETF по определению пропускной способности сети (Network Capacity) и массовой передачи (Bulk Transport Capacity или BTC) предпринимаются и развиваются более 20 лет. За это время сообщество специалистов по производительности стало свидетелем разработки информационных определений в [RFC3148] для модели BTC (Framework for Bulk Transport Capacity), [RFC5136] для производительности сети и максимальной производительности на уровне IP (Maximum IP-Layer Capacity), а также экспериментальных определений показателей и методов в Model-Based Metrics for Bulk Transport Capacity [RFC8337].

Этот документ заново рассматривает вопрос показателей пропускной способности сети (Network Capacity Metrics), обсуждавшийся в [RFC3148], а затем в [RFC5136]. Показатели Maximum IP-Layer Capacity и Bulk Transfer Capacity [RFC3148] (полезная пропускная способность или goodput) отличаются один от другого. Максимальная пропускная способность на уровне IP похожа на теоретическую цель полезной пропускной способности. В [RFC5136] задано множество показателей, таких как доступная пропускная способность (Available Capacity). Измерения зависят от применяемого пути через сеть и варианта применения. Здесь основным вариантом является оценка максимальной пропускной способности (Maximum Capacity) одной или нескольких сетей, где подписчик получает определённые гарантии производительности, иногда называемые доступом в Internet, или где проверяются ограничения технологии, применяемой на тестируемом пути. Например, для пользователя услуги 1 Гбит/с нужно гарантировать эту пропускную способность у пользователя, поставщика услуг (Service Provider или SP) и, возможно, других участников. Когда тест подтверждает согласованный в подписке уровень производительности, узкое место может оказаться в другой части пути.

Этот документ признает важность определения максимальной пропускной способности на уровне IP в то время, когда скорость подписки на Internet резко возросла, - это определение практично и эффективно для измерительного сообщества, а также пользователей Internet. Определения показателей предназначены для активных измерений (Active Methods of Measurement) [RFC7799] и для каждого показателя заданы методы измерения.

В наиболее прямом активном измерении производительности на уровне IP будут применяться пакеты IP, но на практике нужны также транспортные заголовки для прохождения через трансляторы адресов и портов. UDP предлагает наиболее прямую возможность оценки и с исследовании измерений для понимания пригодности UDP в качестве базового транспортного протокола Internet [copycat] авторы обнаружили, что значительная часть протестированных путей поддерживает UDP. По этой причине были заменены некоторые заявления о взаимодействии [LS-SG12-A] [LS-SG12-B], где указаны лабораторные и полевые тесты, поддерживающие применение UDP для измерения производительности на уровне IP.

Этот документ также признаёт изменения модели показателей производительности IP (IP Performance Metrics или IPPM) [RFC2330], опубликованные с 1998 г. В частности, используется [RFC7312] для расширенной модели потоков и выборки (Advanced Stream and Sampling Framework) и [RFC8468] для обновлений сосуществования IPv4, IPv6 и IPv4-IPv6.

В Приложении A описан алгоритм настройки нагрузки с использованием псевдокода, а в Приложении B рассмотрена совместимость алгоритма с [RFC8085].

1.1. Уровни требований

Ключевые слова **должно** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **не рекомендуется** (NOT RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с BCP 14 [RFC2119] [RFC8174] тогда и только тогда, когда они выделены шрифтом, как показано здесь.

2. Область действия, цели, применимость

Целью этого документа является определение показателей активных измерений, соответствующих методов однозначного определения максимальной пропускной способности на уровне IP и полезных вторичных показателей. Другая цель состоит в согласовании заданного показателя и метода для всей отрасли и этот документ фиксирует согласие IETF, что может привести к изменению спецификаций других органов стандартизации (Standards Development Organization или SDO) через обычные процессы внесения вклада каждого SDO и взаимодействие).

Вторичными целями являются рассмотрение тестовых процедур и интерпретация результатов измерения максимальной пропускной способности на уровне IP (для выявления случаев, когда нужны дополнительные тесты, возможно, с иной конфигурацией). Содействие развитию протокольной поддержки для этого показателя и метода измерений также является целью документа (все протоколы активных тестов, заданные рабочей группой IPPM, основаны на UDP, что соответствует основному требованию этих методов). Разработка вспомогательного протокола для измерения этого показателя в соответствии с заданным методом является важным вкладом в измерения Internet.

Область действия алгоритма регулировки величины нагрузки ограничена определением максимальной пропускной способности на уровне IP в контексте нечастых и кратковременных диагностических измерений. Во время измерений **рекомендуется** прервать иной трафик, потребляющий выделенные ресурсы, чтобы это не снижало точности измерений максимальной пропускной способности.

Основным применением описанных здесь показателей и методов измерения является то же, что описано в разделе 2 [RFC7497]

> В центра внимания находится связанная с доступом часть сети. Пользователи обычно подписываются на двунаправленные услуги доступа [Internet], частично описываемые скоростью в бит/сек.

Кроме того, использование описанного в параграфе 8.1 алгоритма регулировки нагрузки задаёт другие ограничения.

- Алгоритм **должен** применяться лишь в приложениях диагностики и рабочих измерений, как описано здесь.
- Алгоритм **должен** применяться лишь в обстоятельствах, соответствующих разделу 10. Вопросы безопасности.
- Если оператор сети уверен в необходимости проверки пропускной способности на уровне IP, он **может** начать тест с фиксированной скоростью и не применять алгоритм регулирования нагрузки. Однако потребности в диагностических тестах (например, запросы абонентов) строго подразумевают отсутствие такой уверенности и **рекомендуется** применять алгоритм регулировки нагрузки.

Показатели и методы измерения предназначены для измерений в случаях, когда точные сведения о пути неизвестны в диапазоне возможных значений.

- Точная максимальная пропускная способность на уровне IP для подписчика неизвестна (иногда это так, скорости обслуживания могут расти без запроса подписчиков в результате обновлений или для компенсации возможных недооценок при тестировании на основе TCP).
- Неизвестен размер буферов в узких местах (bottleneck).

Алгоритму регулирования нагрузки в измерительной системе **не следует** заранее задавать точное значение пропускной способности для его проверки. Это позволяет получить беспристрастный результат и устраняет возможность нечестных операций, когда нужный результат известен до измерений.

3. Мотивация

Как и для других задач, где многие годы разные SDO работали без согласования, возникли разные решения для показателей и методов измерения. Имеется 5 факторов, которые изменились (или начали меняться) в период 2013-2019 гг., и наличие любого из них на пути измерений требует их учёта.

1. Для многих пользователей доступ в Internet перестал быть узким местом, но они ожидают от провайдеров исполнения условий договоров в части пропускной способности.
2. Скорость передачи и величина задержки важны для удовлетворения пользователей.
3. Роль транспорта UDP растёт в областях, где ранее доминировал протокол TCP.
4. Содержимое и приложения перемещаются ближе к пользователям.
5. Измерениям для шлюзов провайдеров (ISP) уделяется меньше внимания, возможно потому, что в будущем трафик не будет проходить через эти шлюзы.

4. Общие параметры и определения

В этом разделе указаны **требуемые** входные факторы для показателей отправителя или получателя.

Src
: Один из адресов хоста (например, глобально маршрутизируемый адрес IP).

Dst
: Один из адресов хоста (например, глобально маршрутизируемый адрес IP).

MaxHops
: Число интервалов пересылки (hop), через которые конкретный пакет может пройти от Src к Dst (TTL или Hop Limit).

T0
: Время начала интервала измерений (передача первого пакета от источника).

I
: Номинальная продолжительность интервала измерений у получателя (по умолчанию 10 секунд).

dt
: Номинальная продолжительность m равных субинтервалов в рамках I у получателя (по умолчанию 1 секунда).

dtn
: Начало конкретного субинтервала n (одного из m субинтервалов в I).

FT
: Интервал между сообщениями обратной связи о состоянии, содержащими результаты измерений, которые получатель передаёт для управления отправителем. Результаты оцениваются в течение всего теста для определения настройки текущей предлагаемой нагрузки у отправителя (по умолчанию 50 мсек).

Tmax
: Максимальное время ожидания прибытия тестовых пакетов к получателю, достаточно большое, чтобы различать длительные задержки и отбрасывание (потерю) пакетов для предотвращения отчечки распределения задержки в одном направлении.

F
: Число различных потоков, создаваемых методом (по умолчанию 1).

Flow
: Поток пакетов с одинаковым кортежем из n назначенных полей заголовка, которые (в предположении их постоянства) приводят к идентичной трактовке при выборе пути (например, при распределении нагрузки). Отметим, что метки потоков IPv6 **следует** включать в определение потока, если маршрутизаторы соответствуют рекомендациям [RFC6438].

Type-P

Полное описание тестовых пакетов, для которых применяется эта оценка (включая поля указания потока). Отметим, что для заданных ниже тестовых пакетов одним из требований является транспорт UDP. Концепция Type-P аналогичная интересующей совокупности (population of interest) из п. 6.1.1 в [Y.1540].

Payload Content - содержимое пакетов

Один из аспектов параметра Type-P, который может повысить детерминированность измерений. Задание содержимого пакетов помогает обеспечить соответствие показателей и измерений схеме IPPM. Если на пути выполняется сжатие содержимого и тесты предназначены для оценки влияния сжатия на пропускную способность, в качестве содержимого **следует** применять псевдослучайные значения, используя часть сжатого файла или иной метод (см. параграф 3.1.2 в [RFC7312]).

PM

Список основных показателей, таких как потери, задержка и переупорядочение, а также соответствующих целевых порогов производительности. **Должны** указываться хотя бы один основной показатель и целевой порог производительности (например, отсутствие потерь пакетов IP в одном направлении [RFC7680]).

Для нескольких показателей также требуется значение, не являющееся параметром.

T

Время хоста в момент прибытия **первого** пакета, определённое в целевой точке измерения (Destination Measurement Point или MP(Dst)). Могут быть другие пакеты, переданные между хостами отправителя и получателя, поэтому нужно фиксировать прибытие первого измерительного пакета.

Отметим, что формат и разрешение (дискретность) меток времени, порядковые номера и т. п. могут устанавливаться в соответствии со стандартом или реализацией применяемого для теста протокола.

5. Одиночное измерение пропускной способности на уровне IP

В этом разделе заданы требования к одиночному (Singleton) измерению, поддерживающему определения раздела 6. Определения показателей максимальной пропускной способности.

5.1. Формальное название

Показатель называется Type-P-One-way-IP-Capacity, неформальным именем служит IP-Layer Capacity. Отметим, что Type-P зависит от выбранного метода.

5.2. Параметры

В этом параграфе заданы **требуемые** входные факторы для показателя в дополнение к заданным в разделе 4.

Дополнительные параметры не требуются.

5.3. Определения показателей

В этом параграфе заданы **требуемые** аспекты показателя производительности на уровне IP (если не указано иное) для измерений между хостами отправителя и получателя.

Определим пропускную способность на уровне - C(T,dt,PM), как число битов уровня IP (включая заголовок и данные) в пакетах, которые могут быть переданы от хоста Src и корректно получены хостом Dst в течение 1 субинтервала продолжительностью dt. Это значение зависит от хостов Src и Dst, их адресов и пути между хостами. Число этих битов уровня IP для конкретного dt обозначим n0[dtn,dtn+1].

При известном и фиксированном размере пакетов число пакетов в субинтервале измерения dt, умноженное на число битов в заголовке и данных пакета IP будет равно n0[dtn,dtn+1].

В преположении выборки одиночных измерений количество субинтервалов с длительностью dt **должно** быть задано натуральным числом m, так что T+I = T + m*dt при dtn+1 - dtn = dt для 1 <= n <= m.

Параметр PM представляет другие показатели производительности (см. 5.4. Время кругового обхода и потери в одном направлении), результаты их измерения **нужно** собирать в интервале измерения пропускной способности уровня IP и связывать с соответствующим dtn для последующей оценки и указания в отчёте. Пользователю **нужно** задать параметр Tmax в соответствии с требованиями определения каждого показателя.

Математически это определение можно представить для каждого n уравнением

$$C(T,dt,PM) = \frac{(\ n0[dtn,dtn+1]\ )}{dt}$$

Рисунок 1. Уравнение для пропускной способности уровня IP.

n0

Общее число битов уровня IP в заголовках и данных, переданное в пакетах стандартного формата [RFC8468] от хоста Src и корректно полученных хостом Dst в одном непрерывном субинтенрвале dt в течение интервала [T,T+I].

C(T,dt,PM)

Пропускная способность уровня IP равна значению n0, измеренному в любом субинтервале, начиная с dtn, разделённому на продолжительность субинтервала (dt).

PM

Представляет другие показатели производительности (см. 5.4. Время кругового обхода и потери в одном направлении), результаты их измерения **нужно** собирать в интервале измерения пропускной способности уровня IP и связывать с соответствующим dtn для последующей оценки и указания в отчёте.

Все субинтервалы **должны** быть одинаковы. Выбор неперекрывающихся последовательных интервалов dt упрощает реализацию.

Битовая скорость физического интерфейса в измерительных устройствах **должна** быть выше наименьшей скорости на каналах пути измерения C(T,I,PM), т. е. наиболее узкого места.

При измерениях на основе этого определения **нужно** использовать транспортный уровень UDP. Стандартные пакеты описаны в разделе 5 [RFC8468]. При измерениях **следует** применять случайный номер порта-источника отправителя или аналогичный метод, а отклики **следует** передавать с адреса, на который были направлены тестовые пакеты.

Некоторые вопросы влияния сжатия на измерения рассмотрены в разделе 6 [RFC8468].

5.4. Время кругового обхода и потери в одном направлении

RTD[dtn,dtn+1] определяется как время кругового обхода (Sample of the Round-Trip Delay) [RFC2681] для выборки между хостами Src и Dst в течение интервала [T,T+I] (набор непрекрывающихся интервалов dt). Разумный период времени из [RFC2681] в этом документе обозначен Tmax. Статистика, применяемая для получения RTD[dtn,dtn+1], **может** включать минимальное, максимальное, медианное, среднее значение и диапазон = (максмум - минимум). Некоторые из этих параметров статистики нужны для настройки нагрузки (8.1. Алгоритм настройки нагрузки), классификации измерений (8.2. Уточняющее измерение или проверка) и отчётов (9. Форматы отчётов).

OWL[dtn,dtn+1] определяется как потери в одном направлении (One-Way Loss) [RFC7680] для выборки между хостами Src и Dst в течение интервала [T,T+I] (набор непрекрывающихся интервалов dt). Статистика, применяемая для получения OWL[dtn,dtn+1], **может** включать число или долю потерянных пакетов.

Можно измерять и другие показатели в одном направлении - нарушение порядка, дублирование и вариации задержки.

5.5. Обсуждение

См. 6.5. Обсуждение.

5.6. Отчёт о показателе

Пропускную способность уровня IP **следует** сообщать с дискретностью 1 Мбит/с (1000000 бит/с). Связанные значения потерь в одной направлении и задержки кругового обхода для одного одиночного измерения **нужно** указывать в отчёте с подходящей дискретностью и единицами измерения.

Отдельные измерения пропускной способности **можно** сообщать в стиле, описанном в разделе 9. Форматы отчётов.

6. Определения показателей максимальной пропускной способности

В этом разделе заданы требования к компонентам поддержки показателя Maximum IP-Layer Capacity.

6.1. Формальное название

Показатель называется Type-P-One-way-Max-IP-Capacity, и имеет неформальное имя Maximum IP-Layer Capacity. Отметим, что Type-P зависит от выбранного метода.

6.2. Параметры

В этом параграфе заданы **требуемые** входные факторы для показателя в дополнение к заданным в разделе 4.

Дополнительные параметры не требуются.

6.3. Определение показателя

В этом параграфе заданы **требуемые** аспекты показателя максимальной производительности на уровне IP (если не указано иное) для измерений между хостами отправителя и получателя.

Определим пропускную способность на уровне - Maximum_C(T,dt,PM), как максимальное число битов уровня IP (включая заголовок и данные) в пакетах, которые могут быть переданы от хоста Src и корректно получены хостом Dst в течение всех субинтервалов продолжительностью dt в [T,T+I] и соответствуют критериям PM. Эквивалентным определением будет максимум выборок одиночных измерений размером m C(T,I,PM), собранных в интервале [T,T+I] и удовлетворяющих критериям PM.

Количество субинтервалов с длительностью dt **должно** быть задано натуральным числом m, так что T+I = T + m*dt при dtn+1 - dtn = dt для 1 <= n <= m.

Параметр PM представляет другие показатели производительности (см. 6.4. Время кругового обхода и потери в одном направлении) и результаты их измерения при измерении максимальной пропускной способности уровня IP. **Должен** быть задан хотя бы 1 целевой порог производительности (критерий PM). Если задано несколько показателей и целевых порогов производительности, субинтервал с максимальным числом переданных битов **должен** соответствовать всем целевым порогам производительности. Пользователю **нужно** задать параметр Tmax в соответствии с требованиями определения каждого показателя.

Математически это определение можно представить уравнением

```
                                max  ( n0[dtn,dtn+1] )
                                [T,T+I]
  Maximum_C(T,I,PM) = --------------------------
                                     dt
где
    T                                        T+I
    _____________________________________________
    |   |   |   |   |   |   |   |   |   |   |
dtn=1   2   3   4   5   6   7   8   9  10  n+1
                                       n=m
```

Рисунок 2. Уравнение для максимальной пропускной способности.

n0

Общее число битов уровня IP в заголовках и данных, переданное в пакетах стандартного формата [RFC8468] от хоста Src и корректно полученных хостом Dst в одном непрерывном субинтенрвале dt в течение интервала [T,T+I].

Maximum_C(T,dt,PM)

Максимальная пропускная способность уровня IP равна значению n0, измеренному в любом субинтервале, начиная с dtn, разделённому на постоянную продолжительность субинтервала (dt).

PM

Представляет другие показатели производительности (см. 6.4. Время кругового обхода и потери в одном направлении) и результаты их измерения при измерении максимальной пропускной способности уровня IP. **Должен** быть задан хотя бы 1 целевой порог производительности (критерий PM).

Все субинтервалы **должны** быть одинаковы. Выбор неперекрывающихся последовательных интервалов dt упрощает реализацию.

В этом определении m субинтервалов можно рассматривать как испытания, в которых хост Src меняет скорость передачи пакетов, отыскивая максимальное значение n0, соответствующее критериям PM, измеренным на хосте Dst в тесте продолжительностью I. Когда хост Src не меняет скорость передачи пакетов, m субинтервалов можно считать попытками оценить стабильность n0 и показатели из списка PM по всем субинтервалам dt в I.

Измерения в соответствии с этим определением **нужно** выполнять для транспортного уровня UDP.

6.4. Время кругового обхода и потери в одном направлении

RTD[dtn,dtn+1] и OWL[dtn,dtn+1] определены в параграфе 5.4. Здесь интервалы тестирования (RTD[T,I] и OWL[T,I]) увеличены в соответствии с выборками пропускной способности.

Интервал dtn,dtn+1, где наблюдается Maximum_C(T,I,PM), является отчетным субинтервалом для RTD[dtn,dtn+1] и OWL[dtn,dtn+1] в рамках RTD[T,I] и OWL[T,I].

Можно измерять и другие показатели в одном направлении - нарушение порядка, дублирование и вариации задержки.

6.5. Обсуждение

Если применяется кондиционирование (например, формовка, правила) трафика на пути, где измеряется Maximum_C(T,I,PM), **следует** выбирать разные dt и выполнять измерения в нескольких интервалах [T,T+I]. Продолжительность каждого dt **следует** выбирать так, чтобы она была кратна возрастающим целочисленным значениям k, умноженным на задержку сериализации Path MTU (PMTU) на физическом интерфейсе, где предполагается кондиционирование трафика. Это должно предотвратить восприятие устройчивых к пикам настроенных одиночных измерений как действительных результатов Maximum_C(T,I,PM).

Maximum_C(T,I,PM) без указания на перегрузку в узких местах, будь то рост задержки, потреря пакетов или маркировка ECN[1] в течение интервала измерения I будет, скорей всего, иметь заниженное значение.

6.6. Отчёт о показателях

Значение пропускной способности уровня IP **следует** указывать с разрешением не хуже 1 Мбит/с в мегабитах за секунду (Mbps - 1000000 бит/с). **Нужно** указывать соответствующие значения потерь в одном направлении и времени кругового обхода с осмысленной дискретностью и единицами измерения.

Когда в выборке демонстрируются и воспроизводятся режимы пропускной способности, максимальную пропускную способность уровня IP **нужно** указывать для каждого режима вместе с временем относительно начала потока, когда этот режим наблюдался. Бимодальная максимальная пропускная способность уровня IP, наблюдавшаяся с некоторыми службами, иногда называется турбо-режимом (turbo mode), предназначенным для ускоренной доставки коротких передач или сокращения начального времени буферизации для некоторых видеопотоков. Отметим, что режимы с длительностью меньше dt не будут обнаружены.

В некоторых технологиях передачи имеется несколько методов работы, которые могут активироваться при улучшении или ухудшении условий в канале, и эти методы передачи могут определять максимальную пропускную способность уровня IP. Примеры этого включают СВЧ-модуляторы для прямой видимости или модемы сотовых сетей, где изменения могут быть вызваны перемещением пользователя из одной области видимости в другую. Работа с разными методами передачи может наблюдаться в течение некоторого времени, но режимы максимальной производительности уровня IP не будут активироваться детерминированно, как у описанном выше турбо-режиме.

7. Одиночное измерение битовой скорости уровня IP у отправителя

В этом разделе приведены требования к компонентам поддержки показателя битовой скорости уровня IP у отправителя. Этот показатель помогает убедиться, что отправитель действительно создаёт желаемую скорость в процессе тестирования, а измерения происходят на интерфейсе между хостом Src и сетевым путём (или как можно ближе к хосту Src). Это не является показателем производительности.

7.1. Формальное название

Показатель называется Type-P-IP-Sender-Bit-Rate или неформально IP-Layer Sender Bit Rate. Отметим, что Type-P зависит от выбранного метода.

7.2. Параметры

В этом параграфе даны **требуемые** входные факторы для задания показателя в дополнение у указанным в разделе 4.

S

Продолжительность интервала измерения у источника.

st

Минимальная продолжительность N субинтервалов в S (по умолчанию st = 0,05 сек.).

stn

Начало конкретного субинтервала n из N субинтервалов в S.

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

S **нужно** задавать значение больше I, в первую очередь для учёта активизации пути по запросу или вводной части тестирования, а также задержки в пути.

Значение st **следует** делать короче субинтервала dt и одного порядка с FT, в ином случае измерение скорости будет включать настройку скорости и дополнительное сглаживание по времени, возможно сглаживание интервала, содержащего Maximum IP-Layer Capacity (с потерей актуальности). Параметр st не имеет значения при передаче источником с фиксированной скоростью в интервале S.

7.3. Определение показателя

В этом параграфе обсуждаются **требуемые** аспекты показателя IP-Layer Sender Bit Rate (если не указано иное) для измерения в заданном источнике по пакетам, адресованным указанному хосту получателя и соответствующим требуемому Type-P.

Определим битовую скорость уровня IP у отправителя B(S,st) как число битов в пакетах уровня IP (включая заголовок и данные), которые передаются от источника с парой адресов Src и Dst в течение одного непрерывного субинтервала st в интервале измерений S (S **нужно** задавать больше I) и подсчёт пакетов фиксированного размера в течение одного субинтервала st даёт также число битов уровня IP в любом из интервалов [stn,stn+1].

При измерениях в соответствии с этим определением **нужно** применять транспорт UDP. Отклики от хоста Dst, полученные хостом Src в интервале измерения [stn,stn+1] **не следует** применять для изменения кондиционирования трафика от Src в течение этого интервала (настройка скорости происходит на границах интервалов st).

7.4. Обсуждение

Битовые скорости у отправителя и получателя **следует** оценивать как часть измерения пропускной способности уровня IP, в ином случае может возникнуть неожиданное ограничение скорости передачи, ведущее к онибке измерения максимальной пропускной способности уровня IP.

7.5. Отчёт о показателе

Битовую скорость уровня IP у источника **нужно** сообщать с осмысленной дискретностью в Мбит/с (1000000 бит/с). Отдельные измерения битовой скорости уровня IP у источника рассматриваются в разделе 9. Форматы отчётов.

8. Метод измерения

В соответствии с архитектурой метода **требуются** хосты в роли источника (Src) и получателя (Dst) с путём измерений и путём возврата между ними.

Продолжительность теста I **должна** ограничиваться в рабочей сети, поскольку метод является активным и вероятно будет вызывать перегрузку на пути от Src к Dst в процессе измерения.

8.1. Алгоритм настройки нагрузки

Описанный в этом параграфе алгоритм недопустимо применять в качестве аплгоритма контроля перегрузок (Congestion Control Algorithm или CCA). Как отмечено в разделе 2. Область действия, цели, применимость, целью алгоритма настройки нагрузки является помощь при определении максимальной пропускной способности уровня IP в контексте кратковременных и нечастых диагностических измерений. Нужен компромисс между продолжительностью теста (объёмом тестовых данных) и агрессивностью алгоритма (темпом нарастания и снижения скорости). Значения параметров, выбранные ниже, обеспечивают проверенный баланс между этими факторами.

Нужна подготовленная (администратором теста) таблица, задающая все скорости, которые будут поддерживаться в тесте (от R1 до Rn по возрастанию, соответствующие проиндексированным строкам таблицы). **Рекомендуется** начинать со скорости 0,5 Мбит/с (индекс 0), использовать скорость 1 Мбит/с с индексом 1 и далее продолжать наращивание по 1 Мбит/с до 1 Гбит/с. От 1 до 10 Гбит/с **рекомендуется** шаг 100 Мбит/с, а выше 10 Гбит/с **рекомендуется** шаг 1 Гбит/с. Можно задаёт более высокое начальное значение IP-Layer Sender Bit Rate, если администратор теста уверен, что Maximum IP-Layer Capacity превышает это начальное значение, а продолжительность тестирования и суммарный тестовый трафик имеют важное значение. Таблице скоростей передачи **следует** задавать скорости по обе стороны от максимума («вилка»), ограничивая по возможности в окрестностях максимума приращение скорости величиной 500 Кбит/с.

Скорость определяется размером дейтаграмм (ss), их числом (cc) в блоке (burst) продолжительностью tt (по умолчанию 100 мксек, близко к тактовому интервалу системы). Хотя выгодно использовать дейтаграммы как можно большего размера, может оказаться разумным некоторое сокращение размера, позволяющее разместить заголовки вторичного протокола и/или туннелирование без фрагментации на уровне IP. Выбор нового значения скорости указывается от текущей строки, например,

```
Rx+1:       отправитель использует следующую строку таблицы.
Rx-10:      отправитель использует строку за 10 до текущей.
```

В начале теста отправитель передаёт со скоростью R1, а получатель запускает таймер обратной связи на время FT (ожидание входящей дейтаграммы). При получении дейтаграммы проверяется отсутствие аномалий в порядковом номере (потери, нарушение порядка, дубликаты и т. п.) и измеряется задержка (одностороняя и круговая). Эти сведения накапливаются до завершения отсчёта FT, а затем отправителю передаётся сообщение о состоянии с собранными сведениями. Собранную статистику получатель сбрасывает, чтобы собирать новую в следующем интервале обратной связи. При получении отправителем сообщения обратной связи тот оценивает сведения для корректировки текущего значения скорости (Rx).

Если обратная связь показывает отсутствие аномалий в порядковых номерах **и** диапазон задержки меньше нижнего порога, скорость передачи увеличивается. Если до этого не было подтверждения перегрузки (см. ниже), скорость можно увеличить более чем на 1 шаг (например, Rx+10). Это позволяет быстрее достичь скорости, близкой к максимальной. Если же перегрузка была подтверждена, скорость увеличивается лишь на 1 шаг (Rx+1). Однако после достижения порога (такого как 1 Гбит/с) скорость каждый раз увеличивается на 1 шаг независимо от перегрузки.

Если обратная связь указывает аномалии порядковых номеров **или** диапазон задержки выше верхнего порога, скорость передачи снижается. **Рекомендуется** устанавливать порог 10 для пропуска номеров, 30 мсек для нижней границы диапазона задержки и 90 мсек для верхней. Кроме того, если перегрузка теперь в первый раз подтверждается при обработке сообщения обратной связи, предлагаемая нагрузка снижается более чем на один шаг (например, Rx-30). Такое однократное снижение предназначено для компенсации быстрого роста в начале. В остальных случаях скорость снижается лишь на один шаг(Rx-1).

Если обратная связь показывает отсутствие аномалий в порядковых номерах **и** диапазон задержки между нижним и верхним порогом, предлагаемая нагрузка не изменяется. Это даёт время для стабилизации и обратной связи, более точно представляющей текущие условия.

В конечном итоге вывод о перегрузке делается на основании аномалии в порядковых номерах **и/или** диапазоне задержки больше верхнего порога в течение трёх последовательных интервалов обратной связи. Описанный выше алгоритм проиллюстрирован в Приложении B к ITU-T Recommendation Y.1540 версии 2020 [Y.1540] и реализован в Приложении A к этому документу (Приложение A. Псевдокод алгоритма настройки нагрузки).

Алгоритм настройки нагрузки **должен** включать таймеры, останавливающие тест при неожиданном прекращении потока принимаемых пакетов. Значения тайм-аутов приведены в таблице 1 вместе другими параметрами и переменными, описанными в этом параграфе. Ниже указаны операции с неочевидными параметрами.

load packet timeout - тайм-аут нагрузочных пакетов

Таймер ожидания пакетов нагрузки **нужно** сбрасывать в заданное значение при получении каждого нагрузочного пакета. При возникновении тайм-аута получателю нужно завершить работу и больше не отправлять сообщений.

feedback message timeout - тайм-аут сообщений обратной связи

Таймер ожидания сообщений обратной связи **нужно** сбрасывать в заданное значение при получении каждого сообщения обратной связи. При возникновении тайм-аута отправителю нужно завершить работу и больше не передавать пакетов.

Таблица 1. Параметры для алгоритма настройки нагрузки.

Параметр	***Значение по умолчанию***	***Тестируемые значения или диапазоны***	***Ожидаемый безопасный диапазон (не проверен полностью, другие значения НЕ РЕКОМЕНДУЮТСЯ)***
FT, интервал обратной связи	50 мсек	20 мсек, 50 мсек, 100 мсек	20 мсек <= FT <= 250 мсек, большие значения могут снижать рост скорости и вызывать отказ при поиске максимума
Тайм-аут для обратной связи (остановка теста)	L*FT, L=20 (1 сек. при FT=50 мсек)	L=100 с FT=50 мсек (5 сек.)	0,5 сек. <= L*FT <= 30 сек., верхний предел только для очень ненадёжных путей
Тайм-аут ожидания пакета (остановка теста)	1 сек.	5 сек.	0,250-30 сек., верхний предел только для очень ненадёжных путей
Индекс таблицы 0	0,5 Мбит/с	0,5 Мбит/с	При скорости теста <= 10 Гбит/с
Индекс таблицы1	1 Мбит/с	1 Мбит/с	При скорости теста <= 10 Гбит/с
Размер (шаг) индекса таблицы	1 Мбит/с	1 Мбит/с <= rate <= 1 Гбит/с	Совпадает с тестируемым
Размер (шаг) индекса таблицы, скорость > 1 Гбит/с	100 Мбит/с	1 Гбит/с <= rate <= 10 Гбит/с	Совпадает с тестируемым
Размер (шаг) индекса таблицы, скорость > 10 Гбит/с	1 Гбит/с	Не тестируется	>10 Гбит/с
ss, размер данных UDP в байтах	Нет	<=1222	Рекомендуется максимальное значение, позволяющее избежать фрагментирования, использование слишком малого значения может необоснованно ограничивать отправителя
cc, число пакетов в блоке (burst)	Нет	1 <= cc <= 100	Совпадает с тестируемым. По мере необходимости cc варьируется для создания желаемого максимума скорости передачи. Размер буфера у отправителя может ограничивать cc в реализации
tt, продолжительность блока	100 мксек	100 мксек, 1 мсек	Доступный диапазон значений tick (параметр HZ)
Нижний порог диапазона задержки	30 мсек	5 мсек, 30 мсек	Совпадает с тестируемым
Верхний порог диапазона задержки	90 мсек	10 мсек, 90 мсек	Совпадает с тестируемым
Порог числа последовательных ошибок	10	0, 1, 5, 10, 100	Совпадает с тестируемым
Порог для отчёта о последовательных ошибках	3	2, 3, 4, 5	Используются значения >1 для исключения ложных промежуточных ошибок
Быстрое увеличение в шагах индекса таблицы	10	10	2 <= число шагов <= 30
Быстрое снижение в шагах индекса таблицы	3 * быстрое увеличение	3 * быстрое увеличение	Совпадает с тестируемым

С принятыми по умолчанию параметрами число шагов в таблице для скоростей меньше 10 Гбит/с составит 1090 (без индекса 0). Соответствующий откат отправителя на условия в сети возникает при отсутствии 1 или нескольких сообщений обратной связи.

Если отправитель не получает обратной связи в течение времени больше тайм-аута Lost Status Backoff = UDRT + (2+w)*FT, где: UDRT = верхний порог диапазона задержки (по умолчанию 90 мсек), FT = интервал ожидания обратной

связи (по умолчанию 50 мсек), w = число тайм-аутов (исходно w=0 и растет на 1 при каждом тайм-ауте, приём сообщения сбрасывает счётчик), начиная с момента приёма отправителем последнего сообщения (любого типа), предлагаемую нагрузку **нужно** снижать как в случае обратной связи, указывающей аномалии в порядковых номерах **или** диапазон задержки больше верхнего порога (см. выше) с текущими значениями переменных найтройки нагрузки. Это означает, что потеря сообщений обратной связи, **или** ошибки в порядковых номерах, **или** вариации задержки могут приводить к снижению скорости и подтверждению перегрузки.

Рекомендуемым начальным значением w является 0 при условии интервала кругового обхода (Round-Trip Time или RTT) меньше значения FT. Значение RTT больше FT является веской причиной должным образом увеличить начальное значение w. Переменную w **нужно** увеличивать на 1 при каждом тайм-ауте Lost Status Backoff. Таким образом, при FT = 50 мсек и UDRT = 90 мсек потеря сообщения обратной связи может быть зафиксирована через 190 мсек после успешного сообщения, вторая - еще через 50 мсек (итого, 240 мсек) и т. д.

Если перегрузка впервые подтверждается тайм-аутом Lost Status Backoff, предлагаемая нагрузка сокращается больше чем на 1 шаг (например, Rx-30). Это однократное снижение предназначено для компенсации быстрого начального роста. В остальных случаях предлагаемая нагрузка снижается не 1 шаг (Rx-1).

В Приложении B обсуждается соответствие применимых обязательных требований [RFC8085] целям для показателя и метода измерения пропускной способности уровня IP, включая описанный здесь алгоритм.

8.2. Уточняющее измерение или проверка

Нужно откалибровать оборудование, измеряющее пропускную способность уровня IP, чтобы убедиться в достаточно точности измерений и соответствии оборудования диапазону измерений (скорость обработки, пропускная способность интерфейса и т. п.).

При оценке максимальной скорости, как указано для показателя, могут получаться завышенные значения, пока буферы на пути не будут заполнены. Другими причинами могут быть блоки (burst) следующих друг за другом пакетов с добавленными в сети интервалами при малом значении интервала измерения (dt), согласованном с блоками. Эти «искусственные» значения могут давать неустойчивые результаты измерения пропускной способности при поиске максимума. Такая ситуация отличается от скоростей бимодальных служб (см. 6.6. Отчёт о показателях), для которых характерна многосекундная длительность (значительно больше измеренного RTT) и повторяющееся поведение.

Есть много способов решения проблемы ложного максимума. Принятое по умолчанию значение для одиночных измерений (dt = 1 сек.) подтвердило свою практическую ценность во время испытаний этого метода, позволяя охарактеризовать скорости бимодальных служб, и имеет очевидную согласованность с единицами измерения (Мбит/с).

Другой подход взят из раздела 24 в [RFC2544] и обсуждения там продолжительности измерений, когда за относительно короткими испытаниями, выполняемыми как часть поиска, следуют более долгие измерения для получения финального результата. В рабочей сети одиночные измерения и выборки (термины для испытаний и тестов Lab Benchmarking) должны быть ограничены по продолжительности из-за возможного влияния на обслуживание. Однако достаточную ценность имеет повторение выборки с фиксированной скоростью передачи, определенной при поиске Maximum IP-Layer Capacity, для получения других показателей производительности, измеренных в то же время.

Уточняющим измерением для результата поиска являетс я последующее измерение со скоростью 99,x% от максимальной пропускной способности уровня IP в интервале I или в течение неопределённого времени. Применяется тот же показатель максимальной пропускной способности, а уточнением результата является выборка без сверхпороговых потерь пакетов или тенденции роста минимальной задержки в последующих одиночных измерениях (или каждом dt измерительного интервала I). Выборки, демонстрирующие сверхпороговые потери пакетов или рост занятости очереди, требуют повторного поиска и/или теста при сниженной фиксированной скорости у отправителя.

Как в любом активном тесте пропускной способности, продолжительность тестирования должна быть короткой. 10-секундные тесты в каждом направлении передачи сегодня являются нормой. По умолчанию принят интервал измерения I в 10 секунд. Сочетание быстрого метода поиска с учётом перегрузок и координации между пользователем и сетью вносит уникальный вклад в рабочее тестирование. Показатель и метод измерения максимальной пропускной способности IP для оценки производительности очень сильно отличаются от классических показателя и методов определения пропускной способности [RFC2544] использованием настройки нагрузки в масштабе времени, близком к реальному, чувствительной к потерям и задержке, а также ограниченной продолжительностью. Измерения пропускной способности в соответствии с [RFC2544] могут создавать сохраняющуюся перегрузку в течение продолжительного времени. Отдельные испытания в тесте, управляемом двоичным поиском, могут длиться по 60 секунд на каждом шаге, а окончательное испытание может быть ещё более долгим. Это сильно отличается от «нормальных» уровней трафика, но перегрузки не вызывают проблем в изолированной среде. Опасения, высказанные в [RFC6815], состояли в том, что методы [RFC2544] будут распространены на рабочие сети, поэтому авторы призвали сообщество разработчиков стандартов найти показатели и методы, подобные описанным в этом документе.

8.3. Вопросы измерений

В общем случае широко распространённые измерения, описанные в этом документе, столкнутся с широко распространенным поведением. Хорошим примером служит бимодальное поведение пропускной способности IP, уже отмеченное в параграфе 6.6. Отчёт о показателях.

В общем случае **рекомендуется** размещать тестовые точки ближе к предусмотренным для измерения каналам, насколько это возможно практически (имеются ограничения на число тестовых конечных точек с разных точек зрения, например, трафик управления и измерений). Тестирующий оператор **должен** установить значение для параметра MaxHops на основе ожидаемой длины пути. Этот параметр будет препятствовать сильному отклонению измерительного трафика от предусмотренного пути.

Измеряемый путь может иметь состояние, основанное на множестве факторов и параметра «время суток» (Time of day) для начала теста может быть недостаточно. Для повторяющегося тестирования может потребоваться знать время от начала измерительного потока и устройство самого потока, включая объём уже переданного трафика, при котором наблюдалась смена состояния, по времени, числу переданных байтов или их комбинации. Пакеты нагрузки и сообщения обратной связи **должны** включать порядковые номера, это поможет измерениям.

Как отмечено в [RFC7312], может встречаться множество типов формовщиков трафика (shaper) и технологий доступа к связи по запросам, оказывающих существенное влияние на измерения. Методы **должны** быть готовы обеспечить передачу короткой преамбулы для активации доступа к связи по запросу и исключить преамбулу из результатов теста.

Во время измерений могут возникать условия, вызывающие потери пакетов независимо от тестового потока.

1. Перегрузка соединительного или магистрального интерфейса может проявляться как распределенные по времени потери пакетов в тестовом потоке из-за гораздо более скоростных интерфейсов на магистрали.
2. Потеря потоков из-за применения случайного упреждающего обнаружения (Random Early Detection или RED) или иного активного управления очередями может (но не обязательно) влиять на измерительный поток при наличии одновременного конкурирующего трафика (другие потоки).
3. Могут быть лишь незначительные вариации задержки, независимые от скорости передачи.
4. Постоянный конкурирующий трафик на пути измерения, включающем общую среду передачи, может вызывать случайные потери пакетов в тестовом потоке.

Можно смягчить влияние этих условий, используя гибкость настройки тестовой нагрузки, описанной в параграфе 8.1.

Если продолжительность блока измерительных пакетов близка или меньше продолжительности блока (burst) формовщика или ограничителя (policer) на пути, может измеряться скорость линии, а не вносимый формовщиком или ограничителем предел пропускной способности. При подозрении на такие условия **следует** применять другие конфигурации.

Как правило, результаты будут зависеть от характеристик передаваемого потока, измерительное сообщество давно знает об этом и должно учитывать в первую очередь. Хотя по умолчанию для тестирования применяется 1 поток (F=1), применение нескольких потоков может давать преимущества по указанным ниже причинам.

1. Тестовые хосты могут создавать более высокую нагрузку, нежели с одним потоком, или могут применяться параллельные тестовые хосты для создания по одному потоку от каждого.
2. Может присутствовать агрегирование потоков (балансировка на основе потоков) и потребуется несколько потоков, чтобы занаять агрегат.
3. Правила доступа в Internet могут ограничивать пропускную способность уровня Ipв зависимости от типа пакетов Type-P, возможно резервируя пропускную способность для разных типов потоков.

Каждый поток будет контролироваться своей реализацией алгоритма настройки нагрузки (поиск).

Очевидно, что не имеет смысла запускать одновременно более одного теста (независимо от числа потоков в них) для измерения максимальной пропускной способности на одном пути. **Нужно** ограничивать число одновременных независимых тестов одним.

Тесты механизмов перехода с IPv4 на IPv6 могут быть основанием для измерения максимальной пропускной способности. Если переданные и принятые пакеты IPv4 и IPv6 имеют стандартный формат, это следует разрешать (изменение размера заголовков легко учитывается на уровне пакета).

Следует ожидать развития методов по мере продолжения тестов. В ITU-T опубликовано дополнение (Supplement 60) to к рекомендациям серии Y Interpreting ITU-T Y.1540 maximum IP-layer capacity measurements [Y.Sup60], являющееся результатом продолжающегося тестирования показателя. Эти результаты улучшаю описанные здесь методы.

9. Форматы отчётов

Результаты одиночных измерений пропускной способности уровня IP **следует** сопровождать контекстом измерения:

- временная метка (особенно для dtn с максимумом);
- источник и получатель (адрес IP или иной значимый идентификатор);
- другие внутренние параметры теста (4. Общие параметры и определения);
- иные параметры, такие как «тест в движении» или другие факторы, связанные с контекстом измерения;
- достоверность результата (с указанием случаев прерывания теста или отказов при попытке измерения);
- поле для указания необычных обстоятельств и поле для целей, игнорируемых (маскируемых) при последующей обработке.

Результаты измерения максимальной пропускной способности уровня IP **следует** представлять в табличном формате. В таблицу **следует** включать столбцы, указывающие фазу теста и число потоков в этой фазе. В остальных столбцах **следует** указывать агрегат всех потоков, включая Maximum IP-Layer Capacity, Loss Ratio, минимум и максимум RTT, а также другие показатели, имеющие аналогичную значимость.

Как отмечено в параграфе 6.6, бимодальные (или мультимодальные) максимумы **нужно** указывать раздельно для каждого режима.

Таблица 2. Результаты измерения максимальной пропускной способности на уровне IP.

Фаза	***Число потоков***	***Максимальная пропускная способность на уровне IP (Мбит/сек)***	***Частота потерь***	***RTT min (мсек)***	***RTT max (мсек)***
Поиск	1	967,31	0,0002	30	58
Проверка	1	966,00	0,0000	30	38

Статические и конфигурационные параметры

Результаты измерения максимальной пропускной способности уровня IP **должно** сопровождать время субинтервала dt, а также другие параметры из раздела 4. Общие параметры и определения.

Результаты измерения максимальной пропускной способности уровня IP **должны** сопровождать показатели списка PM, соответствующие субинтервалу, где наблюдалась максимальная пропускная способность для каждой фазы теста.

Скорость передачи на уровне IP у отправителя **следует** указывать в табличном формате. В таблицу **следует** включать столбцы для фазы теста и каждого индивидуального (пронумерованного) потока или агрегата потоков в этой фазе. В соответствующем столбце **следует** указывать конкретный субинтервал для скорости передачи stn каждого потока и агрегата. В заключительном столбце **следует** указывать результаты IP-Layer Sender Bit Rate для каждого использованного потока или агрегата всех потоков.

Таблица 3. Результаты измерения скорости отправителя на уровне IP (2 потока и st = 0,05 сек.).

Фаза	***Число потоков или агрегат***	***stn (сек)***	***Скорость передачи отправителем Rate (Мбит/с)***
Поиск	1	0,00	345
Поиск	2	0,00	289
Поиск	Агрегат	0,00	634
Поиск	1	0,05	499
Поиск	...	0,05	...

Статические и конфигурационные параметры

Результаты измерения скорости на уровне IP у отправителя **должна** сопровождать длительность интервала st.

Должны также указываться значения остальных параметров из раздела 4. Общие параметры и определения.

9.1. Форматы данных конфигурации и отчётов

В рамках гармонизации этого показателя и метода измерений с другими SDO форум по широкополосной связи (Broadband Forum или BBF) поделился опытом задания информационной модели и модели данных для конфигурации и отчётов. Эти модели согласуются с параметрами показателя и принятыми по умолчанию значениями, заданными в этом документе. В [TR-471] представлена информационная модель, которая применялась при подготовке полной модели данных в соответствующей работе BBF. В BBF также внимательно рассмотрены вопросы, входящие в компетенцию форума, такие как размещение измерительных систем в архитектуре доступа в Internet. Например, требования к дискретности временных меток, влияющие на выбор протокола тестирования, приведены в таблице 2 [TR-471].

10. Вопросы безопасности

С активными показателями и измерениями связана долгая история вопросов безопасности. К этому документу применимы соображения безопасности для активных измерений на рабочих путях, например, из [RFC4656] и [RFC5357].

При рассмотрении приватности участников измерений или поставщиков измерительного трафика сведения, доступные возможным наблюдателям на пути, существенно сокращаются при использовании активных методов, выходящих за рамки это работы. С пассивными наблюдениями пользовательского трафика для измерительных целей связано множество проблем безопасности. Читателям рекомендуется обратиться к соображениям безопасности, рассмотренным в Large-scale Measurement of Broadband Performance (LMAP) Framework [RFC7594], где охвачены активные и пассивные методы.

Имеются некоторые новые соображения, связанные с описанными здесь измерениями пропускной способности.

1. **Требуется** взаимодействие хостов отправителя и получателя и согласие на тестирование пути между этими хостами. Хосты играют роль Src или Dst.
2. **Требуется** наличие инициируемого пользователем установочного согласования между взаимодействующими хостами, которое позволяет межсетевым экранам контролировать трафик, идущий на порт управления (ожидаемый и аутентифицированный) или эфемерные порты, создаваемые по мере необходимости. Межсетевые экраны, защищающие каждый из хостов, могут продолжать обычную работу.
3. **Рекомендуется** аутентификация (клиент-сервер) и защита целостности для сообщений обратной связи при измерениях.
4. Хосты **должны** ограничивать число одновременных тестов во избежание исчерпания ресурсов и получения неточных результатов.
5. Скорость отправки **должна** ограничиваться. Это можно сделать с использованием заранее подготовленной таблицы предлагаемой нагрузки (8.1. Алгоритм настройки нагрузки). Рекомендуемый алгоритм поиска ведёт к росту скорости от минимальной в соответствии с таблицей.
6. Подписчики услуг с ограниченным объёмом данных, выполняющие обширное тестирование, могут столкнуться с влиянием средств контроля трафика у поставщика услуг. Тестирование с использованием измерительных хостов сервис-провайдера **следует** ограничивать по частоте и/или суммарному объёму тестового трафика (например, **следует** отказаться от длительного тестирования с большим значением I).

Точная спецификация этих свойств оставлена на будущее.

11. Взаимодействие с IANA

Документ не требует действий IANA.

12. Литература

12.1. Нормативные документы

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2330] Paxson, V., Almes, G., Mahdavi, J., and M. Mathis, "Framework for IP Performance Metrics", RFC 2330, DOI 10.17487/RFC2330, May 1998, <https://www.rfc-editor.org/info/rfc2330>.

[RFC2681] Almes, G., Kalidindi, S., and M. Zekauskas, "A Round-trip Delay Metric for IPPM", RFC 2681, DOI 10.17487/RFC2681, September 1999, <https://www.rfc-editor.org/info/rfc2681>.

[RFC4656] Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and M. Zekauskas, "A One-way Active Measurement Protocol (OWAMP)", RFC 4656, DOI 10.17487/RFC4656, September 2006, <https://www.rfc-editor.org/info/rfc4656>.

[RFC4737] Morton, A., Ciavattone, L., Ramachandran, G., Shalunov, S., and J. Perser, "Packet Reordering Metrics", RFC 4737, DOI 10.17487/RFC4737, November 2006, <https://www.rfc-editor.org/info/rfc4737>.

[RFC5357] Hedayat, K., Krzanowski, R., Morton, A., Yum, K., and J. Babiarz, "A Two-Way Active Measurement Protocol (TWAMP)", RFC 5357, DOI 10.17487/RFC5357, October 2008, <https://www.rfc-editor.org/info/rfc5357>.

[RFC6438] Carpenter, B. and S. Amante, "Using the IPv6 Flow Label for Equal Cost Multipath Routing and Link Aggregation in Tunnels", RFC 6438, DOI 10.17487/RFC6438, November 2011, <https://www.rfc-editor.org/info/rfc6438>.

[RFC7497] Morton, A., "Rate Measurement Test Protocol Problem Statement and Requirements", RFC 7497, DOI 10.17487/RFC7497, April 2015, <https://www.rfc-editor.org/info/rfc7497>.

[RFC7680] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Loss Metric for IP Performance Metrics (IPPM)", STD 82, RFC 7680, DOI 10.17487/RFC7680, January 2016, <https://www.rfc-editor.org/info/rfc7680>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8468] Morton, A., Fabini, J., Elkins, N., Ackermann, M., and V. Hegde, "IPv4, IPv6, and IPv4-IPv6 Coexistence: Updates for the IP Performance Metrics (IPPM) Framework", RFC 8468, DOI 10.17487/RFC8468, November 2018, <https://www.rfc-editor.org/info/rfc8468>.

12.2. Дополнительная литература

[copycat] Edeline, K., Kühlewind, M., Trammell, B., and B. Donnet, "copycat: Testing Differential Treatment of New Transport Protocols in the Wild", ANRW '17, DOI 10.1145/3106328.3106330, July 2017, <https://irtf.org/anrw/2017/anrw17-final5.pdf>.

[LS-SG12-A] "Liaison statement: LS - Harmonization of IP Capacity and Latency Parameters: Revision of Draft Rec. Y.1540 on IP packet transfer performance parameters and New Annex A with Lab Evaluation Plan", From ITU-T SG 12, March 2019, <https://datatracker.ietf.org/liaison/1632/>.

[LS-SG12-B] "Liaison statement: LS on harmonization of IP Capacity and Latency Parameters: Consent of Draft Rec. Y.1540 on IP packet transfer performance parameters and New Annex A with Lab & Field Evaluation Plans", From ITU-T-SG-12, May 2019, <https://datatracker.ietf.org/liaison/1645/>.

[RFC2544] Bradner, S. and J. McQuaid, "Benchmarking Methodology for Network Interconnect Devices", RFC 2544, DOI 10.17487/RFC2544, March 1999, <https://www.rfc-editor.org/info/rfc2544>.

[RFC3148] Mathis, M. and M. Allman, "A Framework for Defining Empirical Bulk Transfer Capacity Metrics", RFC 3148, DOI 10.17487/RFC3148, July 2001, <https://www.rfc-editor.org/info/rfc3148>.

[RFC5136] Chimento, P. and J. Ishac, "Defining Network Capacity", RFC 5136, DOI 10.17487/RFC5136, February 2008, <https://www.rfc-editor.org/info/rfc5136>.

[RFC6815] Bradner, S., Dubray, K., McQuaid, J., and A. Morton, "Applicability Statement for RFC 2544: Use on Production Networks Considered Harmful", RFC 6815, DOI 10.17487/RFC6815, November 2012, <https://www.rfc-editor.org/info/rfc6815>.

[RFC7312] Fabini, J. and A. Morton, "Advanced Stream and Sampling Framework for IP Performance Metrics (IPPM)", RFC 7312, DOI 10.17487/RFC7312, August 2014, <https://www.rfc-editor.org/info/rfc7312>.

[RFC7594] Eardley, P., Morton, A., Bagnulo, M., Burbridge, T., Aitken, P., and A. Akhter, "A Framework for Large-Scale Measurement of Broadband Performance (LMAP)", RFC 7594, DOI 10.17487/RFC7594, September 2015, <https://www.rfc-editor.org/info/rfc7594>.

[RFC7799] Morton, A., "Active and Passive Metrics and Methods (with Hybrid Types In-Between)", RFC 7799, DOI 10.17487/RFC7799, May 2016, <https://www.rfc-editor.org/info/rfc7799>.

[RFC8085] Eggert, L., Fairhurst, G., and G. Shepherd, "UDP Usage Guidelines", BCP 145, RFC 8085, DOI 10.17487/RFC8085, March 2017, <https://www.rfc-editor.org/info/rfc8085>.

[RFC8337] Mathis, M. and A. Morton, "Model-Based Metrics for Bulk Transport Capacity", RFC 8337, DOI 10.17487/RFC8337, March 2018, <https://www.rfc-editor.org/info/rfc8337>.

[TR-471] Morton, A., "Maximum IP-Layer Capacity Metric, Related Metrics, and Measurements", Broadband Forum TR-471, July 2020, <https://www.broadband-forum.org/technical/download/TR-471.pdf>.

[Y.1540] ITU-T, "Internet protocol data communication service — IP packet transfer and availability performance parameters", ITU-T Recommendation Y.1540, December 2019, <https://www.itu.int/rec/T-REC-Y.1540-201912-I/en>.

[Y.Sup60] ITU-T, "Interpreting ITU-T Y.1540 maximum IP-layer capacity measurements", ITU-T Recommendation Y.Sup60, October 2021, <https://www.itu.int/rec/T-REC-Y.Sup60/en>.

Приложение A. Псевдокод алгоритма настройки нагрузки

В этом приложении представлен псевдокод реализации алгоритма, описанного в параграфе 8.1.

```
Rx = 0              # Текущая скорость передачи (строка таблицы)

seqErr = 0          # Измеренное значение, учитывающее потери,
                    # нарушение, дубликаты (проявляются как аномалии
                    # в порядковых номерах)

seqErrThresh = 10   # Порог счётчика seqErr учитывающего потери,
                    # нарушение порядка и дублирование (все выглядят
                    # как ошибки в порядковых номерах)

delay = 0           # Измеренное время круговой задержки (RTD), мсек

lowThresh = 30      # Нижний порог диапазона RTD, мсек

upperThresh = 90    # Верхний порог диапазона RTD, мсек

hSpeedThresh = 1    # Порог для смены скорости передачи
                    # (такой как 1 Мбит/с и 100 Мбит/с), Гбит/с

slowAdjCount = 0    # Измеренное число последовательных отчётов,
                    # указывающих потери и/или задержку выше
                    # upperThresh

slowAdjThresh = 3   # Порог slowAdjCount для фиксации перегрузки
                    # Следует устанавливать > 1 во избежание
                    # ложных временных потерь.

highSpeedDelta = 10 # Число строк таблицы для перемещения за 1 шаг
                    # быстрой корректировки нагрузки

maxLoadRates = 2000 # Максимальный индекс таблицы (число строк)

if ( seqErr <= seqErrThresh && delay < lowThresh ) {
        if ( Rx < hSpeedThresh && slowAdjCount < slowAdjThresh ) {
                        Rx += highSpeedDelta;
                        slowAdjCount = 0;
        } else {
                        if ( Rx < maxLoadRates - 1 )
                                        Rx++;
        }
} else if ( seqErr > seqErrThresh || delay > upperThresh ) {
        slowAdjCount++;
        if ( Rx < hSpeedThresh && slowAdjCount == slowAdjThresh ) {
                        if ( Rx > highSpeedDelta * 3 )
                                        Rx -= highSpeedDelta * 3;
                        else
                                        Rx = 0;
        } else {
                        if ( Rx > 0 )
                                        Rx--;
        }
}
```

Приложение B. Проверка рекомендаций RFC 8085

Параграф 3.1 [RFC8085] (BCP 145), где даны рекомендации по использованию UDP, сосредоточен на контроле перегрузок и задаёт требования уровня **должно** (MUST) и **следует** (SHOULD).

B.1. Оценка обязательных требований

Требование раздела 3 в [RFC8085] гласит:

> Характеристики путей Internet могут меняться в широких пределах ... Поэтому приложениям, где может использоваться Internet, **недопустимо** принимать какие либо допущения о характеристиках конкретного пути. Вместо этого **должны** применяться механизмы, которые позволят безопасно работать в сильно различающихся условиях. Обычно это требует осторожной проверки текущего состояния пути Internet, через который будет организована связь, для достижения устраивающего поведения, которое будет достаточно беспристрастным к другому трафику на том же пути.

Целью алгоритма регулировки нагрузки, описанного в параграфе 8.1, является зондирование сети и обеспечение измерений максимальной пропускной способности уровня IP с минимальными допущениями о тестируемом пути в пределах применения, описанного в разделе 2. Имеется противоречие между целями измерения и минимизацией измерительного трафика (особенно при гигабитных скоростях), а также продолжительности теста (один из факторов, влияющих на беспристрастность алгоритма).

Далее в разделе 3 [RFC8085] приводятся варианты выполнения обязательных требований, но ни один из них не подходит для описанного здесь показателя и метода. Фактически специализированные методы на основе TCP не позволяют получить точности измерения, показанной в сравнительных тестах с работающим кодом [LS-SG12-A] [LS-

SG12-B] [Y.Sup60]. UDP в этих методах применяется в основном для поддержки современных методов передачи в Internet, где требуется транспортный протокол [corycat], показатели основаны на уровне IP, а UDP допускает простое сопоставление с уровнем IP.

В параграфе 3.1.1 [RFC8085] приведены требования к таймеру протокола.

> Выборку задержек **недопустимо** делать из неоднозначных транзакций. Каноническим примером служит протокол, который повторно передаёт данные, но затем не может определить, какая из копий была подтверждена.

Пакеты нагрузки и обратной связи (состояние) **должны** включать порядковые номера, это поможет при измерениях и не потребуется повторной передачи.

> Когда оценка задержки служит для установки таймера, обеспечивающего обнаружение потерь (с повтором или без него) завершение отсчёта таймера **должно** считаться индикацией перегрузки в сети, вызывающим приспосабливать скорость передачи к безопасному умеренному значению ...

Описанные здесь методы используют таймеры для отката скорости передачи при потере сообщений о состоянии (тайм-аут Lost Status Backoff timeout) и прекращения теста при длительной потере связности (тайм-ауты для пакетов обратной связи или нагрузки).

Этот документ не отмечает каких-либо конкретных преимуществ применения явных уведомлений о перегрузке (ECN).

В параграфе 3.2 [RFC8085] обсуждается размер сообщений.

> Для определения подходящего размера данных UDP приложения **должны** вычесть размер заголовка IP (вместе с необязательными заголовками IPv4 или заголовками расширения IPv6), а также размер заголовка UDP (8 байтов) из значения PMTU.

В описанном методе применяется таблица скоростей с максимальным размером данных (payload) UDP, который предполагает существенные издержки на заголовок и позволяет избежать фрагментации.

В параграфе 3.3 [RFC8085] приведены рекомендации по надёжности.

> Приложения, требующие гарантированной доставки, **должны** сами реализовать подходящий механизм.

Показатели и метод измерения пропускной способности уровня IP не требуют гарантированной доставки.

> Приложения, которым нужна упорядоченная доставка, **должны** самостоятельно упорядочивать дейтаграммы.

Показатели и метод измерения пропускной способности уровня IP не требуют восстановления порядка пакетов, предпочтительно измерять переупорядочение, если оно наблюдается [RFC4737].

B.2. Оценка рекомендаций

Целью алгоритма настройки нагрузки является измерение максимальной пропускной способности уровня IP в контексте нечастых и краткосрочных диагностических тестов. Эта цель является глобальным исключением из многих требования уровня **следует** (SHOULD) в [RFC8085], большинство которых предназначены для долгосрочных потоков, которые должны сосуществовать с другим трафиком более или менее беспристрастно. Однако алгоритм (как указано в параграфе 8.1 и Приложении A) реагирует на перегрузку чётко заданным способом.

Рассмотрим это на примере конкретной рекомендации из параграфа 3.1.5 [RFC8085] (относительно влияния измерений RTT и потерь на контроль перегрузок).

> Контролю перегрузок [algorithm], разработанному для UDP, **следует** как можно быстрее реагировать на индикацию перегрузки, а также **следует** учитывать при выборе новой скорости частоту потерь и время отклика.

Алгоритм настройки нагрузки реагирует на измерение потерь и RTT чётким и кратковременным снижением скорости, когда это оправдано, а отклик использует прямые измерения (более точные, чем можно вывести из TCP ACK).

В параграфе 3.1.5 [RFC8085] сказано:

> Реализованной схеме контроля перегрузок **следует** обеспечивать использование пропускной способности (емкости), сравнимое по порядку величины с TCP, чтобы не истощать другие потоки на том же пути с узким местом.

Это требование для сосуществующих потоков, а не для диагностики и нечастых кратковременных измерений. Колебания скорости во время коротких тестов позволяют проходить другим пакетам и не задерживать другие потоки.

По иронии судьбы, специализированные измерения на основе TCP «скорости Internet» также предназначены для обхода этого требования (уровня **следует**) путём запуска множества потоков (например, 9) для увеличения объёма тестовых данных.

Алгоритм настройки нагрузки не может стать механизмом контроля перегрузок в стиле TCP, поскольку у него будут те же недостатки, что и у TCP при попытке измерить максимальную пропускную способность уровня IP, и он не достигнет цели. Результаты упомянутого тестирования [LS-SG12-A] [LS-SG12-B] [Y.Sup60] подтверждают это многократно при сравнении с измерениями по множеству соединений на основе TCP.

Краткий обзор требований [RFC8085] приведён в таблице 4 («Да» в левом столбце указывает совместимость этого документа с требованием, «-» - неприменимость требования).

Таблица 4. Сводка основных рекомендаций RFC 8085.

	Рекомендации RFC 8085	***Параграф***
Да	**Должна** поддерживаться работа по широкому диапазону путей Internet.	3
-	**Следует** применять полнофункциональный транспорт (например, TCP)	
Да	**Следует** контролировать скорость передачи	3.1
-	**Следует** контролировать перегрузки для всего трафика	
	Для передачи больших объёмов данных	3.1.2
-	**следует** рассмотреть реализацию TFRC,	

-	в противном случае **следует** иным способом использовать пропускную способность по аналогии с TCP Для передачи небольших объемов данных	3.1.3
-	**Следует** измерять RTT и передавать не более 1 дейтаграммы за интервал RTT	3.1.1
-	иначе **следует** передавать не более 1 дейтаграммы за 3 секунды.	
-	**Следует** восстанавливать (back-off) таймеры повтора после потери пакетов	
Да	**Следует** поддерживать механизмы для контроля пиков передачи	3.1.6
-	**Можно** реализовать ECN, при этом будет требоваться определённый набор механизмов в приложении, если применяется ECN.	3.1.7
Да	Для DiffServ **не следует** опираться на реализацию PHB	3.1.8
Да	Для путей с поддержкой QoS **можно** отказаться от использования CC	3.1.9
Да	**Не следует** опираться лишь на QoS для пропускной способности	3.1.10
-	Для потоков без управления **следует** реализовать транспортные выключатели	
Да	**Можно** реализовать транспортные выключатели для других приложений	
	Для туннелей с трафиком IP	3.1.11
-	**не следует** применять контроль перегрузок	
-	**должно** корректно обрабатываться поле IP ECN	
	Для туннелей не-IP или не задаваемой трафиком скорости	3.1.11
-	**следует** выполнять CC или применять выключатель	
-	**следует** ограничивать типы трафика, доставляемого через туннель	
Да	**Не следует** передавать дейтаграммы размером больше PMTU	3.2
Да	т. е. **следует** определять PMTU или передавать дейтаграммы меньше минимального PMTU	
-	При использовании PLPMTUD **требуются** конкретные механизмы приложения	
Да	**Следует** обрабатывать потерю, дублирование и нарушение порядка дейтаграмм	3.3
-	**Следует** обеспечивать устойчивость к задержкам доставки до 2 минут	
Да	**Следует** включать контрольную сумму UDP для IPv4	3.4
Да	**Следует** включать контрольную сумму UDP для IPv6; **требуются** конкретные механизмы для нулевой контрольной суммы UDP	3.4.1
-	**Следует** поддерживать механизм защиты от атак извне пути	5.1
-	**Не следует** всегда передавать сообщения keep-alive для промежуточных устройств	3.5
-	При необходимости **можно** использовать keep-alive с интервалом не менее 15 секунд	
Да	Приложениям с ограниченным применением (или контролируемой средой) **следует** указывать эквивалентные механизмы и описывать их применение	3.6
-	Групповым приложениям с большим трафиком **следует** реализовать контроль перегрузок	4.1.1
-	Групповым приложениям с небольшим трафиком **следует** реализовать контроль перегрузок	4.1.2
-	Групповым приложениям **следует** применять безопасное значение PMTU	4.2
Да	**Следует** избегать применения множества портов	5.1.2
Да	**Должны** проверяться адреса отправителей в принятых пакетах	
-	**Следует** проверять данные в сообщениях ICMP	5.2
Да	**Следует** использовать случайный выходной порт или эквивалентный метод, а для приложений «клиент-сервер» следует передавать отклики с адреса, соответствующего запросу	6
-	При необходимости **следует** использовать стандартные протоколы защиты IETF	6

Благодарности

Спасибо Joachim Fabini, Matt Mathis, J. Ignacio Alvarez-Hamelin, Wolfgang Balzer, Frank Brockners, Greg Mirsky, Martin Duke, Murray Kucherawy, Benjamin Kaduk за комментарии к этому документу и смежным темам. Спасибо Magnus Westerlund, Lars Eggert, Zaheduzzaman Sarker за второй раунд рецензирования.

Адреса авторов

Al Morton
AT&T Labs
200 Laurel Avenue South
Middletown, NJ 07748
United States of America
Phone: +1 732 420 1571
Email: acm@research.att.com

Rüdiger Geib
Deutsche Telekom
Heinrich Hertz Str. 3-7
64295 Darmstadt
Germany
Phone: +49 6151 5812747
Email: Ruediger.Geib@telekom.de

Len Ciavattone
AT&T Labs
200 Laurel Avenue South
Middletown, NJ 07748
United States of America
Phone: +1 732 420 1239
Email: lencia@att.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru