Internet Engineering Task Force (IETF)
Request for Comments: 9522
Obsoletes: 3272
Category: Informational
ISSN: 2070-1721

A. Farrel, Ed.
Old Dog Consulting

January 2024

Overview and Principles of Internet Traffic Engineering

Обзор и принципы организации трафика Internet

Аннотация

В этом документе описываются принципы организации трафика (traffic engineering или TE) в сети Internet. Документ призван способствовать лучшему пониманию вопросов, связанных с организацией трафика в сетях IP и сетях, поддерживающих их, а также обеспечить основу для развития средств организации трафика в Internet. Обсуждаются также принципы, архитектура и методология для оценки и оптимизации производительности действующих сетей.

Эта работа была впервые опубликована как RFC 3272 в мае 2002 г. Документ заменяет RFC 3272, полностью обновляя текст для приведения в соответствие с передовым опытом организации трафика Internet и включения ссылок на соответствующие свежие работы IETF.

Статус документа

Документ не относится к категории Internet Standards Track и публикуется для информации.

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 документах BCP можно найти в разделе 2 в RFC 7841.

Информацию о текущем статусе документа, ошибках и способах обратной связи можно найти по ссылке https://www.rfc-editor.org/info/rfc9522.

Оглавление

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

1. Введение

В этом документе описываются принципы Internet TE. Целью документа является формулировка общих вопросов и принципов Internet TE, а также предоставление при необходимости руководств, рекомендаций и вариантов для разработки предварительно планируемых (offline) и динамических (online) возможностей Internet TE и систем поддержки.

Несмотря на максимальную эффективность Internet TE при сквозном применении, в этом документе основное внимание уделено TE в конкретном домене, таком как автономная система (Autonomous System или AS). Большая часть трафика Internet, как правило, исходит из одной AS и завершается в другой, поэтому в документ включён обзор аспектов междоменной организации трафика.

В документ включена терминология и таксономия для описания и понимания базовых концепций Internet TE.

Эта работа была впервые опубликована как RFC 3272 в мае 2002 г. Документ заменяет RFC 3272, полностью обновляя текст для приведения в соответствие с передовым опытом организации трафика Internet и включения ссылок на соответствующие свежие работы IETF. Следует отметить, что примерно 3/5 RFC, упомянутых в этом документе, выпущены после публикации [RFC3272]. В Приложении A приведена сводка отличий этого документа от [RFC3272].

1.1. Что такое организация трафика Internet?

Одной из наиболее важных функций, выполняемых в Internet, является маршрутизация и пересылка трафика от входных узлов к выходным. Поэтому одной из наиболее характерных функций организации трафика Internet является управление и оптимизаций функций маршрутизации и пересылки для направления трафика через сеть. Организация трафика Internet определяется как аспект устройства сети Internet, связанный с задачами оценки и оптимизации производительности действующих сетей IP. Организация трафика включает применение технологических и научных принципов для измерения, определения характеристик, моделирования и управления трафиком Internet [RFC2702] [AWD2].

Важнейшее значение имеет именно производительность сети, воспринимаемая конечными пользователями. Видимые пользователям характеристики сети - это появляющиеся (emergent) в сети свойства, которые характеризуют сеть в целом. Поэтому основной задачей поставщиков услуг (сервис-провайдеров) является улучшение таких свойств с учётом экономических соображений. Это достигается путём выполнения ориентированных на трафик требований к производительности при надёжном использовании сетевых ресурсов без чрезмерных потерь. К ориентированным на трафик показателям производительности относятся задержки и их вариации, доля теряемых пакетов и пропускная способность.

Internet TE реагирует на события в сети (отказы каналов или узлов, сообщённая или спрогнозированная перегрузка, плановое обслуживание, снижение качества сервиса, плановые изменения картины трафика и т. п.). Управление пропускной способностью реагирует на события в интервалах от нескольких дней до нескольких лет. Функции управления маршрутизацией работают в интервалах от миллисекунд до дней. Функции обработки на уровне пакетов работают с очень высоким временным разрешением (до миллисекунд), реагируя на статистические показатели поведения трафика в реальном масштабе времени.

Таким образом, аспекты оптимизации TE можно рассматривать с точки зрения управления, которое может быть реактивным или упреждающим. В проактивном случае система управления TE выполняет упреждающие действия для защиты от прогнозируемых нежелательных состояний сети, например, организуя резервные пути. Могут также выполняться действия, обеспечивающие более желательное будущее состояние сети. В реактивном варианте система управления исправляет возникшие проблемы и адаптируется к событиям в сети, таким как изменение маршрутизации в результате отказа.

Другой важной задачей Internet TE является содействие надёжной работе сети [RFC2702]. Надёжность работы может быть обеспечена путём предоставления механизмов, улучшающих целостность сети, и внедрения правил для повышения живучести сети. Это снижает уязвимость служб при сбоях, возникающих из-за ошибок, отказов и неисправностей в инфраструктуре сети.

Аспекты оптимизации TE могут быть реализованы за счёт управления пропускной способностью и трафиком. В этом документе управление пропускной способностью включает в себя планирование, а также управление маршрутизацией и ресурсами. Особый интерес представляю такие сетевые ресурсы, как пропускная способность каналов, ёмкость буферов и вычислительные ресурсы. В этом документе управление трафиком включает:

1. функции управления трафиком на узлах, такие как кондиционирование, управление очередями, планирование;
2. другие функции, регулирующие потоки трафика в сети или осуществляющие арбитраж доступа к сетевым ресурсам для разных потоков трафика или пакетов.

Одной из основных задач Internet TE является реализация средств автоматизированного управления, которые быстро и экономически эффективно адаптируются к значительным изменениям состояния сети, сохраняя её стабильность. Оценка производительности позволяет оценить эффективность методов TE, а результаты такой оценки могут применяться для выявления имеющихся проблем, направления повторной оптимизации сети и прогнозирования возможных проблем. Однако этот процесс может быть долгим и не всегда подходит для реагирования на краткосрочные изменения в сети.

Оценку производительности можно выполнить разными способами. Наиболее примечательными методами включают аналитику, моделирование и эмпирические методы на основе измерений.

Организация трафика бывает двух видов:

- фоновый процесс, постоянно отслеживающий трафик и условия в сети и оптимизирующий выделение ресурсов для повышения производительности;
- заранее спланированное распределение трафика, которое считается оптимальным.

В последнем случае любые отклонения от оптимального распределения (например, вызванные обрывом волокна) устраняются после ремонта без дополнительной оптимизации. Однако эта форма TE полагается на представление о том, что планируемое состояние сети оптимально, поэтому в таком режиме имеется два уровня TE:

- задача планирования TE для обеспечения оптимального распределения трафика;
- задачи маршрутизации и пересылки, поддерживающие потоки трафика в соответствии с запланированным распределением.

Как правило, концепции и механизмы TE должны быть достаточно конкретными и чётко определёнными для выполнения известных требований, а также достаточно гибкими и расширяемыми для адаптации к непредвиденным запросам в будущем (см. параграф 6.1).

1.2. Компоненты TE

Как указано в параграфе 1.1, организация трафика Internet обеспечивает оптимизацию производительности сетей IP при экономном и надёжном использовании сетевых ресурсов. Такая оптимизация поддерживается на уровне управления/контроллера и в плоскости данных/пересылки.

Ключевыми элементами TE, требуемыми для любого решения, являются:

1. правила (политика);
2. управления путями;
3. управления ресурсами.

Некоторые решения TE полагаются на эти элементы в большей или меньшей степени. Сохраняются споры о том, может ли решение называться TE, если оно не включает все указанные элементы. В рамках этого документа считается, что все решения TE должны включать те или иные аспекты всех этих элементов. Иные решения можно считать неполным ((partial) TE и они тоже рассматриваются в этом документе.

Политика позволяет выбирать пути (включая следующий узел) на основе сведений, выходящих за рамки базовой достижимости. В ранних определениях политики маршрутизации (например, [RFC1102] и [RFC1104]) рассмотрено применение политики маршрутизации для ограничения доступа к сетевым ресурсам на уровне агрегирования. BGP служит примером широко распространённого механизма применения таких правил (см. [RFC4271] и [RFC8955]). В контексте TE решения на основе правил принимаются в плоскости управления или на контроллерах и определяют выбор путей. Примеры этого приведены в [RFC4655] и [RFC5394]. Решения TE могут включать механизмы для распространения и/или применения правил, но определения конкретных правил остаётся за операторами сети.

Управление путями - это возможность пересылать пакеты с использованием дополнительных сведений, а не только информации о следующем узле (next hop). Примеры такого управления включают заданные отправителем маршруты IPv4 [RFC0791], явные маршруты RSVP-TE [RFC3209], маршрутизацию по сегментам (Segment Routing или SR) [RFC8402], цепочки сервисных функций (Service Function Chaining) [RFC7665]. Управления путями для TE может поддерживаться через протоколы плоскости управления, путём кодирования в заголовках плоскости данных или с использованием сочетания этих методов. Это включает управление с помощью контроллера с использованием ориентированного на сеть протокола управления.

Управление ресурсами обеспечивает пересылку и управление с учётом ресурсов. Примерами ресурсов являются пропускная способность, буферы и очереди, которыми можно управлять для контроля потерь и задержек. Резервирование ресурсов является частью управления ресурсами и обеспечивает в масштабе домена согласованное представление о сетевых ресурсах, используемых конкретными потоками. Определение используемых ресурсов может выполняться грубо или на очень тонком уровне. Отметим, что согласованное представление существует на уровне плоскости управления или контроллере, но не в плоскости данных. Представление может включать лишь данные учёта, но обычно в него входят возможности принимать, отвергать или реклассифицировать потоки на основе правил. Учёт может выполняться на основе любого сочетания статических представлений о потребностях в ресурсах и использования динамических механизмов сбора требований (например, через RSVP-TE [RFC3209]) или сведений о доступности ресурсов (например, через расширения OSPF для GMPLS [RFC4203]).

Распределение ресурсов - это аспект управления ресурсами в плоскости данных. Оно обеспечивает выделение конкретным потокам определённых ресурсов узлов и каналов. Примеры ресурсов включают буферы, механизмы применения правил и формирования скорости, которые обычно поддерживаются с помощью очередей. Распределение ресурсов включает также сопоставление потоков (классификацию) с определенным набором выделенных ресурсов. Методы классификации и дискретность управления ресурсами зависят от технологии. Примеры включают дифференцированное обслуживание (Diffserv) с отбрасыванием и перемаркировкой [RFC4594], MPLS-TE [RFC3209], пути с коммутацией по меткам (Label Switched Path или LSP) на основе GMPLS [RFC3945], а также решения на основе контроллеров [RFC8453]. Это уровень управления ресурсами не является обязательным, но важен для сетей, которые хотят поддерживать правила контроля перегрузок для контроля или регулирования предлагаемого трафика с целью предоставления разных уровней обслуживания и смягчения проблем перегрузки, а также для сетей, желающих контролировать задержки определённых потоков трафика.

1.3. Сфера применения

Данный документ предназначен для внутридоменного применения TE, поскольку это является практическим уровнем технологии TE, существующим в настоящее время в сети Internet, т. е. документ описывает TE в рамках данной AS сети Internet. В документе обсуждаются концепции, связанные с управлением трафиком внутри домена, включая управление маршрутизацией, выделение ресурсов на макро- и микроуровне, а также связанные с этим вопросы координации управления.

В документе описываются и классифицируются методы, уже применяемые или разрабатываемые для Internet TE, а также взаимодействие этих методов и варианты, в которых они могут быть полезны.

Хотя основное внимание в этом документе уделяется внутридоменной организации трафика, в разделе 7 приведён высокоуровневый обзор междоменного TE. Междоменная организация трафика Internet очень важна для повышения производительности планетарной инфраструктуры Internet.

Когда это возможно, соответствующие требования из имеющихся документов IETF и других организаций встроены в текст путём ссылок.

1.4. Терминология

В этом параграфе определены термины, полезные для Internet TE. Определения относятся к данному документу и в других документах могут иметь иной смысл.

Busy hour - час высокой загрузки
: Часовой период в определённом интервале времени (обычно 24 часа), когда нагрузка в сети или подсети является наибольшей.

Congestion - перегрузка
: Состояние сетевого ресурса, при котором трафик превышает возможности ресурса в течение определённого интервала времени. Незначительная перегрузка может быть полезна для обеспечения работы сетевых ресурсов на полную мощность, что особенно актуально для устройств на границе сети, где желательно обеспечить максимально возможное обслуживание сетевого трафика. Если разрешать перегрузку внутри сети (например, при длительном превышении входного трафика над выходным), это будет негативно влиять на трафик пользователей.

Congestion avoidance - предотвращение перегрузок
: Подход к контролю перегрузок, при котором предпринимаются попытки избежать возникновения перегрузок. В основном это относится к перегрузкам в сети, хотя может быть частью контроля перегрузок по запросам трафика.

Congestion response - реагирование на перегрузку
: Подход, при котором предпринимаются попытки устранить уже возникшие проблемы, связанные с перегрузкой.

Constraint-based routing - маршрутизация на основе ограничений
: Класс протоколов маршрутизации, в которых для принятия решений о выборе маршрута учитываются конкретные атрибуты трафика, ограничения сети или правил. Такая маршрутизация применима для отдельных потоков и агрегатов потоков. Это является обобщением маршрутизации на основе QoS.

Demand-side congestion management - контроль перегрузок по запросам

Схема контроля перегрузок на основе регулирования или кондиционирования предлагаемой нагрузки.

Effective bandwidth - эффективная пропускная способность

Минимальная пропускная способность, которую нужно предоставить потоку или агрегату потоков для обеспечения им «приемлемого качества обслуживания». Более строгое определение приведено в [KELLY].

Egress node - выходной узел

Устройство (маршрутизатор), через которое трафик выходит из сети в направлении получателя (хост, сервер и т. п.) или другой сети.

End-to-end - сквозной

Этот термин зависит от контекста и часто применяется к потокам трафика от исходного отправителя до конечного получателя. Другой термин - «от границы до границы» (edge-to-edge) часто применяется для описания потоков трафика от входа в домен или сеть до выхода оттуда. Однако в некоторых случаях (например, при наличии сервисного интерфейса между сетью и её клиентом или при прохождении пути через несколько доменов, контролируемых одним процессом) термин «сквозной» применяется для обозначения полной работы службы, которая может включать конкатенацию операций «от границы до границы». Таким образом, в контексте TE термин «сквозной» может относиться к полному пути TE, но не к полному пути трафика от приложения-источника до конечного получателя.

Hotspot - «горячая точка»

Элемент сети или подсистема, со значительно большим уровнем перегрузки, чем в других местах.

Ingress node - входной узел

Устройство (маршрутизатор), через которое трафик входит в сеть от источника (хост) или из другой сети.

Metric - показатель

Параметр, определяемый в стандартных единицах измерения.

Measurement methodology - методология измерений

Повторяемый метод измерения, служащий для определения одного или нескольких интересующих показателей.

Network congestion - перегрузка сети

Перегрузка на неком узле или канале сети, которая столь велика, что вызывает неприемлемую задержку в очередях или потерю пакетов. Перегрузка сети может негативно влиять на сквозной трафик или трафик edge-to-edge, поэтому могут внедряться схемы TE для балансировки трафика в сети и предотвращения перегрузки.

Network survivability - жизнеспособность сети

Способность обеспечивать предписанный уровень QoS для имеющихся служб при заданном числе отказов в сети.

Offered load - предлагаемая нагрузка

Предлагаемая нагрузка или предлагаемый нагрузочный трафик (offered traffic load) - это количественная мера объёма трафика, представляемого для передачи через сеть по сравнению с передаточными возможностями этой сети. Этот термин заимствован из теории очередей и значение 1 для предлагаемой нагрузки указывает, что сеть может передавать весь предложенный трафик, но не более того.

Offline traffic engineering - автономная организация трафика

Система организации трафика, существующая вне сети.

Online traffic engineering - внутренняя организация трафика

Система организации трафика, существующая в сети, которая обычно реализуется на действующих элементах сети или дополняет их.

Performance measures - показатели производительности

Показатель, обеспечивающий количественную или качественную оценку интересующих систем или подсистем.

Performance metric - показатель производительности

Параметр производительности, указанный в стандартных единицах измерения.

Provisioning - предоставление

Процесс выделения или настройки сетевых ресурсов в соответствии с неким запросом.

Quality of Service (QoS) - качество обслуживания

QoS [RFC3198] - это механизмы, применяемые в сети для достижения определённых целей по доставке трафика конкретной службы в соответствии с параметрами, заданными соглашением об уровне обслуживания (SLA). «Качество» характеризуется доступностью услуги, задержками и их вариациями, пропускной способностью и долей теряемых пакетов. На уровне сетевого ресурса QoS относится к набору возможностей, позволяющих сервис-провайдеру приоритизировать трафик и управлять пропускной способностью и задержкой в сети.

QoS routing - маршрутизация QoS

Класс систем маршрутизации, выбирающих пути для потоков на основе требований к QoS для потока.

Service Level Agreement (SLA) - соглашение об уровне обслуживания

Соглашение между поставщиком услуг и клиентом, гарантирующее конкретные уровни производительности и надёжности при определённой стоимости.

Service Level Objective (SLO) - цель уровня обслуживания

Основной элемент SLA между провайдером и клиентом. SLO согласовываются как меры оценки работы поставщика услуг и обеспечивают способ предотвращения споров между сторонами из-за недопонимания.

Stability - стабильность

Рабочее состояние, когда сеть не переходит из одного режима а другой (и обратно).

Supply-side congestion management - контроль перегрузок на основе предложений

Схема контроля перегрузок, предоставляющая дополнительные ресурсы сети для решения возникших или ожидаемых проблем с перегрузкой.

Traffic characteristic - характеристики трафика

Описание временного поведения или атрибутов данного потока или агрегата трафика.

Traffic-engineering system - система организации трафика

Набор объектов, механизмов и протоколов, применяемых совместно в целях организации трафика.

Traffic flow - поток трафика

Поток пакетов между двумя конечными точками, которых можно охарактеризовать определенным образом. Базовая классификация потока трафика использует квинтет из адресов и портов отправителя и получателя, а также идентификатора протокола. Потоки могут быть как мелкими и короткими, так и очень большими. Описываемые в документе методы TE будут, скорее всего, эффективней для больших потоков. В некоторых формах TE потоки могут агрегироваться и рассматриваться как единое целое, что позволяет применять TE к мелким потокам, образующим агрегат.

Traffic mapping - отображение трафика
Распределение создаваемой трафиком нагрузки между (заранее созданными) путями для выполнения определённых требований.

Traffic matrix - матрица трафика
Представление запросов трафика разных сочетаний абстрактных узлов отправителей и получателей. Абстрактный узел может состоять из одного или нескольких элементов сети.

Traffic monitoring - отслеживание трафика
Процесс наблюдения за характеристиками трафика в данной точке сети и сбора сведений для анализа и последующих действий.

Traffic trunk - транк трафика
Совокупность потоков трафика одного класса, пересылаемых об общему пути. Транк может характеризоваться узлами входа и выхода, а также набором атрибутов, указывающих характеристики поведения и требования к сети.

Workload - рабочая нагрузка
Оценка объёма работы, которую нужно выполнить в сети для удовлетворения потребностей в трафике (насколько загружена сеть). Иногда применяется термин traffic workload.

2. Основы

Целью Internet является быстрая, эффективная и экономичная передача пакетов IP от входных узлов к выходным. Кроме того, в средах с множеством классов обслуживания (например, Diffserv, см. параграф 5.1.1.2) параметры совместного использования ресурсов сети должны быть подобающим образом заданы и настроены в соответствии с правилами предпочтений и моделями обслуживания для разрешения конфликтов при получении ресурсов между пакетами, проходящими через сеть. Таким образом, необходимо рассмотреть вопросы конкуренции между потоками с одним (конфликты внутри класса) и разными (конфликты между классами) классами обслуживания.

2.1. Контекст Internet TE

Контекст организации трафика Internet включает несколько частей.

1. Контекст сетевого домена, определяющий рассматриваемую область действия и, в частности, ситуации, где возникают задачи TE. Контекст домена сети включает структуру сети, правила, характеристики, ограничения, атрибуты качества и критерии оптимизации.
2. Контекст задачи включает общие и конкретные вопросы, решаемые TE. Этот контекст включает идентификацию и абстрагирование относящихся к делу свойств, представление, формулировку и спецификацию требований к пространству решений, а также спецификацию желаемых свойств приемлемых решений.
3. Контекст решения, предлагающий способы решения вопросов, идентифицированных в контексте задачи. Этот контекст включает анализ, оценку вариантов, рекомендации и решение.
4. Контекст реализации и применения, где воплощаются принятые решения. Этот контекст включает планирование, организацию и выполнение.

Контекст Internet TE и различные варианты решения задач рассматриваются в последующих параграфах.

2.2. Контекст сетевого домена

Размер сетей IP варьируется от небольших кластеров маршрутизаторов, расположенных в данном месте, до тысяч соединённых между собой маршрутизаторов, коммутаторов и других компонентов, распределенных по планете.

На самом базовом уровне абстракции сеть IP можно рассматривать как распределенную динамическую систему, состоящую из:

- набора соединённых между собой ресурсов, обеспечивающих транспортные услуги для трафика IP с учётом некоторых ограничений;
- системы запрашивания (спроса), предоставляющей нагрузку (трафик), доставляемую через сеть;
- системы реагирования, состоящей из сетевых процессов, протоколов и соответствующих механизмов, облегчающих перемещение трафика через сеть (см. [AWD2]).

Элементы и ресурсы сети могут иметь конкретные характеристики, ограничивающие способы обработки запросов на доставку трафика. Кроме того, сетевые ресурсы могут быть оснащены механизмами управления трафиком, определяющими способы обслуживания потребностей. Эти механизмы могут служить для:

- управления действиями по обработке пакетов в рамках данного ресурса;
- арбитража доступа различных пакетов к ресурсу;
- регулирования поведения трафика на данном ресурсе.

Система обеспечения и управления конфигурацией может разрешать внешним или внутренним элементам устанавливать настройки механизмов управления трафиком для осуществления контроля над реагированием элементов сети на внешние и внутренние воздействия.

Детали передачи пакетов по сети задаются правилами, устанавливаемыми администраторами сетей через системы управления конфигурацией и обеспечения на основе политики. Обычно типы услуг, предоставляемых сетью, зависят также от технологии и характеристик элементов сети и протоколов, превалирующих моделей служб и полезности, а также от возможности сетевых администраторов транслировать правила в конфигурацию сети.

Сети в Internet имеют две важных характеристики:

- предоставление услуг в реальном масштабе времени;
- работа в быстро меняющихся средах.

Динамические характеристики сетей IP и IP/MPLS могут рассматриваться как часть изменений в потребностях, взаимодействий между разными сетевыми протоколами и процессами, быстрого развития инфраструктуры, нуждающегося в постоянном включении новых технологий и элементов сети, а также временных и сохраняющихся отказов в системах.

Пакеты борются за получение ресурсов сети в процессе их передачи через сеть. Сетевой ресурс считается перегруженным, если в течение некоторого времени скорость прибытия пакетов превышает выходную производительность ресурса. Перегрузка в сети может приводить к задержке и даже отбрасыванию некоторых прибывших пакетов.

Перегрузка сети увеличивает транзитные задержки и их вариации, может приводить к потере пакетов и снижать предсказуемость сетевых услуг. Хотя перегрузка может быть полезным инструментом на входных граничных узлах, перегрузка в сети крайне нежелательна. Борьба с перегрузками в сети при разумных затратах является одной из основных задач Internet TE, хотя сохранение разумных затрат может потребовать компромисса с другими целями.

Эффективное совместное использование ресурсов сети множеством потоков трафика является основной эксплуатационной предпосылкой для Internet. Фундаментальной задачей при эксплуатации сетей является повышение эффективности использования ресурсов при минимизации вероятности возникновения перегрузок.

Сеть Internet должна функционировать при наличии разных классов трафика со своими требованиями к обслуживанию. Это требование разъяснено в архитектуре дифференцированного обслуживания (Differentiated Services или Diffserv) [RFC2475]. В этом документе описывается, как пакеты можно группировать в разные агрегаты трафика, чтобы каждый из агрегатов имел свой набор поведенческих характеристик или общий набор требований к доставке, которые могут задаваться явно или неявно. Два наиболее важных требования к доставке трафика приведены ниже.

- Ограничения пропускной способности могут задаваться статистически как пиковые и минимальные скорости, размер пиков или в форме некого детерминированного указания пропускной способности.
- Требования QoS могут задаваться:
  - ограничениями целостности, такими как потери пакетов;
  - временными ограничениями, такими как ограничение времени доставки каждого пакета (задержка) или последовательности пакетов, относящихся к одному потоку трафика (вариации задержки).

2.3. Контекст задачи

Имеется несколько проблем, связанных с работой сети, которые похожи на описанные в предыдущем параграфе. В этом параграфе анализируется контекст задачи в части TE. Идентификация, абстрагирование, представление и измерение свойств сети, связанных с TE, имеют важное значение. Особую важность представляет формулировка задач, связанных с попытками организации трафика, например:

- способы идентификации требований к пространству решений;
- способы задания желаемых свойств решений;
- способы фактического решение задач;
- способы измерения и характеризации эффективности решений.

Другой класс задач связан со способами измерения и оценки параметров соответствующих состояний сети. Эффективная организация трафика основана на подобающей оценке предлагаемого трафика, а также на представлении о базовой топологии и соответствующих ограничениях ресурсов. Для автономного (offline) планирования требуется полное представление о топологии сети или её части, где планируется организовать трафик.

Ещё один класс задач связан со способами характеризации состояния сети и оценки её производительности. Задача оценки производительности имеет два аспекта, один из которых связан с оценкой производительности сети на системном уровне, другой - с оценкой производительности на уровне ресурсов, уделяющей внимание отдельным ресурсам сети.

В этом документе характеристики сети на системном уровне называются макросостояниями, а на уровне ресурсов - микросостояниями. Характеристики системного уровня также называют возникающими (emergent) свойствами сети. Соответственно, схемы TE, имеющие дело с оптимизацией производительности сети на системно уровне называются макро-TE, а на уровне ресурсов - микро-TE. При определённых обстоятельствах производительность на системном уровне можно вывести из производительности на уровне ресурсов, используя соответствующие правила композиции, в зависимости от конкретных показателей производительности, представляющих интерес.

Другой фундаментальный класс задач связан со способами эффективной оптимизации производительности сети. Оптимизация производительности может включать трансляцию решений для конкретных задач TE в конфигурацию сети. Оптимизация может также включать ту или иную степень контроля за управлением ресурсами, маршрутизацией и добавлением пропускной способности.

2.3.1. Перегрузки и их последствия

Перегрузка сети является одной из наиболее серьёзных проблем в контексте работы сетей IP. Элемент сети считается перегруженным, если он испытывает устойчивую перегрузку в течение определённого интервала времени. Хотя перегрузка на границе сети может быть полезна для доставки сетью как можно большего объёма трафика, перегрузка в сети почти всегда ведёт к снижению качества обслуживания конечных пользователей. Схемы предотвращения перегрузок и реагирования на них могут включать правила для спроса и предложения. Правила для спроса могут ограничивать доступ к перегруженным ресурсам или динамически регулировать спрос для смягчения ситуаций при перегрузке. Правила для предложения могут расширять или добавлять пропускную способность сети, чтобы приспособиться к предлагаемому трафику. Правила для предложения могут также перераспределять ресурсы сети для смены распределения трафика по инфраструктуре. Перераспределение трафика и ресурсов служит для повышения эффективной пропускной способности сети.

В данном документе основное внимание уделено схемам контроля перегрузок, входящим в сферу действия сети, а не схемам, зависящим от чувствительности и адаптивности конечных систем. Иными словами, в документе рассматриваются аспекты контроля перегрузок и решения, на которые могут влиять элементы управления, работающие в сети, и действия сетевых администраторов и систем управления сетью.

2.4. Контекст решения

Контекст решения для Internet TE включает анализ, оценку вариантов и выбор между вариантами действий. Обычно контекст решения основывается на выводах о текущем или будущем состоянии сети и принятии решений, которые могут включать предпочтения для того или иного набора действий. Более конкретно, контекст решения требует обоснованных оценок рабочей нагрузки, характеристики состояния сети, вывода решений, которые могут формулироваться явно или неявно, и возможно, выполнения набора управляющих воздействий. Действия по управлению могут включать манипуляции с параметрами, связанными с маршрутизацией, приобретением тактических мощностей и контролем над функциями управления трафиком. Ниже приведён список инструментов, которые могут применяться в контексте решения Internet TE

- Набор правил, целей и требований (которые могут зависеть от контекста) для оценки и оптимизации производительности сети.
- Набор интерактивных (online) а в некоторых случаях и автономных (offline) инструментов и механизмов для измерения, определения характеристик моделирования и управления трафиком, управления размещением и распределением сетевых ресурсов, а также управления отображением и распределением трафика по инфраструктуре.
- Набор ограничений для рабочей среды, сетевых протоколов и самой системы TE.
- Набор количественных и качественных методов и методик для абстрагирования, формулировки и решения задач TE.
- Набор административных параметров управления, которыми можно манипулировать с помощью системы управления конфигурацией. Такая система сама может включать подсистему управления конфигурацией, хранилище конфигурации и подсистему аудита конфигураций.
- Набор рекомендаций для оценки, оптимизации и повышения производительности сети.

Определение характеристик трафик по измерениям или оценкам очень полезно в сфере решений TE. Оценки трафика можно получить из сведений о подписках клиентов, проекций и моделей трафика, фактических измерений. Измерения могут выполняться на различных уровнях, например для агрегатов трафика или отдельных потоков. Измерения на уровне потока или небольшого агрегата трафика могут выполняться на граничных узлах, где трафик входит в сеть или выходит из неё. Измерения для больших агрегатов могут выполняться в ядре сети.

Для исследования производительности и поддержки планирования в имеющихся или будущих сетях может выполняться анализ маршрутизации с целью определения путей, которые протоколы маршрутизации будут выбирать для различных потребностей трафика, и определения загрузки сетевых ресурсов при маршрутизации трафика через сеть. Анализ маршрутизации включает выбор путей через сеть, распределение трафика по нескольким возможным маршрутам, мультиплексирование трафика IP через транки (при их наличии) и базовую инфраструктуру сети. Модель топологии сети можно получить из нескольких источников:

- документы по архитектуре сети;
- проект сети;
- конфигурационные файлы маршрутизаторов;
- базы данных маршрутизации, такие как базы сведений о состоянии каналов протоколов внутренней маршрутизации (Interior Gateway Protocol или IGP);
- таблицы маршрутизации;
- автоматизированные средства обнаружения и сбора сведений о топологии сети.

Сведения о топологии можно также получить от серверов мониторинга и обеспечения сети.

Маршрутизация в работающих сетях IP может управляться административно на разных уровнях абстракции, включая манипуляции с атрибутами BGP и метрикой IGP. Для ориентированных на пути технологий, таких как MPLS, дополнительное управление маршрутизацией может выполняться манипуляциями с соответствующими параметрами TE и ресурсов, а также административными ограничениями в правилах. В контексте MPLS путь явно маршрутизируемого LSP можно рассчитать и организовать разными способами:

- вручную;
- автоматически и интерактивно (online) с использованием процессов маршрутизации по ограничениям на маршрутизаторах с коммутацией по меткам (Label Switching Router или LSR);
- автоматически и автономно (offline) с использованием элементов маршрутизации по ограничениям, реализованных во внешних системах поддержки TE.

2.4.1. Борьба с перегрузками

Минимизация перегрузок является важнейшим аспектом организации трафика Internet. В этом параграфе приведён обзор общих подходов, предложенных или использованных для борьбы с перегрузками. Правила контроля перегрузок можно классифицировать по описанным ниже критериям (более подробный анализ схем контроля перегрузок приведён в [YARE95]).

1. Контроль перегрузок на основе временных рамок отклика.

- Долгий (от недель до месяцев). Расширение ёмкости сети за счёт добавления оборудования, маршрутизаторов и каналов занимает время и сравнительно дорого, что нужно учитывать при планировании пропускной способности. Расширение основывается на оценках и прогнозах будущего развития и распределения трафика. Такие обновления обычно занимают недели, месяцы и даже годы.

- Средний (от минут до дней). Например, указанные ниже правила управления.

  a) Настройка параметров протоколов маршрутизации для направления трафика в некоторые сегменты сети или от них.

  b) Организация и настройка явно маршрутизируемых LSP в сетях MPLS для направления транков трафика от потенциально перегруженных ресурсов сети на более предпочтительные маршруты.

  c) Перенастройка логической топологии сети для более точного соответствия пространственному распределению трафика с использованием, например, ориентированной на пути технологии, такой как MPLS LSP или оптические канальные трассы.

  Когда эти схемы являются адаптивными, они полагаются на системы измерений, которые отслеживают изменения в распределении трафика, нагрузки и расхода ресурсов сети, обеспечивая сигналы для интерактивных и автономных механизмов и инструментов TE, чтобы те могли запускать в сети управляющие воздействия. Механизмы и инструменты TE могут реализованы распределенным или централизованным образом. Централизованная схема может иметь полную информацию о состоянии сети и давать более оптимальные решения, однако в таких схемах имеется центральная точка отказа и они менее расширяемы по сравнению с распределенными схемами. Кроме того, используемые централизованной схемой сведения могут быть устаревшими и не отражающими фактического состояния сети. Целью этого документа не являются рекомендации по выбору централизованной или распределенной схемы, такой выбор остаётся за администраторами сети в соответствии с конкретными потребностями.

- Короткий (минуты и меньше). Эта категория включает функции обработки и события на уровне пакетов в интервале нескольких периодов кругового обхода. Она включает также механизмы маршрутизаторов, такие как активное и пассивное управление буферами. Все эти механизмы служат для контроля перегрузок и сигнализации о них конечным системам, чтобы те могли адаптивно регулировать скорость передачи трафика в сеть. Общеизвестной схемой управления очередями, особенно для такого отзывчивого трафика, как TCP, является случайное упреждающее обнаружение (Random Early Detection или RED) [FLJA93]. Во время перегрузки (но до заполнения очереди) схема RED выбирает поступающие пакеты для «маркировки» по вероятностному алгоритму с учётом среднего размера очереди. Маршрутизатор, не использующий явных уведомлений о перегрузке (Explicit Congestion Notification или ECN) [RFC3168], может просто отбрасывать помеченные пакеты для снижения перегрузки и неявного уведомления получателя о ней. С другой стороны, если маршрутизатор и конечный хост поддерживают ECN, они могут устанавливать поле ECN в заголовках пакетов и конечный хост может действовать на основе такой маркировки. Было предложено несколько вариантов RED для поддержки различных уровней предпочтения при отбрасывании в средах с разными классами [RFC2597]. RED обеспечивает предотвращение перегрузок не хуже, чем управление очередями Tail-Drop (TD)- отбрасывание прибывающих пакетов только после заполнения очереди. Важно, что RED снижает вероятность синхронизации пиков повторной передачи в сети и повышает уровень беспристрастности для сеансов трафика с разной реакцией. Однако RED, сам по себе, не может предотвратить перегрузки и отсутствие беспристрастности, связанные с источниками, не реагирующими на RED, например, некоторыми «жадными» потоками с некорректным поведением. Для повышения производительности и беспристрастности при наличии невосприимчивого трафика были предложены другие схемы. Некоторые из таких схем, например, отбрасывание в самой длинной очереди (Longest Queue Drop или LQD) и динамическое мягкое разделение со случайным отбрасыванием (Dynamic Soft Partitioning with Random Drop или RND) [SLDC98], были предложены как теоретические основы и обычно не поддерживаются в имеющейся коммерческой продукции, тогда как другие, например, примерно беспристрастное отбрасывание (Approximate Fair Dropping или AFD) [AFD03], были реализованы на практике. Рекомендации при применению схем активного управления очередями (Active Queue Management или AQM) представлены в [RFC7567], где рекомендуются алгоритмы AQM, подобные опубликованным IETF в [RFC8290], [RFC8033], [RFC8034], [RFC9332], на RED все ещё подходит для каналов со стабильной пропускной способностью при тщательной настройке.

2. Реактивные и упреждающие схемы контроля перегрузок.

   - Правила реактивного контроля перегрузок (восстановление) реагируют на возникшие перегрузки. Все описанные выше правила для краткосрочного и среднесрочного реагирования можно отнести к числу реактивных. Они основываются на отслеживании и идентификации возникающей в сети перегрузки с инициированием соответствующих действий для облегчения ситуации. Реактивные схемы контроля перегрузок могут быть одновременно упреждающими.

   - Правила проактивного контроля (предсказание, предотвращение) применяют упреждающие действия для предотвращения перегрузки на основе оценок и предсказаний возможности перегрузки (например, прогнозы картины трафика). Некоторые из правил, указанных для долгосрочных и среднесрочных масштабов, относятся к числу упреждающих. Превентивная политики не обязательно реагирует незамедлительно на возникающие перегрузки. Вместо этого рассматриваются прогнозы спроса на транспортировку и распределение нагрузки и могут предприниматься действия по предотвращению возможных в будущем перегрузок. Схемы, описанные для краткосрочных интервалов, также могут служить для предотвращения перегрузки, поскольку отбрасывание или маркировка пакетов до фактического заполнения очередей приведёт к снижению соответствующего восприимчивого трафика. Превентивный контроль перегрузок может быть в то же время и реактивным.

3. Контроль перегрузок со стороны спроса и предложения.

- Правила контроля перегрузок со стороны спроса увеличивают эффективную пропускную способность для трафика, чтобы контролировать или снижать перегрузку. Это может быть достигнута путём увеличения пропускной способности или балансирования распределения трафика по сети. Планирование пропускной способности нацелено на обеспечение физической топологии и связанной с этим полосы каналов для соответствия или превышения оценочного уровня трафика с учётом его распределения, прогнозов и бюджетных или иных ограничений. Если фактическое распределение трафика не соответствует топологии, выведенной при планировании пропускной способности, трафик можно сопоставить с топологией через механизмы протоколов маршрутизации, применение ориентированных на пути технологий (например, MPLS LSP и оптические трассы) для изменения логической топологии или с помощью иных механизмов перераспределения нагрузки.
- Правила контроля перегрузок со стороны предложения контролируют или регулируют предлагаемый трафик для снижения перегрузки. Например, некоторые из указанных выше краткосрочных механизмов, а также механизмы контроля и формирования скорости пытаются регулировать предлагаемую нагрузку.

2.5. Контекст реализации и применения

Рабочий контекст Internet TE характеризуется постоянными изменениями, которые происходят на разных уровнях абстракции. Для контекста реализации нужно эффективное планирование, организация и исполнение. Аспекты планирования могут включать определение набора предварительных действий для достижения желаемых целей. Организация включает назначение и распределение ответственности между компонентами системы TE, а также координацию действий для достижения желаемых целей TE. Исполнение включает измерения и применение корректирующих действий для достижения и поддержания желаемых целей TE.

3. Модели процессов TE

В этом разделе описывается общая модель процесса, отражающая высокоуровневые практические аспекты организации трафика Internet в рабочем контексте. Модель процесса описывается как последовательность действий, которые должны быть выполнены для оптимизации производительности работающей сети (см. также [RFC2702] и [AWD2]). Модель процесса может быть реализована явно или неявно программным процессом или человеком.

Модель процесса TE является интерактивной [AWD2]. Четыре фазы модели, указанные ниже, повторяются как непрерывная последовательность.

1. Определение соответствующих правил управления, регулирующих работу сети.
2. Получение результатов измерения в работающей сети.
3. Анализ состояния сети и определение характеристик нагрузочного трафика. Упреждающий анализ обнаруживает потенциальные проблемы, которые могут проявиться в будущем. Реактивный анализ обнаруживает имеющиеся проблемы и определяет их причины.
4. Оптимизация производительности сети, включая процесс принятия решений о выборе и исполнении набора действий из возможных по результатам трёх предыдущих этапов вариантов. Действия по оптимизации могут включать применение методов контроля предлагаемого трафика и распределения трафика по сети.

3.1. Компоненты модели

Основные компоненты модели процесса организации трафика указаны ниже.

1. Измерения имеют решающее значение для работы TE. Рабочее состояние сети можно достоверно узнать лишь с помощью измерений. Измерения также важны для оптимизации, поскольку они обеспечивают обратную связь для подсистем управления TE. Эти данные служат для адаптивной оптимизации производительности сети в ответ на события и и побудительные причины в сети и за её пределами. Измерения для поддержки функций TE могут выполняться на разных уровнях абстракции. Например, измерения могут служить для получения характеристик на уровне пакетов, потоков, пользователей или клиентов, агрегатов трафика, компонентов и сети в целом.
2. Важными аспектами Internet TE являются моделирование, анализ и имитация. Моделирование включает создание абстрактного или физического представления, отражающего соответствующие характеристики трафика и атрибуты сети. Модель сети - это абстрактное представление сети, отражающее соответствующие свойства, атрибуты и характеристики сети. Средства имитации сетей очень полезны для TE. Из-за сложности реалистического количественного представления поведения сети некоторые аспекты производительности можно эффективно изучить лишь с помощью имитационного моделирования.
3. Оптимизация производительности сети включает устранение сетевых проблем путём нахождения и реализации решений. Оптимизация производительности сети может быть корректировочной или полной. При корректирующей оптимизации цель состоит в устранении возникших или зарождающихся проблем. Целью полной оптимизации является повышение производительности сети даже при отсутствии проблем или их ожидания.

4. Таксономия систем TE

В этом разделе представлена краткая классификация систем организации трафика не основе стилей и подходов к TE, описанных ниже.

- В зависимости от времени, состояния или событий.
- Автономные и интерактивные (online).
- Централизованные и распределенные.
- По локальным или глобальным сведениям.

- Предписывающие и описательные.
- С открытым и закрытым контуром.
- Тактические и стратегические.

4.1. Управление по времени, состояниям и событиям

Методики организации трафика можно классифицировать как зависящие от времени, состояния или событий. В этом документе все схемы TE считаются динамическими. Статическая организация трафика не предполагает применения методики и алгоритмов TE, она является свойством планирования сети и не имеет реактивной и гибкой природы TE.

В зависящих от времени методиках TE применяются исторические сведения, основанные на периодических вариациях трафика (например, по времени суток), для предварительного программирования маршрутизации и других механизмов TE. Кроме того, может учитываться клиентская подписка и прогнозы трафика. Предварительно планируемая маршрутизация обычно изменяется сравнительно медленно (например, 1 раз в день). Зависящие от времени алгоритмы не пытаются приспособиться к краткосрочным вариациям трафика или изменениям условий в сети. Примером зависящего от времени алгоритма является централизованный оптимизатор, где на вход системы поступает матрица трафика и требования с несколькими классами QoS, как описано в [MR99]. Другим примером является приложение сбора данных о трафике Internet [AJ19], позволяющие применять алгоритмы машинного обучения для идентификации картин трафика в собранных с течением времени сведениях о трафике Internet и извлечения информации для принятия решений и повышения эффективности и продуктивности рабочих процессов.

В TE по состоянию планы маршрутизации приспосабливаются к текущему состоянию сети, что даёт дополнительные сведения о вариациях фактического трафика (т. е. возмущениях относительно регулярных вариаций), которые невозможно предсказать по историческим данным. Примером TE в зависимости от состояния является маршрутизация на основе ограничений, работающая в сравнительно долгосрочном масштабе времени. Примером для сравнительно коротких интервалов является алгоритм распределения нагрузки, описанный в [MATE]. Состояние сети может определяться на основе параметров, рассылаемых маршрутизаторами в лавинном режиме. Другой подход заключается в том, что конкретный маршрутизатор с адаптивной системой TE передаёт пробные пакеты по заданному пути для сбора сведений об этом пути. В [RFC6374] заданы расширения протоколов для сбора сведений о производительности сетей MPLS. Ещё один подход состоит в сборе системой управления сведений напрямую из элементов сети с использованием методов сбора данных телеметрии через публикацию и подписку [RFC7923]. Своевременный сбор и распространение данный очень важны для адаптивных TE. Зависящие от времени алгоритмы подходят для предсказуемых изменений трафика, а алгоритмы на основе состояний могут требоваться для повышения эффективности сети и обеспечения устойчивости к смене состояний сети.

Методы TE, зависящие от событий, также могут применяться для выбора пути TE. Эти методы отличаются от методов TE, зависящих от времени и состояний, по способу выбора путей. Алгоритмы являются адаптивными и распределенными по своей природе и обычно используют модели обучения для поиска в сети хороших путей TE. В то время как модели TE на основе состояний для выбора путей TE применяют лавинную рассылку в доступной полосе канале (available-link-bandwidth или ALB) [E.360.1], методам TE на основе событий не нужна лавинная рассылка ALB. Вместо этого они обычно находят пропускную способность по моделям обучения, как в методе STT (success-to-the-top) [RFC6601]. Лавинная рассылка ALB может отнимать много ресурсов, поскольку для неё нужна пропускная способность на передачу маршрутных анонсов состояний каналов и процессорное время для обработки этих анонсов. Кроме того, издержки на анонсы ALB и их обработку могут ограничивать размер области (area) и AS. Результаты моделирования показывают, что методы TE на основе событий могут приводить к снижению издержек ALB без потери пропускной способности сети [TE-QoS-ROUTING].

Полнофункциональные системы TE вероятно будут использовать все аспекты зависящих от времени, состояния и событий методов, как описано в параграфе 4.3.1.

4.2. Автономные и интерактивные системы

Для организации трафика требуется расчёт маршрутных планов, который можно выполнить автономно или в сети (online). Автономный расчёт подходит для случаев, когда планы не требуется исполнять в реальном масштабе времени. Например, маршрутные планы на основе предсказаний, можно рассчитывать автономно. Обычно автономные расчёты применяются также для широкого поиска в многомерном пространстве решений.

Интерактивный расчёт требуется в случаях, когда маршрутные планы должны приспосабливаться к изменениям состояний сети в зависящих от состояния алгоритмах. В отличие от автономных расчётов (которые могут быть объёмными), интерактивный расчёт ориентирован на сравнительно простые и быстрые операции для выбора маршрута, тонкой настройки выделения ресурсов и распределения нагрузки.

4.3. Централизованные и распределенные системы

При централизованном управлении существует центральный орган, определяющий маршрутные планы и, возможно, другие параметры управления TE от имени каждого маршрутизатора. Этот орган может периодически собирать от всех маршрутизаторов сведения о состоянии сети и передавать им маршрутные данные. Цикл обновления для обмена сведениями в обоих направлениях является очень важным параметром, напрямую влияющим на производительность управляемой сети. Для централизованного управления могут потребоваться значительные вычислительные ресурсы и пропускная способность.

При распределенном управлении маршруты выбирает каждый маршрутизатор автономно на основе своего представления о состоянии сети. Сведения о состоянии сети маршрутизаторы могут получать с помощью зондирования или от других маршрутизаторов на периодической основе через анонсы состояний каналов. Сведения о состоянии сети могут распространяться и в исключительных ситуациях. Примеры протокольных расширений, применяемых для анонсирование сведений о состоянии каналов, определены в [RFC5305], [RFC6119], [RFC7471], [RFC8570], [RFC8571]. См. также параграф 5.1.3.9.

4.3.1. Гибридные системы

На практике большинство систем TE буду представлять гибрид централизованного и распределенного управления. Например, популярный в MPLS подход к TE заключается в применении центрального контроллера на основе элемента расчёта пути (Path Computation Element или PCE) с учётом состояний, а протоколы маршрутизации служат для принятия локальных решений на маршрутизаторах внутри сети. Локальные решения позволяют быстрее реагировать на события в сети, но могут возникать конфликты между решениями разных маршрутизаторов.

В сетевых операциях для систем TE может также применяться гибрид автономных и сетевых расчётов. Пути TE могут рассчитываться заранее на основе сведений о стабильной сети и планируемых потребностях трафика, но изменяться в активной сети в зависимости от изменений состояния и нагрузочного трафика. Кроме того, реакция на события в сети может быть рассчитана автономно для быстрого реагирования без добавочных расчётов или определяться интерактивно (online) в зависимости от природы событий.

4.3.2. Соображения для программно-определяемых сетей (SDN)

Как отмечено в параграфе 5.1.2.2, одним из основных факторов развития программно-определяемых сетей (Software-Defined Networking или SDN) является отделение плоскости управления сетью от плоскости данных [RFC7149]. Однако SDN может также обеспечивать централизованное управление ресурсами и облегчать взаимодействие приложений с сетью через интерфейс прикладных программ (Application Programming Interface или API), как описано в [RFC8040]. Сочетание этих свойств обеспечивает гибкость сетевой архитектуры, позволяя приспосабливать сетевые требования к разным приложениям верхнего уровня. Это часто называют программируемой сетью [RFC7426].

Централизованное управление SDN помогает улучшить использование ресурсов сети по сравнению с распределенным управлением, где локальные правила зачастую могут преобладать над целями маршрутизации всей сети. В среде SDN плоскость данных пересылает трафик нужным адресатам. Однако до попадания трафика в плоскость данных логически централизованная плоскость управления SDN часто определяет путь, по которому трафик приложения пойдёт в сети. Поэтому плоскости управления SDN нужно знать топологию и возможности базовой сети, а также текущие состояния ресурсов узлов и каналов.

С использованием основанной на PCE схемы управления SDN [RFC7491] топологию сети можно раскрыть путём запуска пассивного экземпляра OSPF или IS-IS, а также через BGP Link State (BGP-LS) [RFC9552]) для генерации базы организации трафика (Traffic Engineering Database или TED) (см. параграф 5.1.3.14). PCE служит для расчёта пути (см. параграф 5.1.3.11) на основе TED и доступной пропускной способности, а затем может быть выполнена оптимизация пути на основе запрошенных целевых функций [RFC5541]. Когда подходящий путь рассчитан, программирование явного сетевого пути можно выполнить с использованием протокола сигнализации, проходящего по всему пути [RFC3209], или поэтапно (per-hop) с непосредственным программированием каждого узла контроллером SDN [RFC8283].

Используя централизованный подход к управлению сетью, можно получить дополнительные преимущества, включая глобальную одновременную оптимизацию (Global Concurrent Optimization или GCO) [RFC5557]. Запрос расчёта пути GCO будет использовать топологию сети и сигнальные запросы пути вместе с соответствующими ограничениями для оптимального размещения в сети. Поэтому расчёты на основе GCO можно применять для пересчёта имеющихся сетевых путей с целью оптимизации трафика и снижения перегрузок.

4.4. Локальные и глобальные сведения

Для алгоритмов организации могут требоваться локальные и глобальные сведения о состоянии сети.

Локальные сведения содержат данные о состоянии части домена, например, о пропускной способности и потерях пакетов на определённом пути или о состоянии и возможностях сетевого канала. Локальных сведений может быть достаточно для некоторых экземпляров распределенного управления TE.

Глобальные сведения содержат информацию о состоянии всего домена TE. Примеры таких сведений включают глобальную матрицу трафика и данные о загрузке каждого канала в интересующем домене. Глобальные сведения обычно требуются для централизованного управления, но могут быть нужны и распределенным системам TE.

4.5. Предписывающие и описательные системы

Системы TE можно разделить на предписывающие и описательные.

В предписывающей организации трафика оцениваются варианты и рекомендуется курс действий. Такие системы можно дополнительно разделить на корректирующие и совершенствующие (perfective). Корректирующие TE предписывают курс действий для устранения имеющихся или прогнозируемых аномалий. Совершенствующие TE предписывают курс действий для развития и повышения производительности сети даже при отсутствии аномалий.

Описательная организация трафика характеризует состояние сети и оценивает влияние различных правил, не рекомендуя каких-либо конкретных действий.

4.5.1. Сети на основе намерений

Одним из способов выразить запрос на обслуживание является намерение (intent). Сети на основе намерений проще в управлении и эксплуатации, требуя лишь минимального вмешательства. Намерения определены в [RFC9315].

> Набор операционных целей (которым следует соответствовать сети) и результатов (которые сети следует обеспечивать), определённый декларативно без указания способов их достижения и реализации.

Намерение представляет данные и функциональную абстракцию, что позволяет пользователям и операторам не заботиться о низкоуровневой конфигурации устройств и механизмах для реализации данного намерения. Такой подход концептуально проще для пользователя, но может быть менее выразительным в плане ограничений и рекомендаций.

Сети на базе намерений применимы для TE, поскольку многие цели высокого уровня можно указать как намерение (например, распределение нагрузки, предоставление услуги, устойчивость к отказам). Намерения преобразуются системой управления в действия TE внутри сети.

4.6. Системы с открытым и закрытым контуром

Управлением организацией трафика с открытым контуром считается управление, не использующее обратной связи о текущем состоянии сети. Однако управляющие действия могут использовать свои локальные сведения для целей учёта. Управление организацией трафика с закрытым контуром использует данные обратной связи о состоянии сети. Эти данные могут иметь форму текущих измерений или свежих исторических записей.

4.7. Тактические и стратегические системы

Тактическая организация трафика нацелена на решение конкретны проблем производительности (например, горячих точек), возникающих в сети, с тактической точки зрения без учёта общих стратегических задач. Без надлежащего планирования и понимания тактическая TE имеет тенденцию к ситуативным решениям. Стратегические подходы к задачам TE применяет более организованное и системное рассмотрение с учётом ближайших переспектив и долгосрочных последствий предпринятых действий и правил.

5. Обзор методов TE

В этом параграфе дан краткий обзор связанных с TE подходов, предложенных и реализованных в телекоммуникационных и компьютерных сетях с использованием протоколов и архитектуры IETF, по трём категориям:

- механизмы TE, соответствующие определениям из параграфа 1.2;
- подходы, основанные на этих механизмах TE;
- методы, применяемые в этих подходах и механизмах TE.

Рассмотрение не претендует на полноту и предназначено в основном для освещения имеющихся подходов к TE в сети Internet. Исторический обзор TE в телекоммуникационных сетях приведён в разделе 4 [RFC3272], а в параграфе 4.6 этого документа представлен обзор некоторых ранних подходов к TE, разработанных другими органами стандартизации. В задачи этого документа не входит анализ истории TE или перечисление связанной с TE работы других органов стандартизации (Standards Development Organization или SDO).

5.1. Обзор проектов IETF, связанных с TE

В этом параграфе приводится обзор ряда мероприятий IETF, относящихся к организации трафика Internet. Некоторые из этих технологий широко внедрены, другие получили меньшее распространение, а некоторые ещё не одобрены или находятся в стадии разработки.

5.1.1. Механизмы IETF TE

5.1.1.1. Интегрированные услуги

В IETF разработана модель интегрированных услуг (Integrated Services или Intserv), требующая предварительного выделения ресурсов, таких как пропускная способность и буферы, для данного потока трафика с целью исполнения запрошенных этим трафиком требований QoS. Модель Intserv включает компоненты, дополняющие принятые в модели обслуживания по возможности (best-effort), такие как классификаторы и планировщики пакетов, а также контроль допуска. Классификаторы пакетов служат для идентификации потоков, получающих определённый уровень обслуживания, планировщики реализуют планирование обслуживания разных потоков пакетов в соответствии с обязательствами QoS, а контроль допуска служит для проверки наличия у маршрутизатора ресурсов, требуемых для нового потока.

Основной проблемой модели Intserv является расширяемость [RFC2998], особенно в больших общедоступных сетях IP, где могут одновременно существовать миллионы активных потоков трафика. Уведомления о наступающей перегрузке (Pre-Congestion Notification или PCN) [RFC5559] решают проблему расширяемости Intserv за счёт инструментального контроля допуска (и прерывания потоков при отказах) между граничными узлами. Узлы между границами межсетевого обмена не применяют операций на уровне потоков, а граничные узлы могут использовать для потока или агрегата потоков протокол резервирования ресурсов (Resource Reservation Protocol или RSVP).

Примечательной особенностью модели Intserv является необходимость явной сигнализации требований QoS от конечных систем к маршрутизаторам [RFC2753]. Протокол RSVP (см. параграф 5.1.3.2) реализует эту функцию и является критически важным для модели Intserv.

5.1.1.2. Дифференцированные услуги

Целью дифференцированного обслуживания (Differentiated Services или Diffserv) в рамках IETF было создание расширяемых механизмов классификации трафика по агрегатам поведения, в конечном итоге приводящих к разной трактовке поведения каждого агрегата, особенно в случаях нехватки ресурсов (таких как пропускная способность каналов и ёмкость буферов) [RFC2475]. Одним из основных мотивов создания Diffserv была разработка дополнительных механизмов дифференциации услуг в Internet для смягчения проблем расширяемости, присущих модели Intserv.

В Diffserv применяется поле DS заголовка IP (6 битов) которое исходно служило октетом типа обслуживания (Type of Service или TOS). Поле DS служит для указания режима пересылки, который следует обеспечивать пакету на транзитных узлах [RFC2474]. Diffserv включает концепцию групп поведения на этапах пересылки (Per-Hop Behavior или PHB). С помощью PHB можно задать несколько классов обслуживания, используя различные правила классификации, контроля, формирования и планирования трафика.

Для применения конечным пользователем услуг Diffserv, предоставляемых его сервис-провайдером (Internet Service Provider или ISP), ему может потребовать заключение соглашения об уровне обслуживания (SLA) с ISP. В SLA может явно или неявно включаться соглашение о кондиционировании трафика (Traffic Conditioning Agreement или TCA), задающее правила классификации, измерения, маркировки, отбрасывания и формовки трафика.

Пакеты классифицируются с возможным применением правил и формовки на входе в сеть Diffserv. При прохождении пакетов через границу между доменами Diffserv значение поля DS в них может изменяться в соответствии с

имеющимися между доменами соглашениями. Diffserv разрешает лишь конечное число классов обслуживания, которые можно указать в поле DS. Основным преимуществом подхода Diffserv по сравнению с моделью Intserv является расширяемость. Ресурсы выделяются по классам и объем информации о состояниях пропорционален числу классов, а не числу потоков приложений.

После планирования сети и маркировки пакетов на её границе модель Diffserv решает вопросы управления трафиком на каждом узле (per-hop). Модель управления Diffserv состоит из набора управляющих механизмов микро-TE. Для обеспечения приемлемого качества обслуживания в сетях Diffserv нужны и другие возможности TE, такие как управление пропускной способностью (включая управление маршрутизацией). Для лучшего представления Diffserv в рамках всего домена введена концепция поведения в домене (Per-Domain Behavior) [RFC3086].

Процедуры Diffserv применимы и в контексте MPLS (см. параграф 6.8).

5.1.1.3. SR Policy

SR Policy [RFC9256] является развитием SR (см. параграф 5.1.3.12) для расширения возможностей TE в SR. Это модель, позволяющая создавать на узле упорядоченный список сегментов для реализации правил маршрутизации от источника с конкретным намерением направлять трафик от этого узла. SR Policy указывается триплетом <headend, color, endpoint>, где headend - это IP-адрес узла, на котором установлены правила, endpoint - IP-адрес цели правил, а color — индекс, связывающий SR Policy с намерением (например, малая задержка).

Головной узел (headend) получает уведомления о правилах SR и связанных с ними путях SR через конфигурацию или расширения протоколов, таких как коммуникационный протокол PCE (Path Computation Element Communication Protocol или PCEP) [RFC8664] или BGP [SR-TE-POLICY]. Каждый путь SR состоит из списка сегментов (путь SR с маршрутизацией от источника) и головной узел использует параметры endpoint и color для классификации пакетов в соответствии с SR Policy и определения пути для их пересылки. Если правила SR связаны с набором путей SR, каждый из путей имеет все для взвешенного распределения нагрузки. Кроме того, с набором полей SR может быть связано несколько SR Policy, чтобы по одним путям передавалось несколько потоков трафика.

С каждым путём-кандидатом, связанным с SR Policy или с самой политикой SR можно связать идентификатор привязки (SR Binding SID или BSID). Головной узел устанавливает запись с ключом BSID в плоскости пересылки и назначает ей действие по направлению пакетов, соответствующих записи, на выбранный путь SR Policy. Это направление можно реализовать разными способами.

По SID
: Идентификатор сегмента (SID) входящих пакетов совпадает с локальным BSID на головном узле.

По адресату
: Входящие пакеты соответствуют маршруту BGP/Service, указывающему SR Policy.

По потоку
: Входящие пакеты соответствуют массиву пересылки (например, классический 5-tuple), указывающему SR Policy.

По правилам
: Входящие пакеты соответствуют правилу маршрутизации, направляющему их в SR Policy.

5.1.1.4. TE на основе транспорта L4

В дополнение к механизмам TE на основе IP могут рассматриваться подходы на основе транспорта L4 в соответствующем контексте развёртывания (например, ЦОД и многодомные системы). Например, 3GPP определяет сервисные функции направления, коммутации и расщепления трафика доступа (Access Traffic Steering, Switching, and Splitting или ATSSS) [ATSSS], как указано ниже.

Access Traffic Steering - направление трафика доступа
: Выбор сети доступа для нового потока и передача трафика этого потока через выбранную сеть доступа.

Access Traffic Switching - коммутация трафика доступа
: Перенос всех пакетов действующего потока из одной сети доступа в другую с использованием в каждый момент лишь одной сети доступа.

Access Traffic Splitting - расщепление трафика доступа
: Пересылка пакетов потока через несколько сетей доступа одновременно.

Плоскость управления предоставляет хостам и конкретным сетевым устройствам набор правил, определяющих, каким потокам можно использовать услуги ATSSS. Трафик, соответствующий правилам ATSSS, может распределяться между сетями доступа в одном из 4 режимов, указанных ниже.

Active-Standby - активный-резервный
: Трафик передаётся через конкретный (активный) доступ и коммутируется в другой (резервный) при недоступности активного.

Priority-based - по приоритетам
: Системам доступа назначается приоритет, указывающий, какая из них используется первой. Трафик соответствующего потока направляется в доступ и наивысшим приоритетом, пока не возникает перегрузка, после чего пересылается в доступ со следующим уровнем приоритета.

Load-Balancing - распределение нагрузки
: Трафик распределяется между сетями доступа в заданном процентном отношении (например, 75% и 25%).

Smallest Delay - наименьшая задержка
: Трафик пересылается через систему доступа с наименьшим временем кругового обхода (round-trip time или RTT).

Для управления ресурсами хосты и сетевые устройства поддерживают такие средства, как контроль перегрузок, измерение RTT и планирование пакетов.

С целью предоставления услуг ATSSS для трафика TCP применяются Multipath TCP [RFC8684] и 0-RTT Convert Protocol [RFC8803], для UDP - Multipath QUIC [QUIC-MULTIPATH] и Proxying UDP in HTTP [RFC9298]. Отметим, что QUIC поддерживает процедуру переноса соединений, позволяющую партнёрам менять свои транспортные координаты L4 (адреса IP и номера портов) без прерывания базового соединения QUIC. Расширения протокола контроля перегрузок для дейтаграмм (Datagram Congestion Control Protocol или DCCP) [RFC4340] с поддержкой операций по нескольким путям, заданы в [MULTIPATH-DCCP].

5.1.1.5. Детерминированные сети

Архитектура детерминированных сетей (Deterministic Networking или DetNet) [RFC8655] предназначена для приложений с критическими требованиями по времени и надёжности. Многоуровневая архитектура сосредоточена в основном на развитии возможностей служб DetNet в плоскости данных [RFC8938]. Подуровень сервиса DetNet обеспечивает набор функций репликации, устранения и упорядочения пакетов (Packet Replication, Elimination, and Ordering Functions или PREOF) для сквозной гарантии обслуживания. Подуровень пересылки DetNet обеспечивает соответствующие гарантии пересылки (низкие потери, ограниченная задержка, упорядоченная доставка), используя механизмы выделения ресурсов и явного задания маршрутов. Разделение на два подуровня обеспечивает гибкую адаптацию возможностей DetNet к ряду механизмов TE, таких как IP, MPLS, SR. Ещё более важна связь между чувствительными ко времени сетями (Time Sensitive Networking или TSN) IEEE 802.1 [RFC9023], развёрнутыми в системах промышленного управления и автоматизации (Industry Control and Automation Systems или ICAS).

DetNet можно считать специализированным вариантом TE, поскольку архитектура обеспечивает явный набор оптимизированных путей с выделением запрашиваемых ресурсов. Приложение DetNet может указать атрибуты QoS или поведение трафика с помощью любой комбинации функций DetNet, описанных на подуровнях. Затем они распространяются и предоставляются с применением механизмов управления и обеспечения, приспособленных для организации трафика.

Для DetNet требуется значительный объём сведений о состояниях для поддержки дисциплин очередей по большому числу отдельных потоков. Это может оказаться достаточно сложным для работы сети при некоторых событиях, таких как отказы, изменение объёма трафика или перераспределение ресурсов. Поэтому в DetNet рекомендуется поддерживать агрегаты потоков, однако при этом все равно требуется большое число сигналов управления для организации и поддержки потоков DetNet.

Отметим, что в DetNet могут проявляться некоторые проблемы расширяемости, отмеченные для Intserv в параграфе 5.1.1.1, но область действия DetNet обычно меньше и такие проблемы менее часты.

5.1.2. Подходы IETF на основе механизмов TE

5.1.2.1. Оптимизация трафика прикладного уровня

В этом документе описаны различные механизмы TE, доступные в сети, однако в общем случае распределенные приложения (в частности, «жадные» до пропускной способности приложения P2P, применяемые, например, для совместного использования файлов) не могут использовать эти методы напрямую. В соответствии с [RFC5693] приложения могут значительно улучшить распределение и качество трафика за счёт взаимодействия с внешними источниками, знающими топологию сети. Решение задачи оптимизации трафика прикладного уровня Application-Layer Traffic Optimization или ALTO) означает, с одной стороны, внедрение службы ALTO для предоставления приложениям сведений о базовой сети (например, структуры базового размещения и предпочтений для сетевых путей), а с другой - усовершенствование приложений в плане применения этих данных для лучшего, чем случайный, выбора конечных точек с которыми организуются соединения.

Основная функция ALTO основана на абстрактных планах сети, обеспечивающих упрощённое представление, но содержащих достаточно сведений для их эффективного использования приложениями. На основе этих планов строятся дополнительные службы. В [RFC7285] описан протокол, реализующий услуги ALTO как интерфейс публикации сведений, позволяющий сети публиковать информацию о себе для сетевых приложений. Эта информация может включать местоположение узлов сети, группы соединений между узлами, упорядоченными по стоимости в соответствии с настраиваемой детализацией, а также свойства конечных хостов. Сведения, публикуемые протоколом ALTO, следует делать полезными для сети и приложений. Протокол ALTO имеет соответствующее REST устройство и представляет свои запросы и отклики с использованием модульного представления JSON [RFC8259] путём деления публикации сведений ALTO на множество услуг ALTO (например, Map Service, Map-Filtering Service, Endpoint Property Service, Endpoint Cost Service).

В [RFC8189] задана новая служба, позволяющая клиенту ALTO извлекать несколько показателей стоимости через один запрос для отфильтрованной карты стоимости ALTO и карту стоимости конечных точек. [RFC8896] расширяет службу сведений о стоимости ALTO, чтобы приложения могли решать не только «где» присоединиться, но и «когда» это делать. Это полезно для приложений, которым нужна массовая передача данных с её планированием (например, в часы малой загрузки). В [RFC9439] введены показатели производительности сети, включая задержку и её вариации, долю теряемых пакетов, число интервалов пересылки (hop) и пропускную способность. Сервер ALTO может выводить и агрегировать такие показатели из BGP-LS (см. параграф 5.1.3.10), IGP-TE (см. параграф 5.1.3.9) или инструментов управления и раскрывать эти сведения, чтобы позволить приложениям определить, где подключаться к сети на основе критериев производительности. Рабочая группа ALTO оценивает использование свойств TE в сети при принятии решений о новых вариантах использования, таких как пограничные расчёты и объединение ЦОД.

5.1.2.2. Визуализация и абстрагирование сети

Одной из основных движущих сил SDN [RFC7149] является отделение плоскости управления сетью от плоскости данных. Это разделение можно обеспечить для сетей TE путём развития MPLS (параграф 5.1.3.3) и GMPLS (параграф 5.1.3.5) и элементов PCE (параграф 5.1.3.11). Одним из преимуществ SDN является логическая централизация управления, позволяющая полностью видеть базовые сети. Централизованное управление в SDN помогает улучшить использование ресурсов сети по сравнению с распределенным управлением.

Абстракция и управление сетями TE (Abstraction and Control of TE Networks или ACTN) [RFC8453] задаёт иерархическую архитектуру SDN, описывающую функциональные элементы и методы согласования ресурсов в нескольких доменах для предоставления композитных услуг с организацией трафика. ACTN облегчает составные междоменные соединения и предоставляет их пользователю. Основные задачи ACTN указаны ниже.

- Абстрагирование ресурсов базовой сети и способов их предоставления приложениям и клиентам вышележащих уровней.
- Виртуализация базовых ресурсов для использования клиентами, приложениями и службами. Создание виртуализованных сред позволяет операторам видеть и контролировать многодоменные сети как единую виртуализованную сеть.

- Представление сетей клиентам в виде виртуальной сети через открытые и программируемые интерфейсы.

Управляемая ACTN инфраструктура создаётся из сетевых ресурсов организации трафика, которые могут включать статистическую пропускную способность для пакетов, физические источники плоскости пересылки (такие как длины волн и временные интервалы), а также возможности пересылки и кросс-соединений. Виртуализация в ACTN позволяет клиентам и приложениям (арендаторам) использовать и независимо управлять выделенными виртуальными ресурсами сети как будто они являются их собственными. Сеть ACTN «расслоена» (sliced) и арендаторам предоставляется лишь частичное и абстрагированное представление топологии базовой физической сети.

5.1.2.3. Расслоение сети

IETF Network Slice - это логическая топология сети, соединяющая множество конечных точек с использованием набора общих или выделенных ресурсов сети [NETWORK-SLICES]. Ресурсы применяются для выполнения конкретных SLO, заданных клиентами.

Сетевые слои, сами по себе, не являются конструкциями TE, однако оператор сети, предлагающий такое расслоение, будет, скорей всего, применять множество инструментов TE для управления своей сетью и предоставления услуг. Слои сети IETF определены так, что они не зависят от базовой инфраструктуры соединений и применяемой технологии. С точки зрения клиента IETF Network Slice выглядит как матрица связности VPN с дополнительными сведениями об уровне обслуживания, который нужен клиенту между конечными точками. С точки зрения оператора IETF Network Slice выглядит как набор инструкций по маршрутизации и туннелированию с резервированием сетевых ресурсов, требуемых для обеспечения уровня обслуживания, заданного в SLO. Концепция IETF Network Slice согласуется и расширенными VPN [ENHANCED-VPN].

5.1.3. Методы IETF, используемые TE

5.1.3.1. Маршрутизация на основе ограничений

Маршрутизация на основе ограничений относится к классу систем маршрутизации, рассчитывающим маршруты через сеть с учётом соблюдения набора требований и ограничений. В самом общем случае такая маршрутизация может также стремиться к оптимизации производительности сети при минимизации затрат. Ограничения и требования могут вноситься самой сетью или административными правилами. Ограничения могут включать пропускную способность, число интервалов пересылки (hop), задержку и инструменты политики, такие как атрибуты классов ресурсов, а также зависимые от домена атрибуты некоторых сетевых технологий и контекста, ограничивающие пространство решений функции маршрутизации. Ориентированные на пути технологии, такие как MPLS, делают маршрутизацию на основе ограничений осуществимой и привлекательной для IP-сетей общего пользования.

Концепция маршрутизации на основе ограничений в контексте требований MPLS-TE в сетях IP была впервые описана в [RFC2702] и привела к разработке MPLS-TE [RFC3209], как указано в параграфе 5.1.3.3.

В отличие от маршрутизации на основе QoS (например, [RFC2386], [MA], [PERFORMANCE-ROUTING]), где обычно решаются задачи маршрутизации отдельных потоков трафика для выполнения предписанных требований QoS при условии доступности сетевых ресурсов, маршрутизация на основе ограничений применима к агрегатам и потокам трафика с возможностью выполнения широкого круга ограничений, включая ограничения политики.

5.1.3.1.1. Гибкие алгоритмы IGP

Обычный подход к маршрутизации в сети IGP основан на определении протоколами IGP «кратчайших путей» через сеть на основе лишь метрики IGP, назначенной для каналов. С таким подходом часто связаны ограничения - трафик может сгущаться в направлении адресата, что может вызывать перегрузку, а также невозможно направить трафик по путям в соответствии со сквозным качеством, запрашиваемым приложениями.

Для преодоления этих ограничений широко распространены различные виды TE, как описано в этом документе, где компоненты TE отвечают за расчёт пути с учётом дополнительных показателей и/или ограничений. Такие пути (или туннели) нужно помещать в таблицы пересылки маршрутизаторов в дополнение или на замену рассчитанных IGP путей. Основным недостатком таких подходов TE является усложнение протоколов и управления, а также необходимость поддержки состояний в сети.

Гибкие алгоритмы IGP [RFC9350] позволяют протоколам IGP строить пути через сеть с учётом ограничений, определяя следующий узел (hop) на основе ограничений. Целью гибких алгоритмов является снижение сложности TE за счёт разрешения протоколу IGP выполнять некоторые базовые расчёты TE. Гибкий алгоритм включает набор расширений IGP, который позволяет маршрутизатору передавать TLV:

- описывающие набор ограничений для технологии;
- определяющие тип расчёта;
- описывающие тип метрики для расчёта лучших путей через топологию с ограничениями.

Заданная комбинация типов расчёта и метрики, а также ограничений называется определением гибкого алгоритма (Flexible Algorithm Definition или FAD). Маршрутизатор, передающий такой набор TLV, также назначает конкретный идентификатор (гибкий алгоритм) такой заданной комбинации.

Имеется два варианта применения гибкого алгоритма в сетях IP [RFC9502] и сетях SR [RFC9350]. В первом случае гибкий алгоритм рассчитывает пути к адресам IPv4 или IPv6, во втором - к Prefix SID (см. параграф 5.1.3.12).

Некоторые примеры, где использование гибких алгоритмов может быть полезно, приведены ниже.

- Расширение функций показателей производительности IP [RFC5664], где в сети может быть организована конкретная маршрутизация на основе ограничений (гибкий алгоритм) по результатам измерения производительности.
- Формирование «базовой» сети с использованием гибких алгоритмов и реализация «наложенной» сети с использованием методов TE. Такой подход позволяет использовать вложенную комбинацию гибкого алгоритма и расширений TE для IGP (см. параграф 5.1.3.9).

- Гибкие алгоритмы в SR-MPLS (параграф 5.1.3.12) могут служить основой для простого создания топологии в стиле TE без компонентов TE на маршрутизаторах и использования PCE (см. параграф 5.1.3.11).
- Поддержка сетевых слоёв (slice) [NETWORK-SLICES] где SLO конкретного IETF Network Slice может гарантироваться гибким алгоритмом или с использованием гибкого алгоритма может быть создана отфильтрованная топология (Filtered Topology) [NETWORK-SLICES] в стиле топологии TE.

5.1.3.2. RSVP

RSVP - это протокол сигнализации «мягкого состояния» (soft-state) [RFC2205]. Протокол поддерживает резервирование ресурсов для индивидуальных и групповых потоков по инициативе получателей. RSVP был разработан как сигнальный протокол модели интегрированных услуг (IntServ, см. параграф 5.1.1.1) для приложений, передающих требования QoS в сеть с целью резервирования соответствующих ресурсов QoS [RFC2205].

В RSVP отправитель трафика или узел-источник передаёт сообщение Path получателю трафика с теми же адресами отправителя и получателя, которые будет использовать отправитель. Сообщение Path включает:

- спецификацию трафика от отправителя, описывающую характеристики трафика;
- шаблон отправителя, задающий формат трафика;
- необязательная спецификация анонса, используемая для поддержки концепции «один путь с анонсами» (One Pass With Advertising или OPWA) [RFC2205]

Каждый промежуточный маршрутизатор на пути пересылает сообщение Path следующему узлу, указанному протоколом маршрутизации. Принявший сообщение Path получатель отвечает сообщением Resv, включающим дескриптор потока, служащим для запроса резервирования ресурсов. Сообщение Resv проходит к отправителю или узлу-источнику по обратному направлению пути, пройденного сообщением Path. Каждый промежуточный маршрутизатор на пути может принять или отклонить резервирование по запросу из сообщения Resv. Если маршрутизатор отклоняет запрос, он отправляет получателю сообщение об ошибке и сигнальный процесс прерывается. Если запрос воспринят, для потока резервируется пропускная способность и буферное пространство, а в маршрутизаторе для потока устанавливается соответствующее состояние.

Одна из проблем исходной спецификации [RFC2205] связана с расширяемостью. Это обусловлено тем, что резервирование требовалось для микропотоков, что обычно вело к линейному росту числа состояний, поддерживаемых элементами сети, по мере увеличения числа потоков. Эти проблемы описаны в [RFC2961], где предложены изменения и расширения RSVP для смягчения проблемы расширяемости и возможности применять RSVP в качестве универсального протокола сигнализации для Internet. Например, RSVP можно расширить с целью резервирования ресурсов для агрегатов потоков [RFC3175], создания явных MPLS LSP (см. параграф 5.1.3.3) и выполнения других сигнальных функций в Internet. В [RFC2961] также описан механизм для сокращения числа сообщений Refresh, требуемых для поддержки созданных сессий RSVP.

5.1.3.3. MPLS

MPLS - это схема пересылки, включая расширения традиционных протоколов плоскости управления IP. MPLS расширяет модель маршрутизации Internet и улучшает пересылку пакетов и управление путями [RFC3031].

На входе в домен MPLS маршрутизаторы LSR делят пакеты IP по классам эквивалентности пересылки (Forwarding Equivalence Class или FEC) на основе множества факторов, включая, например, сочетание сведений из заголовков IP в пакетах и локальные данные маршрутизации, поддерживаемые LSR. Затем в начало стека меток MPLS каждого пакета добавляется метка, соответствующая FEC. Запись стека меток MPLS имеет размер 32 бита и содержит 20-битовое поле метки.

LSR принимает решения о пересылке с использованием метки, добавленной в начало пакета, как индекса локальной записи о следующем узле пересылки по меткам (Next Hop Label Forwarding Entry или NHLFE). Затем пакет обрабатывается в соответствии с NHLFE. Входящая метка заменяется исходящей (смена меток - label swap) и пакет может пересылаться следующему LSR. Перед выходом пакета из домена MPLS его (верхняя) метка MPLS может быть удалена. LSP - путь между входным и выходным LSR, по которому проходит пакет с меткой. Путь явного LSP определяется узлом отправителя (входным) LSP. В MPLS для организации LSPможет применяться сигнальный протокол, такой как RSVP или протокол распространения меток (Label Distribution Protocol или LDP).

MPLS является мощной технологией для Internet TE за счёт поддержки явных LSP, позволяющих эффективно реализовать маршрутизацию на основе ограничений в сетях IP [AWD2]. Требования для TE в MPLS описаны в [RFC2702]. Расширения RSVP для поддержки организации явных LSP рассмотрены в [RFC3209] и параграфе 5.1.3.4.

5.1.3.4. RSVP-TE

RSVP-TE является расширением протокола RSVP (параграф 5.1.3.2) для организации трафика, спецификация расширения задана в [RFC3209]. RSVP-TE позволяет создавать MPLS LSP с организацией трафика (TE LSP), используя строгие или нестрогие пути и учитывая ограничения сети, такие как доступная пропускная способность. Расширение поддерживает сигнализацию LSP по явным путям, которые могут быть заданы административно или рассчитаны подходящим элементом (таким, как PCE, см. параграф 5.1.3.11) на основе требования QoS и правил с учётом превалирующего состояния сети, анонсированного расширением IGP для IS-IS [RFC5305], OSPFv2 [RFC3630], или OSPFv3 [RFC5329]. RSVP-TE позволяет резервировать ресурсы (например, пропускную способность) на пути.

RSVP-TE включает возможность вытеснять LSP на основе приоритета и использовать близость (affinity) каналов для их исключения или включения в LSP. Протокол дополнительно расширен для поддержки быстрой перемаршрутизации (Fast Reroute или FRR) [RFC4090], Diffserv [RFC4124] и двухсторонних LSP [RFC7551]. Расширения RSVP-TE для поддержки GMPLS (см. параграф 5.1.3.5) заданы в параграфе [RFC3473]. Требования к MPLS-TE LSP «один со многими» (point-to-multipoint или P2MP) заданы в [RFC4461], а расширения для организации P2MP MPLS-TE LSP через RSVP-TE - в [RFC4875]. P2MP LSP состоит из множества суб-LSP «от источника к листу» (source-to-leaf или S2L). Для определения путей P2MP LSP важен выбор точек ветвления (на основе возможностей, состояния сети и правил) [RFC5671]

Протокол RSVP-TE был расширен для предоставления в реальном масштабе времени динамических показателей для выбора пути с малой задержкой с использованием расширений IS-IS [RFC8570] и OSPF [RFC7471] на основе простого протокола активных двухсторонних измерений (Simple Two-Way Active Measurement Protocol или STAMP) [RFC8972] и протокола двухсторонних активных измерений (Two-Way Active Measurement Protocol или TWAMP) [RFC5357].

Использование RSVP-TE начиналось с каналов с ограниченной пропускной способностью, однако по мере расширения полосы протокол стал инструментом управления пропускной способностью для её эффективного использования и упреждающего управления ресурсами.

5.1.3.5. GMPLS

GMPLS расширяет протоколы управления MPLS охватывая технологии с разделением по времени (например, SONET/SDH[1], PDH[2], OTN[3]), длине волны (lambda) и пространственной коммутацией (например, входного порта или волокна в выходной порт или волокно) и продолжая поддерживать коммутацию пакетов. GMPLS предоставляет общий набор протоколов управления для всех этих уровней (включая расширения для некоторых технологий), каждый из которых имеет свою плоскость данных или пересылки. GMPLS охватывает сигнализацию и маршрутизацию в плоскости управления и базируется на расширениях TE для MPLS (см. параграф 5.1.3.4).

В GMPLS [RFC3945] исходная архитектура MPLS расширена для включения LSR с плоскостью пересылки на основе коммутации каналов (устройств), поэтому здесь не могут применяться для пересылки данных сведения из заголовков пакетов или ячеек. В частности, такие LSR включают устройства, где коммутация основана на временных интервалах (time slot), длине волны или физических порта. Эти дополнения влияют на базовые свойства LSP - способы запроса и передачи меток, одностороннюю природу MPLS LSP, способы распространения сведений об ошибках и информацию, предоставляемую для синхронизации входных и выходных LSR [RFC3473].

5.1.3.6. IPPM

Рабочая группа IETF по показателям производительности IP (IP Performance Metrics или IPPM) подготовила набор стандартных показателей, которые могут применяться для мониторинга качества, производительности и надёжности служб Internet. Эти показатель могут применять операторы сетей, конечные пользователи и независимые тестировщики для обеспечения пользователей и сервис-провайдеров общим пониманием в части производительности и надёжности облаков компонентов Internet [RFC2330]. Критерии для показателей производительности, разработанных IPPM, описаны в [RFC2330]. Примеры показателей производительности включают потери пакетов в одном направлении [RFC7680], задержку в одном направлении [RFC7679] и показатели связности между парой узлов [RFC2678]. Другие показатели включают измерения второго порядка для задержек и потери пакетов.

Некоторые из показателей производительности IPPM полезны при задании SLA, которые представляют собой наборы SLO, согласованные между пользователями и сервис-провайдерами, где каждая цель указывается сочетанием одного или нескольких показателей производительности, возможно, с заданием неких ограничений.

Рабочая группа IPPM также разрабатывает методы и протоколы измерений для определения показателей.

5.1.3.7. Измерение потоков

Рабочая группа IETF по измерению потоков в реальном масштабе времени (Real Time Flow Measurement или RTFM) разработала архитектуру, определяющую метод указания потоков трафика, а также ряд измерительных компонентов (измерители, считыватели и диспетчеры измерителей) [RFC2722]. Система измерения потоков позволяет проводить измерения и анализ на уровне отдельного потока сетевого трафика. Как отмечено в [RFC2722], система измерения потоков может быть очень полезна в ряде случаев:

- анализ поведения существующих сетей;
- планирование внедрения и расширения сетей;
- количественная оценка производительности сети;
- проверка качества сетевых услуг;
- идентификация работы пользователей с сетью.

Система измерения потоков включает измерители, а также считыватели и диспетчеры измерителей. Измеритель наблюдает за пакетами, проходящими через точку измерения, классифицирует их по группам, собирает сведения об использовании (например, число пакетов и байтов для каждой группы) и сохраняет их в таблице потоков. Группа может представлять любой набор пользовательских приложений, хостов, сетей и т. п. Считыватель извлекает сведения об использовании из различных измерителей для последующего анализа. Диспетчер отвечает за настройку и управление измерителями и считывателями. Инструкции, получаемые измерителем от диспетчера, включают спецификации потоков, параметры настройки измерителя и методы выборки. Инструкции, получаемые от диспетчера считывателем, включают адрес измерителя для сбора данных, частоту сбора и типы потоков для сбора сведений.

Спецификация экспорта данных о потоках IP (IP Flow Information Export или IPFIX) [RFC5470] определяет архитектуру, которая очень похожа на архитектуру RTFM[4] и включает процессы измерения, экспорта и сбора. В [RFC5472] рассмотрена применимость IPFIX и дано сравнение с RTFM, где указано, что архитектурно RTFM имеет дело с устройствами, а IPFIX - с процессами, для уточнения того, что на одной машине может присутствовать множество таких процессов. Протокол IPFIX [RFC7011] получил широкое распространение.

5.1.3.8. Контроль перегрузки конечных точек

В [RFC3124] описан набор механизмов контроля перегрузок для транспортных протоколов, позволяющих также создавать механизмы унификации контроля перегрузок для подмножества индивидуальных соединений (групп перегрузки). Диспетчер контроля перегрузок отслеживает состояние пути для каждой контролируемой им «группы

[1]Synchronous Optical Network / Synchronous Digital Hierarchy - синхронная оптическая сеть / синхронная цифровая иерархия.
[2]Plesiochronous Digital Hierarchy - плезиохронная цифровая иерархия.
[3]Optical Transport Network - оптическая транспортная сеть.
[4]Realtime Traffic Flow Measurement - измерение потоков трафика в реальном масштабе времени. *Прим. перев.*

перегрузок», используя полученные сведения для инструктирования планировщика в части распределения пропускной способности между соединениями этой группы. Концепции, описанные в [RFC3124], и уроки, которые можно извлечь из этой работы, нашли применение в HTTP/2 [RFC9113] и QUIC [RFC9000], а в [RFC9040] описана взаимозависимость блоков управления TCP, лежащая в основе работы диспетчера контроля перегрузок, описанного в [RFC3124].

5.1.3.9. Расширения TE для IGP

В [RFC5305] описаны расширения протокола взаимодействия промежуточных систем (Intermediate System to Intermediate System или IS-IS) для поддержки TE, в [RFC3630] - расширения TE OSPFv2, а в [RFC5329] - для OSPFv3. В IS-IS и OSPF применяется общая концепция расширений TE для распространения параметров TE, таких как тип и идентификатор канала, локальный и удалённый адрес IP, метрика TE, максимальная пропускная способность, максимальная резервируемая пропускная способность, незарезервированная пропускная способность и административная группа. Распространяемые IGP сведения могут служить для создания представления о состоянии и возможностях сети TE (см. параграф 5.1.3.14).

Различия между IS-IS и OSPF заключаются в деталях кодирования и передачи параметров TE.

- В IS-IS используются Extended IS Reachability TLV (тип 22), Extended IP Reachability TLV (тип 135) и Traffic Engineering router ID TLV (тип 134), где для передачи параметров TE служат суб-TLV, описанные в [RFC8570].
- В OSPFv2 используются Opaque LSA [RFC5250] типа 10, а в OSPFv3 - Intra-Area-TE-LSA. В обеих версиях OSPF применяются два TLV верхнего уровня (Router Address и Link TLV), использующие суб-TLV для передачи параметров TE ([RFC7471] для OSPFv2 и [RFC5329] для OSPFv3).

5.1.3.10. BGP - состояние канала

Во многих средах вызывается внешний по отношению к сети компонент для расчётов на основе топологии сети и текущего состояния соединений в ней, включая сведения TE. Эти данные обычно распространяются в пети протоколом IGP (см. параграф 5.1.3.9).

BGP (см. раздел 7) является одним из основных протоколов маршрутизации, объединяющим Internet. BGP-LS [RFC9552] - это механизм, с помощью которого можно собрать из сети сведения о состоянии каналов и TE, а также передать её внешним компонентам с помощью протокола маршрутзации BGP. Механизм подходит для физических и виртуальных каналов IGP и может управляться на основе правил. Сведения, собранные BGP-LS, можно использовать, например, для создания базы TED (параграф 5.1.3.14), используемой PCE (параграф 5.1.3.11), или на серверах ALTO (параграф 5.1.2.1).

5.1.3.11. Элемент расчёта пути

Расчёт путей на основе ограничений является важнейшей частью TE в сетях MPLS и GMPLS. Расчёт путей в больших многодоменных сетях сложен и может требовать специальных расчётных компонентов и кооперации между элементами разных доменов. PCE [RFC4655] - это элемент (компонент, приложение или узел сети), способный рассчитать путь или маршрут через сеть на основе графа сети и зададных для расчёта ограничений. PCE может служить центральным компонентом TE, работающей на основе базы TED (см. параграф 5.1.3.14), с передачей ему ответственности за расчёт путей в сетях MPLS, GMPLS или SR. PCE использует коммуникационный протокол PCEP [RFC5440] для взаимодействия с клиентами расчёта путей (Path Computation Client или PCC), такими как MPLS LSR, для ответов на их запросы расчёта путей или инструктирования их для инициирования новых путей [RFC8281] и поддержки состояния для путей, уже организованных в сети [RFC8231].

PCE являются ключевыми компонентами ряда систем TE. Дополнительные сведения о применимости PCE приведены в [RFC8051], а в [RFC6805] описано использование PCE для определения путей через несколько доменов. PCE могут также применяться в сетях ACTN (см. параграф 5.1.2.2), централированном управлении сетями (Centralized Network Control) [RFC8283] и SDN (см. параграф 4.3.2).

5.1.3.12. Маршрутизация по сегментам (SR)

Архитектура SR [RFC8402] использует парадигмы маршрутизации от источника и туннелирования. Путь передачи пакетов определяется на входе, а на выходе пакет туннелируется. В реализации протокола входной узел направляет пакет с использованием набора инструкций, называемых сегментами, которые включаются в добавляемый в начало пакета заголовок SR - стек меток в случае MPLS или последовательность 128-битовых SID для IPv6. Сегменты указываются идентификаторами SID. Имеется 4 типа SID, относящихся к TE.

- Prefix SID - уникальное в домене маршрутизации значение SID, служащее для идентификации префикса.
- Node SID - Prefix SID с установленным битом N для идентификации узла.
- Adjacency SID указывает смежность в одном направлении.
- Binding SID используется для двух целей:
  1. анонсирование сопоставлений префиксов с SID или метками;
  2. анонсирование пути, доступного для класса эквивалентной пересылки (FEC).

Сегмент может представлять любую инструкцию, основанную на топологии или услуге. SID можно искать в глобальном (домен) или ином конткексте (см., например, контекст меток в разделе 3 [RFC5331]).

Применение правил к SR может превратить SR в механизм TE, как описано в параграфе 5.1.1.3.

5.1.3.13. Построение дерева для явной репликации битового индекса

Явная репликация битового индекса (Bit Index Explicit Replication или BIER) [RFC8279] задаёт инкапсуляцию для групповой пересылки, которая может применяться в транспорте MPLS или Ethernet. Механизм построения дерева для репликации битового индекса (Tree Engineering for Bit Index Explicit Replication или BIER-TE) [RFC9262] представляет собой компонент, который может служить для построения групповой системы организации трафика. BIER-TE сам по себе не обеспечивает полной организации трафика и TE в данном случае с организацией трафика не связано.

В BIER-TE направление трафика обеспечивается путём задания строки битов, присоединяемой к каждому пакету для указания пересылки и репликации пакета а сети. Эта строка направляет трафик внутри сети и является элементом системы организации трафика. Центральный контроллер, которому известны возможности и состояние сети, а также потребности различных потоков трафика, способен выбирать пути групповой пересылки с учётом потребностей и доступных ресурсов, поэтому он отвечает за элменты политики организации трафика.

Управление ресурсами влияет на плоскость пересылки не только в части направления пакетов, заданного для BIER-TE. Оно включает выделение буферов для удовлетворения требований представленного трафика и может включать механизмы применения правил и/или формирования скорости с помощью различных форм очередей. Этот уровень управления ресурсов, хотя и не является обязательным, важен для сетей, которые хотят поддерживать правила контроля перегрузок для контроля или регулирования предлагаемого трафика, чтобы предоставлять различные уровни обслуживания и смягчать проблемы перегрузки. Она важен также для сетей, жалающий контролировать задержки, возникающие для конкретных потоков трафика.

5.1.3.14. Определение и представление состояния TE в сети

Состояния сети, относящиеся к TE, должны сохраняться в системе и представляться пользователю. База TED содержит набор сведений TE об узлах и каналах TE в сети. Она является важным компонентом таких систем TE, как MPLS-TE [RFC2702] и GMPLS [RFC3945]. Для формального определения данных TED и их представления пользователю можно применять язык моделирования данных YANG [RFC7950], как описано в [RFC8795].

5.1.3.15. Интерфейсы управления системой

Система управления TE должна иметь удобный для человека интерфейс, программируемый для оптимизации. Протокол настройки сети (Network Configuration Protocol или NETCONF) [RFC6241] и протокол RESTCONF [RFC8040] обеспечивают программируемые интерфейсы, удобные для человека. Эти протоколы используют сообщения в формате XML или JSON. Если для оптимизации интерфейса управления или экономии пропускной способности нужно сжимать сообщения, можно воспользоваться протоколом грруповых коммуникаций для приложений с ограничениями (Constrained Application Protocol или CoAP) [RFC7390] или gRPC [GRPC], особенно при кодировании сообщений в двоичном формате. Вместе с любым из этих протоколов можно использовать язык моделирования данных YANG [RFC7950] для формального и точного описания данных интерфейса.

Другим вариантом является протокол PCEP [RFC5440], разработанный как вариант интерфейса управления системой TE. Сообщения PCEP передаются в формате TLV и не определяются языком моделирования данных, таким как YANG.

5.2. Распространение содержимого

В Internet доминируют взаимодействия клиент-сервер, особенно для web-трафика и мультимедийных потоков, хотя в будущем могут доминировать более сложные медиа-серверы. Местоположение и производительность основных информационных серверов оказывают существенное влияние на картину трафика в Internet, а также восприятие качества обслуживания пользователями.

Было разработано множество методов динамического распределения нагрузки для повышения производительности реплицируемых информационных серверов. Эти методы могут привести к изменению пространственных характеристик трафика Internet, делая их более динамичными, поскольку эти серверы могут выбираться динамически на основе местоположения клиентов и серверов, относительной загрузки серверов, а также относительной производительности разных частей сети. Этот процесс связывания распределенных серверов с клиентами называется направлением трафика (traffic directing) и является функцией прикладного уровня.

Для более эффективного выбора серверов, размещённых в географически удалённых местах, для работы с клиентами может потребоваться эмпирическая статистика производительности сети. В будущем могут потребоваться системы измерений, предоставляющие такую статистику.

При возникновении перегрузок в сети системам направления и организации трафика следует действовать согласованно. Этот вопрос требует изучения.

Вопросы, связанные с местоположением и репликацией информационных серверов, в частности, web-серверов, важны для организации трафика Internet, поскольку на такие серверы приходится значительная часть трайика Internet.

6. Рекомендации по организации трафика Internet

В этом разделе приведены в общем виде высокоуровневые рекомендации по организации трафика Internet, описывающие средства, требуемые для решения задач и достижения целей TE. В широком смысле рекомендации можно разделить на функциональные и нефункциональные.

- Функциональные рекомендации описывают функции, которые следует поддерживать системам организации трафика. Эти функции нужны для достижения целей TE по решению задач организации трафика.
- Нефункциональные рекомендации связаны с атрибутами качества и характеристик состояния систем TE и могут содержать противоречивые утверждения, а иногда их трудно оценить.

6.1. Базовые нефункциональные рекомендации

Ниже приведены базовые нефункциональные рекомендации по организации трафика Internet. В данном контексте некоторые рекомендации могут быть очень важны, а другие могут быть необязательными. Поэтому на этапе разработки системы TE для работы к конкретном контексте может потребоваться задание приоритетов.

Автоматизация

ПО возможности системе TE следует автоматизировать как можно больше функций TE для сокращения участия человека в анализе сети и управлении её работой. Автоматизация особенно важна в больших сетях общего пользования, поскольку с человеческим фактором эксплуатации сетей связаны значительные расходы и риски вызываемых людьми проблем. Автоматизация может выигрывать от обратной связи с сетью, показывающей

состояние ресурсов сети и текущую загрузку. Кроме того, интеллектуальные функции компонентов TE могут сделать автоматизацию более динамичной и отзывчивой к изменениям в сети.

Гибкость

Системе TE следует разрешать изменение политики оптимизации. В частности, следует обеспечивать опции конфигурации для приспособления системы к конкретной среде. Может быть желательно наличие автономной и интерактивной подсистем TE, которые можно включать и отключать независимо. Системам TE, применяемым в сетях с несколькими классами, следует иметь опции для поддержки оценки и оптимизации по классам.

Функциональная совместимость

По возможности системы TE и их компоненты следует разрабатывать с интерфейсами на основе открытых стандартов для функциональной совместимости с другими системами и компонентами.

Расширяемость

Сети общего пользователя продолжают быстро расти в плане размеров и объёма трафика. Поэтому для сохранения применимости по мере развития сети сестеме TE следует быть расширяемой. В частности, системе TE следует сохранять функциональность по мере роста числа маршрутизаторов и каналов, а также числа потоков и объёма трафика. Системе TE следует иметь расширяемую архитектуру, не оказывать негативного влияния на другие функции и процессы в элементах сети, а также не потреблюять слишком много ресурсов сети при сборе и распространении информации, а также при управлении.

Безопасность

Безопасность критически важна для систем TE. Такие системы обычно осуществляют контроль функциональных аспектов сети для достижения желаемой производительности. Поэтому должны приниматься адекватные меры защиты целостности системы TE, а также адекватные меры защиты сети от уязвимостей, возникающих из-за брешей в безопасности и других нарушения в системе TE.

Простота

Системе TE следует быть максимально простой. Простота пользовательского интерфейса не обязательно означает примерение в системе TE наивных алгоритмов. При использовании сложных алгоритмов и внутренних структур пользовательскому интерфейсу следует максимально скрывать сложности от администраторов сети.

Стабильность

По стабильностью понимается устойчивость сети к осцилляциям (переходам) состояния в разрушительной манере, которые могут приводить к перенаправлению трафика без удовлетворительного разрешения базовых проблем TE, вызывающих безостановочные смены состояния. Стабильность очень важна в системах TE, реагирующих на смену состояний сети. Зависящие от состояния методологии TE обычно включают компромисс между оперативностью и стабильностью. Настоятельно рекомендуется при таких компромиссах отдавать предпочтение стабильности (особенно в магистральных IP-сетях общего пользования).

Удобство использования

Удобство использования - это связанный с людьми аспект систем TE. Это включает простоту внедрения и эксплуатации систем TE. В общем случае желательно иметь систему TE, которую легко развернуть в имеющейся сети. Желательна также простота эксплуатации и обслуживания системы TE.

Видимость

В систему TE следует включать механизмы сбора статистики из сети и анализа статистических данных для определения эффективности работы сети. Производная статистика (матрицы трафика, загрузка каналов, задержки и потери пакетов, а также другие показатели производительности), определяемая из сетевых измерений, может служить индикатором преобладающих в сети условий. Возможности различных компонентов системы маршрутизации являются другим примром данных состояния, которые следует делать доступны для наблюдения.

6.2. Рекомендации по маршрутизации

Управление маршрутизацией является важным аспектом организации трафика Internet. Маршрутизация влияет на многие ключевые показатели производительности, связанные с сетью, такие как пропускная способность, задержки и загрузка сети. В общем случае очень сложно обеспечить высокое качество обслуживания в распределенной сети без эффективного управления маршрутизацией. Желаемой системой маршрутизации для TE будет та, которая учитывает характеристики трафика и ограничения сети, обеспечивая при этом стабильность.

Протоколы IGP, основанные на алгоритмах поиска кратчайших путей (Shortest Path First или SPF), имеют ограниченные возможности управления TE [RFC2702] [AWD2]. Эти ограничения кратко описан ниже.

1. Чистые протоколы SPF не учитывают ограничения сети и характеристики трафика при выборе маршрутов. Например, IGP всегда выбирают кратчайшие пути на основе метрики каналов, заданно администраторами, поэтому распределение нагрузки по неравноценным путям невозможно. Отметим, что метрика каналов назначается на основе выбранных оператором правил, которые могут включать предпотение одних каналов над другими, следовательно, «кратчайший» путь может не быть мерой дальности. Использование кратчайших путей для пересылки трафика может вызывать ряд проблем, указанных ниже.

   - Если трафик от источника к получателю превышает возможности канала на кратчайшем пути, канал (и кратчайший путь) перегружается, а более длинные пути между этими узлами загружаются недостаточно.
   - Кратчайшие пути от разных источников могут накладываться один на другой на некоторых каналах. Если суммарных трафик превышает возможности этих каналов, возникает перегрузка.
   - Проблемы могут также возникать в результате изменения потребностей в трафике с течением времени, если топология сети и конфигурация маршрутизации своевременно не изменяются. Это делает топологию сети и конфигурацию маршрутизации неоптимальной, что может вызывать сохраняющуюся перегрузку.

2. Поддержка нескольких равноценных путей (Equal-Cost Multipath или ECMP) в SPF IGP позволяет распределять трафик по ряду путей с одинаковой стоимостью. Однако ECMP пытается равномерно распределить трафик между равноценными кратчайшими путями. В общем случае ECMP не поддерживает настраиваемое распределение нагрузки между равноуенными путями. В результате трафик на одном из путей может оказаться значительно выше из-за передачи по нему трафика из других источников и может вызывать перегрузку на этом пути. Некоторое смягмение этой проблемы обеспечивает взвешенный ECMP (Weighted ECMP или WECMP, см., например, [EVPN-UNEQUAL-LB]).

3. Смена метрики IGP для управления трафиком обычно оказывает влияние на всю сеть, что может приводить к непредвиденным и нежелательным эффектам. Описанная в разделе 8 работа [FT00] [FT01] может улучшить контроль.

С учётом ограничений нужны средства улучшения функций маршрутизации в сетях IP. Некоторые из таких возможностей указаны ниже.

- Маршрутизация на основе ограничений может быть полезна на общедоступных магистралях IP со сложной топологией. Ограничения могут включать пропускную способность, число пересылок и административные средства, такие как атрибуты классов ресурсов [RFC2702] [RFC2386]. Это позволяет выбирать маршруты, соответствующие заданному набору требований. Маршруты, рассчитанные с учётом ограничений, не обязательно будут кратчайшими. Маршрутизация на основе ограничений лучше всего работает с ориетированными на пути технологиями, поддерживающими явные маршруты, такими как MPLS.

- Машрутизацию на основе правил можно также применять для распределения трафика по инфраструктуре (в том числе трафика best-effort). Например, проблем, связанных с перегрузками из-за неравномерного распределения трафика, можно избежать (или смягчить их), зная атрибуты возможности резервирования пропускной способности каналов сети и задавая требования к пропускной способности при выборе пути.

- Ряд усовершенствований IGP на основе состояния каналов позволяет этим протоколам распространять дополнительные сведения о состоянии каналов, требуемые для маршрутизации на основе ограничений. Расширения для OSPF описаны в [RFC3630], для IS-IS - в [RFC5305]. Некоторые добавочные сведения о состоянии топологии включают атрибуты каналов, такие как доступная для резервирования пропускная способность и атрибуты классов ресурсов (задаваемые административно свойства каналов). Концепция атрибутов класса ресурсов введена [RFC2702]. Дополнительные сведения о топологии передаются в новых TLV и суб-TLV для IS-IS [RFC5305] и в Opaque LSA для OSPF [RFC3630].

- IGP с расширенными сведениями о состоянии каналов могут чаще рассылать лавинные данные, нежели обычный IGP. Это связано с тем, что даже при отсутствии изменений в топологии изменение доступой для резервирования полосы или близости каналов могут инициировать лавинную рассылку расширенным протоколом IGP. Компромисс между своевременностью и объёмом лавинной рассылки обычно достигается с использованием порога, основанного на доле (в процентах) изменений в анонсируемых ресурсах, чтобы избежать излишнего расхода пропускной способности и расчётных ресурсов, а также нестабильности TED.

- В системах TE желательно, чтобы подсистема маршрутизации позволяла настраивать долю трафика для отдельных путей (с одинаковой или разной стоимостью). Это позволит администраторам сетей более гибко контролировать распределение трафика в сетях и может быть очень полезно в определённых ситуациях для предотвращения и смягчения перегрузок. Примеры этого представлены в [XIAO] и [EVPN-UNEQUAL-LB].

- Системе маршрутизации следует также обеспечивать возможность контроля маршрутов для подмножеств трафика без влияния на маршруты другого трафика при наличии достаточных ресурсов. Это позволит более тонко управлять распределением трафика по сети. Например, возможность переноса трафика с одного пути на другой (не воздействуя на остальные пути трафика) позволяет перенаправить трафик по пути с достаточными ресурсами. Ориентированные на пути технологии, такие как MPLS-TE, по своей природе поддерживают такую возможность, как описано в [AWD2].

- Подсистеме маршрутизации следует поддерживать возможность выбора разных путей для разных классов трафика (или агрегатов поведения), если сеть поддерживает несколько классов обслуживания (разных агрегатов поведения).

6.3. Рекомендации по распределению трафика

Распределением трафика называется его направление на (заранее созданные) пути, удовлетворяющие некоторым требованиям. Основанная на ограничениях маршрутизация имеет дело с выбором путей, а распределение трафика - с его направлением по путям, созданным такой маршрутизацией или иными способами. Распределение трафика может выполняться в зависимости от времени или состояния, как описано в параграфе 4.1.

Двумя важными аспектами функции распределения трафика являются способность организовать несколько путей между источником и получателем, а также направления трафика по этим путям в соответствии с заданными правилами. Для этой схемы требуются гибкие механизмы разделения трафика и последующего указания для каждой его части одного из параллельных путей (параллельные транки трафика в [RFC2702]). При распределении трафика по нескольким параллельным путям рекомендуется уделить особое внимание корректному порядку пакетов, относящихся к одному приложению (потоку трафика), на целевом узле параллельных путей.

Механизмам, выполняющим функции распределения трафика, следует стремиться распределять трафик по инфраструктуре сети для минимизации перегрузок. Если суммарный трафик невозможно распределить или функции маршрутизации и отображения не могут должным образом реагировать на изменение условий для трафика, система распределения может применять механизмы краткосрочного контроля перегрузок (управление очередями, планирование и т. п.) для их смягчения. Таким образом, механизмы, выполняющие функции распределени трафика, дополняют имеющиеся механизмы контроля перегрузок. В рабочей сети трафик следует распределять по инфраструктуре сети так, чтобы минимизировать соперничество за ресурсы внутри класса и между классами (см. раздел 2).

При зависимости методов распределения трафика от динамической обратной связи, например, адаптивной организации трафика MPLS (MPLS Adaptive Traffic Engineering или MATE) [MATE], требуется уделять особое внимание стабильности сети.

6.4. Рекомендации по измерениям

Важность измерений в TE обсуждается на протяжении всего этого документа. В систему TE следует включать механизмы измерения и сбора статистики в сети для поддержки работы TE. Для анализа статистики могут потребоваться дополнительные средства. Этим механизмам следует не оказывать негативного воздействия на

точность и целостность собранной статистики. Механизмам сбора статистики следует поддерживать расширяемость по мере роста сети.

Статистику трафика можно разделить на кратковременную (моментальную) и долгосрочную. Долгосрочная статистика очень полезна для организации трафика и может периодически регистрировать загрузку сети (например, почасовые, суточные и недельные изменения профилей трафика), а также тенденции трафика. Аспекты статистики трафика могут описывать характеристики классов обслуживания для сетей, поддерживающих разные классы. Анализ долгосрочной статистики трафика может давать такие сведения, как характеристики в часы максимальной нагрузки, модели роста трафика, сохраняющиеся перегрузки, «горячие точки», дисбаланс загрузки каналов, связанный с аномалиями маршрутизации.

Следует предусматривать механизм построения матриц трафика для кратковременной и долгосрочной статистики. В мультисервисных сетях IP матрицы трафика могут строиться по классам обслуживания. Каждый элемент матрицы трафика представляет статистику потоков трафика между парами абстрактных узлов, которые могут соответствовать маршрутизаторам, группам маршрутизаторов или сайтам в VPN.

В статистике трафика следует предоставлять разумные и надёжные показатели текущего состояния сети в краткосрочном масштабе. Некоторые из таких показателей могут отражать загрузку каналов и состояния перегрузки в сети. Примеры индикаторов перегрузки включают чрезмерную задержку пакетов, потери и высокую загрузку ресурсов. Примеры механизмов распространения таких сведений включают SNMP, инструменты зондирования, FTP, анонсы состояния каналов IGP, NETCONF, RESTCONF и т. п.

6.5. Политика, планирование и контроль доступа

Рекомендации параграфов 6.2 и 6.3 могут быть неоптимальными или неэффективными, если объем трафика по маршруту или пути превышает возможности ресурсов на этом маршруте или пути. Для повышения производительности систем TE может применяться несколько подходов.

- Основополагающим подходом является та или иная форма планирования, где трафик распределяется по маршрутам с учётом доступных на них ресурсов. Планирование может быть централизованным или распределенным и должно учитывать объёмы трафика и доступные ресурсы. Однако такой подход имеет смысл лишь при соответствии трафика запланированным объёмам.
- Потоки трафика могут контролироваться на границах сети. Это простой способ обеспечить соответствие между запланированным и фактическим трафиком. Применяется та или иная форма измерений (см. параграф 6.4) для определения скорости поступления трафика, а избыточный трафик может отбрасываться или пересылаться через сеть как получится (best-effort). Этот подход эффективен лишь при строгом планировании в масштабе всей сети и жёстком подходе к избыточному трафику.
- Контроль допуска к ресурсам - это процесс, в котором узлы сети принимают решение о предоставлении доступа к своим ресурсам на уровне пакетов в соответствии с потоком, к которому пакет относится. Сведения о принадлежности применяются правилами, заданными локально или установленными через систему управления или плоскость управления. В итоге пакет получает доступ к конкретным ресурсам лишь в случае соответствия потока, к которому он относится, заданным правилам.

Для построения эффективной системы TE рекомендуется применять сочетание указанных подходов.

6.6. Живучесть сети

Под живучестью понимается способность сети поддерживать непрерывность обслуживания при возникновении отказов. Это может быть достигнуто путём быстрого восстановления после сбоев и поддержкой после восстановления требуемого уровня QoS для имеющихся служб. Живучесть вызывает серьёзную озабоченность сообщества Internet в связи с необходимостью передачи через Internet критически важного и высокоприоритетного трафика, а также трафика в реальном масштабе времени. Проблему живучести можно решать на уровне устройств путём разработки более надёжных элементов сети или на уровне сети за счёт включения избыточности в архитектуру, проектирование и эксплуатацию сетей. Рекомендуется использовать отказоустойчивость и живучесть в архитектуре, проектировании и эксплуатации систем TE, применяемых для управления сетями IP (особенно общедоступными). Поскольку требования к живучести могут зависеть от контекста, поддержку живучести следует делать гибкой, чтобы её можно было приспособить к разным потребностям. Для обеспечения живучести сетей разработан ряд методов и инструментов, включая быструю перемаршрутизацию MPLS (Fast Reroute) [RFC4090], независимую от топологии быструю перемаршрутизацию по дополнительным путям без петель (Topology Independent Loop-free Alternate Fast Reroute for Segment Routing) [SR-TI-LFA], расширения RSVP-TE для поддержки сквозного [RFC4872] и посегментного [RFC4873] восстановления GMPLS.

Влияние перебоев в обслуживании существенно меняется для разных классов обслуживания в зависимости от продолжительности прерывания, которая может варьироваться от миллисекунд (незначительное влияние на услуги) до секунд (возможное прерывание сессий IP-телефонии и тайм-ауты в основанных на соединениях транзакциях) и даже минут или часов (возможно со значительными социальными и бизнес-последствиями). Перебои разной продолжительности оказывают различное влияние в зависимости от характера прерываемых потоков трафика.

Возможности защиты и восстановления доступны на разных уровнях, поскольку сетевые технологии продолжают развиваться. Оптические сети способны обеспечивать динамическое восстановление в кольцах и многосвязных (mesh) соединениях на уровне длин волн. На уровне SONET/SDH свойства живучести обеспечиваются автоматическим защитным переключением (Automatic Protection Switching или APS), а также самовосстановлением колец и mesh-соединений. Похожая функциональность обеспечивается технологиями канального уровня, такими как Ethernet.

На уровне IP используется перемаршрутизация для восстановления обслуживания после отключения каналов и узлов. Перемаршрутизация на уровне IP происходит после схождения маршрутов, для чего могут потребоваться секунды и даже минуты. Для повышения живучести сетей IP экономически эффективным способом можно применять основанные на путях технологии, такие как MPLS [RFC3469].

Важным аспектом живучести на разных уровнях является то, что технологии этих уровней способны обеспечивать защиту и восстановление в разным уровнем детализации в плане временных масштабов и пропускной способности (от уровня пакетов до уровня длин волн). Возможности защиты и восстановления могут зависеть от класса обслуживания и моделей работы сетей. Координация возможностей защиты и восстановления на разных уровнях для обеспечения живучести сети за разумную цену является сложной задачей и возможна не всегда, поскольку сети на разных уровнях могут относиться к разным административным доменам. Ниже приведены некоторые базовые рекомендации по координации защиты и восстановления.

- Возможности защиты и восстановления на разных уровнях следует координировать так, чтобы обеспечивалась гибкая и недорогая живучесть сети. Одним из способов является предотвращение дублирования функций на разных уровнях. Передача аварийных сигналов и иных индикаторов сбоев с нижних уровней на верхние также может выполняться скоординированно. Ещё одним способом координации защиты и восстановления на разных уровнях является обеспечение временного порядка триггеров разных уровней.
- Пропускная способность, зарезервированная на одном уровне для защиты и восстановления, недоступна (и не видна) на вышележащем уровне в качестве резервной. Применение функций защиты и восстановления на нескольких уровнях может повышать уровень резервирования и отказоустойчивости, однако это может снижать эффективность использования ресурсов сети. Нужно тщательное планирование для обеспечения компромисса между живучестью и оптимальным использованием ресурсов.
- В общем случае желательно иметь схемы защиты и восстановления эффективные по своей природе в плане использования пропускной способности.
- Уведомлениям об отказах в сети следует быть надёжными и своевременными, если они служат триггерами операций защиты и восстановления.
- Следует обеспечивать сигналы тревоги и другие средства мониторинга и оповещения на соответствующих уровнях сети для выполнения на этих уровнях действий по защите и восстановлению.

6.6.1. Живучесть сетей на основе MPLS

Поскольку технология MPLS ориентирована на пути, она потенциально может обеспечивать возможности более быстрой и предсказуемой защиты и восстановления по сравнению с традиционными системами поэтапной (hop-by-hop) маршрутизации IP. Типы защиты в сетях MPLS можно разделить на 4 категории.

Защита канала

Целью защиты канала является защита LSP от сбоев на данном канале. В этом случае резервный (вторичный) LSP проходит по пути, не связанному с рабочим (первичным) LSP на канале, где нужна защита. При отказе защищаемого канала трафик рабочего LSP переносится на защитный LSP в головном устройстве отказавшего канала. Как локальный метод ремонта, защита канала может быть быстрой. Эта форма защиты может быть наиболее подходящей в ситуациях, где некоторые элементы данного пути заведомо менее надёжны, чем другие.

Защита узла

Целью защиты узла является защита LSP от сбоев на данном узле. В этом случае резервный (вторичный) LSP проходит по пути, не связанному с рабочим (первичным) LSP на узле, где нужна защита. Вторичный LSP также развязан с первичным на всех каналах, подключённых к защищаемому узлу. При отказе защищаемого канала трафик рабочего LSP переносится на защитный LSP в LSP восходящего направления, напрямую соединённом с отказавшим узлом. Защита узла охватывает большую часть сети нежели защита канала, но в остальном принципиально не отличается.

Защита пути

Целью защиты пути LSP (сквозная защита) является защита LSP от любых отказов на маршрутизируемом пути. В этом случае защитный LSP полностью развязан с рабочим LSP. Преимуществом этого подходя является защита рабочего LSP резервным от всех возможных отказов узлов и каналов на пути, кроме отказов входного или выходного LSR. Кроме того, защита пути может быть более эффективной в плане использования ресурсов по сравнению с защитой каналов и узлов, применяемой на каждом этапе пути. Однако защита пути может быть медленней защиты каналов и узлов, поскольку для неё требуется распространение сведений об отказах.

Защита сегмента

Домен MPLS может быть разделен на субдомены (домены защиты). Применяется защита пути для каждого LSP от входа в домен до выхода из него. Когда LSP проходит через несколько доменов защиты, механизму защиты внутри домена требуется защищать лишь проходящий внутри домена сегмент LSP. Защита сегментов обычно быстрее сквозной защиты, поскольку восстановление как правило происходит ближе к месту отказа и не требуется дальнего распространения уведомлений об отказе.

Более полное описание восстановления на основе MPLS представлено в [RFC3469] и [RFC6372].

6.6.2. Варианты защиты

Ещё одним вопросом, требующим рассмотрения, является концепция опций защиты. Здесь используются обозначения m:n, где m указывает число защитных LSP, применяемых для защиты n рабочих LSP. Во всех вариантах, кроме защиты 1+1, связанные с защитными LSP ресурсы могут использоваться для трафика best-effort, когда на рабочем LSP не наблюдается проблем.

1:1

Один рабочий LSP защищается (восстанавливается) одним защитным LSP. Трафик передаётся по защищаемому LSP пока событие защиты (восстановления) не перенесёт трафик в защитный LSP.

1:n

Один защитный LSP применяется для защиты (восстановления) n рабочих LSP. Трафик передаётся через n защищаемых рабочий LSP, пока событие защиты (восстановления) не перенесёт трафик отказавшего LSP в защитный. В каждый момент может быть восстановлен лишь один отказавший LSP.

n:1

Один рабочий LSP защищается (восстанавливается) с помощью n защитных LSP, возможно, с распределением нагрузки между ними. Это может быть особенно полезно в случаях невозможности найти один путь, обеспечивающий требования к пропускной способности основного LSP.

1+1

Трафик передаётся одновременно по рабочему и защитному LSP. Выходной маршрутизатор LSR выбирает один или два LSP по локальным правилам (обычно на основе обеспечения целостности). При нарушении трафика (сбой) на одном LSP выходной маршрутизатор переносит весь трафик на другой LSP. Этот подход ведёт к значительному расходу ресурсов сети, но обеспечивает более быстрое восстановление.

6.7. Многоуровневая организация трафика

Сети часто организуются по уровням. Отношения между уровнями могут представлять взаимодействие технологий (например, сеть IP на основе оптической сети) или отношения между операторами (например, сеть клиента, работающая на основе сети провайдера). Отметим, что многоуровневая сеть не требует применения разных технологий, хотя в таких сетях часто применяется та или иная форма инкапсуляции.

Многоуровневая организация трафика сопряжена с рядом проблем, связанных с расширяемостью и конфиденциальностью. Эти вопросы рассмотрены в [RFC7926], где обсуждается обобществление сведений между доменами по правилам фильтрации, агрегирования, абстрагирования и виртуализации. В этом документе также обсуждается, какие из имеющихся протоколов могут поддерживать такой сценарий, с особым упором на BGP-LS (см. параграф 5.1.3.10).

PCE (параграф 5.1.3.11) также является полезным инструментом для многоуровневых сетей, как описано в [RFC6805], [RFC8685] и [RFC5623]. Методы сигнализации для многоуровневой организации трафика описаны в [RFC6107].

Живучесть многоуровневых сетей рассмотрена в параграфе 6.6.

6.8. Организация трафика в средах Diffserv

Растущие требования к поддержке множества классов трафика в Internet (например, best-effort и критически важные данные) требуют от сетей IP разделять трафик по некоторым критериям и обеспечивать преимущественную обработку определенным типам трафика. Большое число потоков может объединяться в несколько поведенческих агрегатов на основе критериев, связанных с базовыми требованиями к производительности в плане потери пакетов, задержек и их вариаций, а также с базовыми полями заголовков в пакетах IP.

Дифференцированное обслуживание (Diffserv) [RFC2475] может применяться для соблюдения соглашений SLA, заданных для разделения потоков трафика. Классы обслуживания могут поддерживаться в среде за счёт конкатенации поведения на этапе пересылки (Per-Hop Behavior или PHB) на пути маршрутизации. PHB задаёт поведение пересылки пакетов на поддерживающих Diffserv узлах и может настраиваться на каждом маршрутизаторе. PHB обеспечивается с помощью механизмов управления буферами и планирования, требуя выполнять на входном узле классификацию и маркировку трафика, а также применение правил и формовку (shaping).

TE может дополнять Diffserv для улучшения использования ресурсов сети. TE может работать сразу для всех классов обслуживания [RFC3270] или по отдельным классам. Первый вариант применяется для лучшего распределения нагрузки между ресурсами сети (описание механизмов для этого приведено в [RFC3270]). Второй вариант рассматривается ниже, поскольку он специфичен для среды Diffserv (организация трафика с поддержкой Diffserv [RFC4124]).

Для некоторых сетей Diffserv может оказаться желательным контроль производительности для некоторых классов обслуживания за счёт связывания числа выделяемых для класса обслуживания ресурсов сети с объёмом трафика данного класса. Такая связь выделения ресурсов с потребностями может быть обеспечения с помощью комбинации нескольких механизмов, например,

- механизмы TE по классам обслуживания, связывающие объем трафика данного класса с выделяемыми этому классу ресурсами;
- механизмы, динамически регулирующие выделенные данному классу ресурсы в соответствии с объёмом трафика этого класса.

Может оказаться желательным ограничение влияния высокоприоритетного трафика на производительность трафика с более низким приоритетом. Это может быть достигнуто, например, путём контроля доли высокоприоритетного трафика, маршрутизируемого через данный канал. Другим вариантом является соответствующее повышение пропускной способности канала, чтобы трафик с низким приоритетом по-прежнему получал надлежащее обслуживание. Когда доли трафика с разными классами обслуживания существенно меняются от маршрутизатора к маршрутизатору, традиционных протоколов маршрутизации IGP и механизмов TE, не различающих классы обслуживания, может оказаться недостаточно. Взамен может быть желательно выполнение TE по классам обслуживания, особенно для функций управления маршрутизацией и отображения. Одним из вариантов реализации этого в домене с поддержкой MPLS и Diffserv является задание LSP по классам обслуживания и отображение трафика каждого класса на один или несколько таких LSP. Для данного класса обслуживания LSP может маршрутизироваться с поддержкой защиты (восстановления) в зависимости от класса по соответствующим правилам.

Для организации трафика по классам обслуживания может потребоваться распространение параметров каждого класса. Обычно некоторые классы имеют те или иные общие совокупные ограничения (например, требования к максимальной пропускной способности), которые не применяются к отдельному классу. Такие классы можно группировать группировать в типы и распространять параметры для типов (а не отдельных классов), что улучшит расширяемость механизмов. Это также позволяет более эффективно использовать пропускную способность для разных классов в рамках одного типа. Тип класса — это набор классов, удовлетворяющих указанным ниже условиям.

- Классы одного типа имеют общие совокупные требования к обеспечению уровня производительности.
- На уровне отдельного класса в рамках одного типа не предъявляется никаких требований. Тем не менее, сохраняется возможность реализовать некие правила приоритизации, чтобы обеспечить преимущественный доступ к пропускной способности путём использования приоритетов вытеснений.

Подробное описание требований к TE с поддержкой Diffserv приведено в [RFC4124].

6.9. Управляемость сети

Соображения, приведённые в параграфе 4.2 для автономной (offline) и интерактивной организации трафика, будут ограниченно полезны, если сетью невозможно эффективно управлять для реализации решений TE и достижения желаемых целей в плане производительности сети.

Добавление пропускной способности является грубым решением проблем TE. Однако оно является простым, может применяться путём создания параллельных каналов, формирующих часть схемы ECMP, и могут обеспечивать преимуществ, если добавление пропускной способности дёшево и несложно. Однако недорогое повышение пропускной способности возможно не всегда и такой подход может оказаться не лучшим. Настройка административных весов и других параметров, связанных с протоколами маршрутизации, обеспечивает более тонкий контроль, но этот подход сложнее и менее точен из-за взаимодействия протоколов маршрутизации в сети.

Механизмы управления могут быть ручными (например, статическая конфигурация), полуавтоматическими (например, сценарии) или полностью автоматизированными (например, системы управления на основе правил). Автоматизированные механизмы особенно полезны в больших сетях. Взаимодействию оборудования разных производителей могут способствовать стандартизованные средства управления (например, модели YANG) для поддержки управляющих функций, требуемых для достижения целей TE.

Функциям управления сетью следует быть защищёнными, надёжными и стабильными, поскольку они часто требуются для обеспечения корректной работы сети в случае отказов (например, при перегрузке или атаках).

7. Междоменное взаимодействие

Междоменная организация трафика оптимизирует производительность для трафика, исходящего из одного административного домена и направленного в другой домен.

BGP [RFC4271] является стандартным протоколом внешних шлюзов, используемым для обмена маршрутными сведениями между автономными системами (AS) в Internet. BGP включает процесс решения, определяющий предпочтения для маршрутов в данную целевую сеть. Два основных аспекта междоменной организации трафика с использованием BGP указаны ниже.

Распространение маршрутов
: Управление импортом и экспортом маршрутов между AS, а также передачей маршрутов между BGP и другими протоколами внутри AS.

Выбор лучшего пути
: Выбор лучшего из имеющихся вариантов пути в данную целевую сеть. Этот выбор осуществляет процесс решения BGP, который определяет точки выхода из данной AS в направлении конкретных целевых сетей с учётом различных факторов и соображений. На процесс выбора пути BGP могут влиять с помощью атрибутов, таких как NEXT_HOP, LOCAL_PREF, AS_PATH, ORIGIN, MULTI_EXIT_DISC (MED), метрика IGP и т. п.

Большинство реализаций BGP предоставляет конструкции, упрощающие организацию сложных правил BGP на основе заданных заранее логических условий. Они могут служить для управления импортом и экспортом входящих и исходящих маршрутов, обмена маршрутами между BGP и другими протоколами, а также влияния на выбор лучшего пути путём манипулирования атрибутами (стандартизованными или фирменными), связанными с процессом решения BGP.

При рассмотрении междоменной TE с BGP следует учитывать, что точка выхода трафика является управляемой, тогда как соединительная точка, через которую входит трафик, обычно не управляется. Поэтому каждая отдельная сеть реализует стратегию TE, нацеленную на эффективную доставку трафика, исходящего от её клиентов, к точкам партнёрства (peering). Большинство правил TE основано на стратегии «ближайшего выхода», где междоменный трафик «выгружается» в точку выхода, ближайшую в направлении целевой AS. Большинство методов манипулирования точкой входа трафика неэффективно или неприемлемо в сообществе партнёров.

Междоменная организация трафика с BGP в целом эффективна, но обычно применяется методом проб и ошибок, поскольку система TE, как правило, знает о доступных сетевых ресурсах лишь в рамках одного домена (AS в данном случае). Для систематического подхода к междоменной TE требуется кооперация доменов. Хорошее для одного домена решение не обязательно будет хорошим и в другом домене. Кроме того, обычно нежелательно разрешать процессу управления одного домена влиять на маршрутизацию и управление трафиком другого домена в его сети.

Туннели MPLS-TE (LSP) могут повышать гибкость выбора точек выхода для междоменной маршрутизации за счёт применения концепции абсолютной и относительной метрики. Если атрибуты BGP определены так, что процесс решения BGP зависит от метрики IGP при выборе точек выхода для междоменного трафика, для некого междоменного трафика, адресованного в данную партнерскую сеть, можно сделать конкретную точку выхода предпочтительной путём организации туннеля TE между маршрутизатором, делающим выбор, и его партнёром с назначением туннелю TE метрики, которая меньше стоимости IGP для всех других точек партнёрства. Расширения протокола RSVP-TE для междоменных систем MPLS и GMPLS описаны в [RFC5151].

Как и внутридоменную TE, междоменную организацию трафика лучше всего реализовать при наличии возможности создания матрицы, отражающей объем трафика из одной AS в другую.

Транспортные протоколы L4 с поддержкой нескольких путей разработаны для обмена трафиком между доменами и обеспечивают возможность некоторого влияния на выбор путей. Для эффективной работы таких протоколов им требуется представление о путях и состояниях сетей в других доменах, но эти сведения могут быть недоступными, неполными и не заслуживающими доверия.

8. Обзор современных методов TE в работающих сетях IP

В этом разделе представлен обзор некоторых современных применений TE в сетях IP с основным вниманием к аспектам управления функцией маршрутизации в рабочем контексте. Цель состоит в предоставлении обзора наиболее распространённых применений без попыток охвата всех.

Сервис-провайдеры применяют множество описанных в этом документе механизмов TE для оптимизации производительности своих сетей IP, хотя некоторые совсем не используют их. Эти методы включают планирование

пропускной способности, в том числе с использованием ECMP, для долгих интервалов, управление маршрутизацией с использованием метрики IGP и MPLS, а также планирование путей и управление ими с использованием MPLS и SR на средних интервалах и механизмы управления трафиком для коротких интервалов.

- Планирование пропускной способности является важным компонентом планирования эффективной IP-сети сервис-провайдером. Это планирование может учитывать местоположение новых каналов и узлов, алгоритмы WECMP, имеющиеся и прогнозируемые картины трафика, затраты, пропускную способность каналов, топологию, решения о маршрутизации и живучесть.

- Оптимизация производительности работающих сетей обычно представляет собой продолжающийся процесс, в ходе которого в сети постоянно собираются сведения о статистике трафика, параметрах производительности и индикаторах отказов. Эти эмпирические данные анализируются и применяются в качестве триггеров для механизмов TE. Для содействия процессам TE путём анализа сценариев до применения новой конфигурации в работающих сетях могут применяться инструменты прогнозирования «что, если» (what-if).

- TE внутри домена с использованием IGP выполняется путём увеличения значений метрики OSPF или IS-IS на перегруженных каналах, пока с них не будет снят достаточный объем трафика. Такому подходу присущи некоторые ограничения, отмеченные в параграфе 6.2. Подходы к TE внутри домена [RR94] [FT00] [FT01] [WANG] принимают на входе матрицу трафика, топологию сети и целевые параметры производительности, давая на выходе значения метрики и коэффициенты распределения нагрузки. Эти процессы позволяют более системно внедрять внутри домена TE с использованием IGP.

Администраторы сетей MPLS-TE задают и настраивают атрибуты каналов и ограничения ресурсов, такие как максимальная резервируемая пропускная способность и атрибуты класса ресурсов для каналов в домене. Для распространения сведений о топологии сети и атрибутах каналов всем маршрутизаторам домена служит протокол IGP на основе состояний каналов, поддерживающий расширения TE (IS-IS-TE или OSPF-TE). Администраторы задают LSP, начинающиеся в каждом маршрутизаторе, указывая для каждого LSP целевой узел и атрибуты LSP, указывающие требования, которые должны быть выполнены при выборе пути. Атрибуты могут включать явный путь для LSP или маршрутизатор-источник может использовать локальный процесс маршрутизации на основе ограничений для расчёта пути LSP. Для создания LSP применяется сигнальный протокол RSVP-TE. Задавая для каналов и LSP подходящие значения пропускной способности можно предотвратить или смягчить перегрузки, связанные с неравномерным распределением трафика.

Атрибуты пропускной способности LSP связаны с требованиями к пропускной способности проходящего по этому LSP трафика. Атрибут трафика для LSP можно изменить с учётом сохраняющейся смены потребностей (рост или сокращение трафика). Если возникает перегрузка сети из-за непредвиденных событий, для смягчения проблемы можно перемаршрутизировать имеющиеся LSP или администратор может добавить LSP для переноса части трафика на другие пути. Доступная для резервирования пропускная способность может быть сокращена на перегруженных каналах, чтобы некоторые LSP были перенесены на другие пути. Матрицу трафика в домене MPLS можно оценить путём отслеживания трафика на LSP. Такую статистику трафика можно использовать для разных целей, включая планирование и оптимизацию сети.

Системы сетевого управления и планирования развиваются, принимая на себя большую часть ответственности за определение путей трафика в сетях TE. Это позволяет получить представление о ресурсах в масштабе всей сети и упрощает координацию использования ресурсов для всех потоков трафика в сети. Первоначальные решения с использованием PCE для расчёта пути от имени сетевых маршрутизаторов уступили место подходу на основе архитектуры SDN. PCE с учётом состояний может отслеживать все LSP в сети и распространять их для более эффективного использования доступных ресурсов. Такой элемент PCE может быть частью оркестратора сети, использующего PCEP или иной интерфейс настройки и управления для инструктирования сигнального протокола или непосредственного программирования маршрутизаторов.

Маршрутизация по сегментам (SR) использует централизованный контроллер TE и плоскость пересылки MPLS или IPv6, но не требует применять сигнальный протокол или протокол плоскости управления для резервирования ресурсов в маршрутизаторах. Резервирование ресурсов выполняется логически внутри контроллера и не распространяется в маршрутизаторы. Пакеты направляются через сеть с использованием SR и это может быть полезно для настройки и оперативного масштабирования.

Как отмечено в разделе 7, входной контроль за распределением входящего в домен трафика обычно отсутствует. Поэтому основной целью междоменной организации трафика является оптимизация распределения исходящего трафика между множеством выходных междоменных каналов. В географически распределенной сети (например, сети международного провайдера) важно сохранение возможности работы региональной сети, когда это требуется, с сохранением преимуществ объединённой в глобальном масштабе сети.

TE между доменами с BGP начинается с размещения партнерских точек соединений, расположенных в непосредственной близости к источникам/получателям трафика и предоставляющих наименее затратные пути через сеть между партнерскими точками и источниками/получателями. Некоторые проблемы принятия решений о местоположении, возникающие в связи с междоменной маршрутизацией, рассмотрены в работе [AWD5].

После задания местоположений и реализации партнерских соединений оператор сети решает, как лучше обрабатывать маршруты, анонсируемые партнёром, а также как распространять маршруты партнёра в своей сети. Одним из способов организации исходящих потоков трафика в сети со множеством партнерских соединений является создание иерархии партнёров. Обычно для пересылки трафика выбираются кратчайшие пути через AS, но могут применяться метрики BGP для предпочтения некоторых партнёров и определённых путей. Предпочтительными считаются партнёры, подключённые по каналам с большей доступной пропускной способностью. Могут потребоваться изменения, например, при взаимодействии с «проблемным партнёром», с которым трудно работать по обновлениям, или запрашивающим слишком высокую цену за подключение к своей сети. В таких случаях для партнёра можно снизить уровень предпочтения. Этот тип изменений может влиять на большой объем трафика и к нему следует прибегать лишь при невозможности достигнуть результатов иным способом.

При наличии нескольких точек выхода в направлении данного партнёра, из которых перегружена лишь одна, не требуется полностью переносить трафик от этого партнёра и достаточно снять его с перегруженного соединения. Это можно сделать с помощью пассивных метрик IGP, а также фильтрации AS_PATH или префиксов.

9. Вопросы безопасности

В общем случае механизмы TE не оказывают влияния на безопасность и этот документ не вносит новых вопросов безопасности.

Безопасность сетей важна и механизмы TE могут иметь преимущества и недостатки, отмеченные ниже.

- TE может использовать туннели, слегка помогающие защитить трафик от просмотра, а в некоторых случаях туннели могут использовать шифрование.
- TE помещает трафик на предсказуемые пути через сеть, что может облегчить поиск и организацию атак.
- TE зачастую усложняет работу сети и управление ею, что может приводить к ошибкам, снижающим уровень безопасности.
- TE позволяет направлять трафик в более защищённые каналы и участки сети.
- TE можно применять для направления трафика через узлы, обеспечивающие дополнительные функции защиты.

Последствия атак на протоколы управления и поддержки, применяемые при работе сетей TE, могут быть значимыми:

- трафик можно перехватить для передачи через специальные узлы для его просмотра и даже доставки в ненужное место;
- трафик можно направить на пути, обеспечивающие качество доставки ниже желаемого;
- можно перегрузить сети или израсходовать их ресурсы.

Поэтому важно применять адекватные механизмы защиты, такие как проверка подлинности, на всех протоколах, используемых для TE.

По деталям используемых путей TE можно определить некоторые аспекты сетей. Например, можно судить о связности каналов, качестве и загрузке отдельных каналов по сведениям о путях трафика и требованиям к сети (путём просмотра управляющих сообщений или отслеживания путей). Такие сведения можно использовать для организации целевых атак (например, нарушения работы важных каналов) или раскрытия коммерчески важных сведений (например, о приближении сети к насыщению). Поэтому операторы могут применять методы маскировки и сокрытия данных из своей сети.

Внешние интерфейсы управления, служащие для поддержки и управления системами TE (см. параграф 5.1.2), обеспечивают гибкость для администраторов и клиентов, но могут вносить риск раскрытия деталей устройства сети для потенциальных злоумышленников. Используемые в таких интерфейсах протоколы должны быть защищены от прослушивания и изменения данных, а доступ к интерфейсам должен сопровождаться проверкой полномочий.

10. Взаимодействие с IANA

Этот документ не требует действий IANA.

11. Литература

[AFD03] Pan, R., Breslau, L., Prabhakar, B., and S. Shenker, "Approximate fairness through differential dropping", ACM SIGCOMM Computer Communication Review, Volume 33, Issue 2, Pages 23-39, DOI 10.1145/956981.956985, April 2003, <https://dl.acm.org/doi/10.1145/956981.956985>.

[AJ19] Adekitan, A., Abolade, J., and O. Shobayo, "Data mining approach for predicting the daily Internet data traffic of a smart university", Journal of Big Data, Volume 6, Number 1, Page 1, DOI 10.1186/s40537-019-0176-5, February 2019, <https://journalofbigdata.springeropen.com/track/pdf/10.1186/s40537-019-0176-5.pdf>.

[ATSSS] 3GPP, "Study on access traffic steering, switch and splitting support in the 5G System (5GS) architecture", Release 16, 3GPP TR 23.793, December 2018, <https://www.3gpp.org/ftp//Specs/archive/23_series/23.793/23793-g00.zip>.

[AWD2] Awduche, D., "MPLS and traffic engineering in IP networks", IEEE Communications Magazine, Volume 37, Issue 12, Pages 42-47, DOI 10.1109/35.809383, December 1999, <https://ieeexplore.ieee.org/document/809383>.

[AWD5] Awduche, D., "An approach to optimal peering between autonomous systems in the Internet", Proceedings 7th International Conference on Computer Communications and Networks (Cat. No. 98EX226), DOI 10.1109/ICCCN.1998.998795, October 1998, <https://ieeexplore.ieee.org/document/998795>.

[E.360.1] ITU-T, "Framework for QoS routing and related traffic engineering methods for IP-, ATM-, and TDM-based multiservice networks", ITU-T Recommendation E.360.1, May 2002, <https://www.itu.int/rec/T-REC-E.360.1-200205-I/en>.

[ENHANCED-VPN] Dong, J., Bryant, S., Li, Z., Miyasaka, T., and Y. Lee, "A Framework for NRP-based Enhanced Virtual Private Network", Work in Progress, Internet-Draft, draft-ietf-teas-enhanced-vpn-17, 25 December 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-enhanced-vpn-17>.

[Err309] RFC Errata, Erratum ID 309, RFC 3272, <https://www.rfc-editor.org/errata/eid309>.

[EVPN-UNEQUAL-LB] Malhotra, N., Ed., Sajassi, A., Rabadan, J., Drake, J., Lingala, A., and S. Thoria, "Weighted Multi-Path Procedures for EVPN Multi-Homing", Work in Progress, Internet-Draft, draft-ietf-bess-evpn-unequal-lb-21, 7 December 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-bess-evpn-unequal-lb-21>.

[FLJA93] Floyd, S. and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, Volume 1, Issue 4, Pages 397-413, DOI 10.1109/90.251892, August 1993, <https://www.icir.org/floyd/papers/early.twocolumn.pdf>.

[FT00] Fortz, B. and M. Thorup, "Internet Traffic Engineering by Optimizing OSPF Weights", Proceedings IEEE INFOCOM 2000, DOI 10.1109/INFCOM.2000.832225, March 2000, <https://www.cs.cornell.edu/courses/cs619/2004fa/documents/ospf_opt.pdf>.

[FT01] Fortz, B. and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World", IEEE Journal on Selected Areas in Communications, DOI 10.1109/JSAC.2002.1003042, May 2002, <https://ieeexplore.ieee.org/document/1003042>.

[GRPC] gRPC Authors, "gRPC: A high performance, open source universal RPC framework", <https://grpc.io>.

[KELLY] Kelly, F., "Notes on effective bandwidths", Oxford University Press, 1996.

[MA] Ma, Q., "Quality-of-Service Routing in Integrated Services Networks", Ph.D. Dissertation, Carnegie Mellon University, CMU-CS-98-138, January 1998, <https://apps.dtic.mil/sti/pdfs/ADA352299.pdf>.

[MATE] Elwalid, A., Jin, C., Low, S., and I. Widjaja, "MATE: MPLS Adaptive Traffic Engineering", Proceedings IEEE INFOCOM 2001, Conference on Computer Communications, Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No. 01CH37213), DOI 10.1109/INFCOM.2001.916625, August 2002, <https://www.yumpu.com/en/document/view/35140398/mate-mpls-adaptive-traffic-engineering-infocom-ieee-xplore/8>.

[MR99] Mitra, D. and K.G. Ramakrishnan, "A case study of multiservice, multipriority traffic engineering design for data networks", Seamless Interconnection for Universal Services, Global Telecommunications Conference, GLOBECOM'99, (Cat. No. 99CH37042), DOI 10.1109/GLOCOM.1999.830281, December 1999, <https://ieeexplore.ieee.org/document/830281>.

[MULTIPATH-DCCP] Amend, M., Ed., Brunstrom, A., Kassler, A., Rakocevic, V., and S. Johnson, "DCCP Extensions for Multipath Operation with Multiple Addresses", Work in Progress, Internet-Draft, draft-ietf-tsvwg-multipath-dccp-11, 12 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-tsvwg-multipath-dccp-11>.

[NETWORK-SLICES] Farrel, A., Ed., Drake, J., Ed., Rokui, R., Homma, S., Makhijani, K., Contreras, L. M., and J. Tantsura, "A Framework for Network Slices in Networks Built from IETF Technologies", Work in Progress[1], Internet-Draft, draft-ietf-teas-ietf-network-slices-25, 14 September 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-teas-ietf-network-slices-25>.

[PERFORMANCE-ROUTING] Xu, X., Hegde, S., Talaulikar, K., Boucadair, M., and C. Jacquenet, "Performance-based BGP Routing Mechanism", Work in Progress, Internet-Draft, draft-ietf-idr-performance-routing-03, 22 December 2020, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-performance-routing-03>.

[QUIC-MULTIPATH] Liu, Y., Ed., Ma, Y., Ed., De Coninck, Q., Ed., Bonaventure, O., Huitema, C., and M. Kühlewind, Ed., "Multipath Extension for QUIC", Work in Progress, Internet-Draft, draft-ietf-quic-multipath-06, 23 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-quic-multipath-06>.

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

[RFC1102] Clark, D., "Policy routing in Internet protocols", RFC 1102, DOI 10.17487/RFC1102, May 1989, <https://www.rfc-editor.org/info/rfc1102>.

[RFC1104] Braun, H., "Models of policy based routing", RFC 1104, DOI 10.17487/RFC1104, June 1989, <https://www.rfc-editor.org/info/rfc1104>.

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, DOI 10.17487/RFC2205, September 1997, <https://www.rfc-editor.org/info/rfc2205>.

[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>.

[RFC2386] Crawley, E., Nair, R., Rajagopalan, B., and H. Sandick, "A Framework for QoS-based Routing in the Internet", RFC 2386, DOI 10.17487/RFC2386, August 1998, <https://www.rfc-editor.org/info/rfc2386>.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, DOI 10.17487/RFC2474, December 1998, <https://www.rfc-editor.org/info/rfc2474>.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, DOI 10.17487/RFC2475, December 1998, <https://www.rfc-editor.org/info/rfc2475>.

[RFC2597] Heinanen, J., Baker, F., Weiss, W., and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, DOI 10.17487/RFC2597, June 1999, <https://www.rfc-editor.org/info/rfc2597>.

[1]Опубликовано в RFC 9543. *Прим. перев.*

[RFC2678] Mahdavi, J. and V. Paxson, "IPPM Metrics for Measuring Connectivity", RFC 2678, DOI 10.17487/RFC2678, September 1999, <https://www.rfc-editor.org/info/rfc2678>.

[RFC2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M., and J. McManus, "Requirements for Traffic Engineering Over MPLS", RFC 2702, DOI 10.17487/RFC2702, September 1999, <https://www.rfc-editor.org/info/rfc2702>.

[RFC2722] Brownlee, N., Mills, C., and G. Ruth, "Traffic Flow Measurement: Architecture", RFC 2722, DOI 10.17487/RFC2722, October 1999, <https://www.rfc-editor.org/info/rfc2722>.

[RFC2753] Yavatkar, R., Pendarakis, D., and R. Guerin, "A Framework for Policy-based Admission Control", RFC 2753, DOI 10.17487/RFC2753, January 2000, <https://www.rfc-editor.org/info/rfc2753>.

[RFC2961] Berger, L., Gan, D., Swallow, G., Pan, P., Tommasi, F., and S. Molendini, "RSVP Refresh Overhead Reduction Extensions", RFC 2961, DOI 10.17487/RFC2961, April 2001, <https://www.rfc-editor.org/info/rfc2961>.

[RFC2998] Bernet, Y., Ford, P., Yavatkar, R., Baker, F., Zhang, L., Speer, M., Braden, R., Davie, B., Wroclawski, J., and E. Felstaine, "A Framework for Integrated Services Operation over Diffserv Networks", RFC 2998, DOI 10.17487/RFC2998, November 2000, <https://www.rfc-editor.org/info/rfc2998>.

[RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <https://www.rfc-editor.org/info/rfc3031>.

[RFC3086] Nichols, K. and B. Carpenter, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", RFC 3086, DOI 10.17487/RFC3086, April 2001, <https://www.rfc-editor.org/info/rfc3086>.

[RFC3124] Balakrishnan, H. and S. Seshan, "The Congestion Manager", RFC 3124, DOI 10.17487/RFC3124, June 2001, <https://www.rfc-editor.org/info/rfc3124>.

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, DOI 10.17487/RFC3168, September 2001, <https://www.rfc-editor.org/info/rfc3168>.

[RFC3175] Baker, F., Iturralde, C., Le Faucheur, F., and B. Davie, "Aggregation of RSVP for IPv4 and IPv6 Reservations", RFC 3175, DOI 10.17487/RFC3175, September 2001, <https://www.rfc-editor.org/info/rfc3175>.

[RFC3198] Westerinen, A., Schnizlein, J., Strassner, J., Scherling, M., Quinn, B., Herzog, S., Huynh, A., Carlson, M., Perry, J., and S. Waldbusser, "Terminology for Policy-Based Management", RFC 3198, DOI 10.17487/RFC3198, November 2001, <https://www.rfc-editor.org/info/rfc3198>.

[RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, DOI 10.17487/RFC3209, December 2001, <https://www.rfc-editor.org/info/rfc3209>.

[RFC3270] Le Faucheur, F., Ed., Wu, L., Davie, B., Davari, S., Vaananen, P., Krishnan, R., Cheval, P., and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, DOI 10.17487/RFC3270, May 2002, <https://www.rfc-editor.org/info/rfc3270>.

[RFC3272] Awduche, D., Chiu, A., Elwalid, A., Widjaja, I., and X. Xiao, "Overview and Principles of Internet Traffic Engineering", RFC 3272, DOI 10.17487/RFC3272, May 2002, <https://www.rfc-editor.org/info/rfc3272>.

[RFC3469] Sharma, V., Ed. and F. Hellstrand, Ed., "Framework for Multi-Protocol Label Switching (MPLS)-based Recovery", RFC 3469, DOI 10.17487/RFC3469, February 2003, <https://www.rfc-editor.org/info/rfc3469>.

[RFC3473] Berger, L., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 3473, DOI 10.17487/RFC3473, January 2003, <https://www.rfc-editor.org/info/rfc3473>.

[RFC3630] Katz, D., Kompella, K., and D. Yeung, "Traffic Engineering (TE) Extensions to OSPF Version 2", RFC 3630, DOI 10.17487/RFC3630, September 2003, <https://www.rfc-editor.org/info/rfc3630>.

[RFC3945] Mannie, E., Ed., "Generalized Multi-Protocol Label Switching (GMPLS) Architecture", RFC 3945, DOI 10.17487/RFC3945, October 2004, <https://www.rfc-editor.org/info/rfc3945>.

[RFC4090] Pan, P., Ed., Swallow, G., Ed., and A. Atlas, Ed., "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", RFC 4090, DOI 10.17487/RFC4090, May 2005, <https://www.rfc-editor.org/info/rfc4090>.

[RFC4124] Le Faucheur, F., Ed., "Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering", RFC 4124, DOI 10.17487/RFC4124, June 2005, <https://www.rfc-editor.org/info/rfc4124>.

[RFC4203] Kompella, K., Ed. and Y. Rekhter, Ed., "OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS)", RFC 4203, DOI 10.17487/RFC4203, October 2005, <https://www.rfc-editor.org/info/rfc4203>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI 10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/info/rfc4271>.

[RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, DOI 10.17487/RFC4340, March 2006, <https://www.rfc-editor.org/info/rfc4340>.

[RFC4461] Yasukawa, S., Ed., "Signaling Requirements for Point-to-Multipoint Traffic-Engineered MPLS Label Switched Paths (LSPs)", RFC 4461, DOI 10.17487/RFC4461, April 2006, <https://www.rfc-editor.org/info/rfc4461>.

[RFC4594] Babiarz, J., Chan, K., and F. Baker, "Configuration Guidelines for DiffServ Service Classes", RFC 4594, DOI 10.17487/RFC4594, August 2006, <https://www.rfc-editor.org/info/rfc4594>.

[RFC4655] Farrel, A., Vasseur, J.-P., and J. Ash, "A Path Computation Element (PCE)-Based Architecture", RFC 4655, DOI 10.17487/RFC4655, August 2006, <https://www.rfc-editor.org/info/rfc4655>.

[RFC4872] Lang, J.P., Ed., Rekhter, Y., Ed., and D. Papadimitriou, Ed., "RSVP-TE Extensions in Support of End-to-End Generalized Multi-Protocol Label Switching (GMPLS) Recovery", RFC 4872, DOI 10.17487/RFC4872, May 2007, <https://www.rfc-editor.org/info/rfc4872>.

[RFC4873] Berger, L., Bryskin, I., Papadimitriou, D., and A. Farrel, "GMPLS Segment Recovery", RFC 4873, DOI 10.17487/RFC4873, May 2007, <https://www.rfc-editor.org/info/rfc4873>.

[RFC4875] Aggarwal, R., Ed., Papadimitriou, D., Ed., and S. Yasukawa, Ed., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", RFC 4875, DOI 10.17487/RFC4875, May 2007, <https://www.rfc-editor.org/info/rfc4875>.

[RFC5151] Farrel, A., Ed., Ayyangar, A., and JP. Vasseur, "Inter-Domain MPLS and GMPLS Traffic Engineering — Resource Reservation Protocol-Traffic Engineering (RSVP-TE) Extensions", RFC 5151, DOI 10.17487/RFC5151, February 2008, <https://www.rfc-editor.org/info/rfc5151>.

[RFC5250] Berger, L., Bryskin, I., Zinin, A., and R. Coltun, "The OSPF Opaque LSA Option", RFC 5250, DOI 10.17487/RFC5250, July 2008, <https://www.rfc-editor.org/info/rfc5250>.

[RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <https://www.rfc-editor.org/info/rfc5305>.

[RFC5329] Ishiguro, K., Manral, V., Davey, A., and A. Lindem, Ed., "Traffic Engineering Extensions to OSPF Version 3", RFC 5329, DOI 10.17487/RFC5329, September 2008, <https://www.rfc-editor.org/info/rfc5329>.

[RFC5331] Aggarwal, R., Rekhter, Y., and E. Rosen, "MPLS Upstream Label Assignment and Context-Specific Label Space", RFC 5331, DOI 10.17487/RFC5331, August 2008, <https://www.rfc-editor.org/info/rfc5331>.

[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>.

[RFC5394] Bryskin, I., Papadimitriou, D., Berger, L., and J. Ash, "Policy-Enabled Path Computation Framework", RFC 5394, DOI 10.17487/RFC5394, December 2008, <https://www.rfc-editor.org/info/rfc5394>.

[RFC5440] Vasseur, JP., Ed. and JL. Le Roux, Ed., "Path Computation Element (PCE) Communication Protocol (PCEP)", RFC 5440, DOI 10.17487/RFC5440, March 2009, <https://www.rfc-editor.org/info/rfc5440>.

[RFC5470] Sadasivan, G., Brownlee, N., Claise, B., and J. Quittek, "Architecture for IP Flow Information Export", RFC 5470, DOI 10.17487/RFC5470, March 2009, <https://www.rfc-editor.org/info/rfc5470>.

[RFC5472] Zseby, T., Boschi, E., Brownlee, N., and B. Claise, "IP Flow Information Export (IPFIX) Applicability", RFC 5472, DOI 10.17487/RFC5472, March 2009, <https://www.rfc-editor.org/info/rfc5472>.

[RFC5541] Le Roux, JL., Vasseur, JP., and Y. Lee, "Encoding of Objective Functions in the Path Computation Element Communication Protocol (PCEP)", RFC 5541, DOI 10.17487/RFC5541, June 2009, <https://www.rfc-editor.org/info/rfc5541>.

[RFC5557] Lee, Y., Le Roux, JL., King, D., and E. Oki, "Path Computation Element Communication Protocol (PCEP) Requirements and Protocol Extensions in Support of Global Concurrent Optimization", RFC 5557, DOI 10.17487/RFC5557, July 2009, <https://www.rfc-editor.org/info/rfc5557>.

[RFC5559] Eardley, P., Ed., "Pre-Congestion Notification (PCN) Architecture", RFC 5559, DOI 10.17487/RFC5559, June 2009, <https://www.rfc-editor.org/info/rfc5559>.

[RFC5623] Oki, E., Takeda, T., Le Roux, JL., and A. Farrel, "Framework for PCE-Based Inter-Layer MPLS and GMPLS Traffic Engineering", RFC 5623, DOI 10.17487/RFC5623, September 2009, <https://www.rfc-editor.org/info/rfc5623>.

[RFC5664] Halevy, B., Welch, B., and J. Zelenka, "Object-Based Parallel NFS (pNFS) Operations", RFC 5664, DOI 10.17487/RFC5664, January 2010, <https://www.rfc-editor.org/info/rfc5664>.

[RFC5671] Yasukawa, S. and A. Farrel, Ed., "Applicability of the Path Computation Element (PCE) to Point-to-Multipoint (P2MP) MPLS and GMPLS Traffic Engineering (TE)", RFC 5671, DOI 10.17487/RFC5671, October 2009, <https://www.rfc-editor.org/info/rfc5671>.

[RFC5693] Seedorf, J. and E. Burger, "Application-Layer Traffic Optimization (ALTO) Problem Statement", RFC 5693, DOI 10.17487/RFC5693, October 2009, <https://www.rfc-editor.org/info/rfc5693>.

[RFC6107] Shiomoto, K., Ed. and A. Farrel, Ed., "Procedures for Dynamically Signaled Hierarchical Label Switched Paths", RFC 6107, DOI 10.17487/RFC6107, February 2011, <https://www.rfc-editor.org/info/rfc6107>.

[RFC6119] Harrison, J., Berger, J., and M. Bartlett, "IPv6 Traffic Engineering in IS-IS", RFC 6119, DOI 10.17487/RFC6119, February 2011, <https://www.rfc-editor.org/info/rfc6119>.

[RFC6241] Enns, R., Ed., Bjorklund, M., Ed., Schoenwaelder, J., Ed., and A. Bierman, Ed., "Network Configuration Protocol (NETCONF)", RFC 6241, DOI 10.17487/RFC6241, June 2011, <https://www.rfc-editor.org/info/rfc6241>.

[RFC6372] Sprecher, N., Ed. and A. Farrel, Ed., "MPLS Transport Profile (MPLS-TP) Survivability Framework", RFC 6372, DOI 10.17487/RFC6372, September 2011, <https://www.rfc-editor.org/info/rfc6372>.

[RFC6374] Frost, D. and S. Bryant, "Packet Loss and Delay Measurement for MPLS Networks", RFC 6374, DOI 10.17487/RFC6374, September 2011, <https://www.rfc-editor.org/info/rfc6374>.

[RFC6601] Ash, G., Ed. and D. McDysan, "Generic Connection Admission Control (GCAC) Algorithm Specification for IP/MPLS Networks", RFC 6601, DOI 10.17487/RFC6601, April 2012, <https://www.rfc-editor.org/info/rfc6601>.

[RFC6805] King, D., Ed. and A. Farrel, Ed., "The Application of the Path Computation Element Architecture to the Determination of a Sequence of Domains in MPLS and GMPLS", RFC 6805, DOI 10.17487/RFC6805, November 2012, <https://www.rfc-editor.org/info/rfc6805>.

[RFC7011] Claise, B., Ed., Trammell, B., Ed., and P. Aitken, "Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of Flow Information", STD 77, RFC 7011, DOI 10.17487/RFC7011, September 2013, <https://www.rfc-editor.org/info/rfc7011>.

[RFC7149] Boucadair, M. and C. Jacquenet, "Software-Defined Networking: A Perspective from within a Service Provider Environment", RFC 7149, DOI 10.17487/RFC7149, March 2014, <https://www.rfc-editor.org/info/rfc7149>.

[RFC7285] Alimi, R., Ed., Penno, R., Ed., Yang, Y., Ed., Kiesel, S., Previdi, S., Roome, W., Shalunov, S., and R. Woundy, "Application-Layer Traffic Optimization (ALTO) Protocol", RFC 7285, DOI 10.17487/RFC7285, September 2014, <https://www.rfc-editor.org/info/rfc7285>.

[RFC7390] Rahman, A., Ed. and E. Dijk, Ed., "Group Communication for the Constrained Application Protocol (CoAP)", RFC 7390, DOI 10.17487/RFC7390, October 2014, <https://www.rfc-editor.org/info/rfc7390>.

[RFC7426] Haleplidis, E., Ed., Pentikousis, K., Ed., Denazis, S., Hadi Salim, J., Meyer, D., and O. Koufopavlou, "Software-Defined Networking (SDN): Layers and Architecture Terminology", RFC 7426, DOI 10.17487/RFC7426, January 2015, <https://www.rfc-editor.org/info/rfc7426>.

[RFC7471] Giacalone, S., Ward, D., Drake, J., Atlas, A., and S. Previdi, "OSPF Traffic Engineering (TE) Metric Extensions", RFC 7471, DOI 10.17487/RFC7471, March 2015, <https://www.rfc-editor.org/info/rfc7471>.

[RFC7491] King, D. and A. Farrel, "A PCE-Based Architecture for Application-Based Network Operations", RFC 7491, DOI 10.17487/RFC7491, March 2015, <https://www.rfc-editor.org/info/rfc7491>.

[RFC7551] Zhang, F., Ed., Jing, R., and R. Gandhi, Ed., "RSVP-TE Extensions for Associated Bidirectional Label Switched Paths (LSPs)", RFC 7551, DOI 10.17487/RFC7551, May 2015, <https://www.rfc-editor.org/info/rfc7551>.

[RFC7567] Baker, F., Ed. and G. Fairhurst, Ed., "IETF Recommendations Regarding Active Queue Management", BCP 197, RFC 7567, DOI 10.17487/RFC7567, July 2015, <https://www.rfc-editor.org/info/rfc7567>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC7679] Almes, G., Kalidindi, S., Zekauskas, M., and A. Morton, Ed., "A One-Way Delay Metric for IP Performance Metrics (IPPM)", STD 81, RFC 7679, DOI 10.17487/RFC7679, January 2016, <https://www.rfc-editor.org/info/rfc7679>.

[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>.

[RFC7923] Voit, E., Clemm, A., and A. Gonzalez Prieto, "Requirements for Subscription to YANG Datastores", RFC 7923, DOI 10.17487/RFC7923, June 2016, <https://www.rfc-editor.org/info/rfc7923>.

[RFC7926] Farrel, A., Ed., Drake, J., Bitar, N., Swallow, G., Ceccarelli, D., and X. Zhang, "Problem Statement and Architecture for Information Exchange between Interconnected Traffic-Engineered Networks", BCP 206, RFC 7926, DOI 10.17487/RFC7926, July 2016, <https://www.rfc-editor.org/info/rfc7926>.

[RFC7950] Bjorklund, M., Ed., "The YANG 1.1 Data Modeling Language", RFC 7950, DOI 10.17487/RFC7950, August 2016, <https://www.rfc-editor.org/info/rfc7950>.

[RFC8033] Pan, R., Natarajan, P., Baker, F., and G. White, "Proportional Integral Controller Enhanced (PIE): A Lightweight Control Scheme to Address the Bufferbloat Problem", RFC 8033, DOI 10.17487/RFC8033, February 2017, <https://www.rfc-editor.org/info/rfc8033>.

[RFC8034] White, G. and R. Pan, "Active Queue Management (AQM) Based on Proportional Integral Controller Enhanced (PIE) for Data-Over-Cable Service Interface Specifications (DOCSIS) Cable Modems", RFC 8034, DOI 10.17487/RFC8034, February 2017, <https://www.rfc-editor.org/info/rfc8034>.

[RFC8040] Bierman, A., Bjorklund, M., and K. Watsen, "RESTCONF Protocol", RFC 8040, DOI 10.17487/RFC8040, January 2017, <https://www.rfc-editor.org/info/rfc8040>.

[RFC8051] Zhang, X., Ed. and I. Minei, Ed., "Applicability of a Stateful Path Computation Element (PCE)", RFC 8051, DOI 10.17487/RFC8051, January 2017, <https://www.rfc-editor.org/info/rfc8051>.

[RFC8189] Randriamasy, S., Roome, W., and N. Schwan, "Multi-Cost Application-Layer Traffic Optimization (ALTO)", RFC 8189, DOI 10.17487/RFC8189, October 2017, <https://www.rfc-editor.org/info/rfc8189>.

[RFC8231] Crabbe, E., Minei, I., Medved, J., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for Stateful PCE", RFC 8231, DOI 10.17487/RFC8231, September 2017, <https://www.rfc-editor.org/info/rfc8231>.

[RFC8259] Bray, T., Ed., "The JavaScript Object Notation (JSON) Data Interchange Format", STD 90, RFC 8259, DOI 10.17487/RFC8259, December 2017, <https://www.rfc-editor.org/info/rfc8259>.

[RFC8279] Wijnands, IJ., Ed., Rosen, E., Ed., Dolganow, A., Przygienda, T., and S. Aldrin, "Multicast Using Bit Index Explicit Replication (BIER)", RFC 8279, DOI 10.17487/RFC8279, November 2017, <https://www.rfc-editor.org/info/rfc8279>.

[RFC8281] Crabbe, E., Minei, I., Sivabalan, S., and R. Varga, "Path Computation Element Communication Protocol (PCEP) Extensions for PCE-Initiated LSP Setup in a Stateful PCE Model", RFC 8281, DOI 10.17487/RFC8281, December 2017, <https://www.rfc-editor.org/info/rfc8281>.

[RFC8283] Farrel, A., Ed., Zhao, Q., Ed., Li, Z., and C. Zhou, "An Architecture for Use of PCE and the PCE Communication Protocol (PCEP) in a Network with Central Control", RFC 8283, DOI 10.17487/RFC8283, December 2017, <https://www.rfc-editor.org/info/rfc8283>.

[RFC8290] Hoeiland-Joergensen, T., McKenney, P., Taht, D., Gettys, J., and E. Dumazet, "The Flow Queue CoDel Packet Scheduler and Active Queue Management Algorithm", RFC 8290, DOI 10.17487/RFC8290, January 2018, <https://www.rfc-editor.org/info/rfc8290>.

[RFC8402] Filsfils, C., Ed., Previdi, S., Ed., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", RFC 8402, DOI 10.17487/RFC8402, July 2018, <https://www.rfc-editor.org/info/rfc8402>.

[RFC8453] Ceccarelli, D., Ed. and Y. Lee, Ed., "Framework for Abstraction and Control of TE Networks (ACTN)", RFC 8453, DOI 10.17487/RFC8453, August 2018, <https://www.rfc-editor.org/info/rfc8453>.

[RFC8570] Ginsberg, L., Ed., Previdi, S., Ed., Giacalone, S., Ward, D., Drake, J., and Q. Wu, "IS-IS Traffic Engineering (TE) Metric Extensions", RFC 8570, DOI 10.17487/RFC8570, March 2019, <https://www.rfc-editor.org/info/rfc8570>.

[RFC8571] Ginsberg, L., Ed., Previdi, S., Wu, Q., Tantsura, J., and C. Filsfils, "BGP - Link State (BGP-LS) Advertisement of IGP Traffic Engineering Performance Metric Extensions", RFC 8571, DOI 10.17487/RFC8571, March 2019, <https://www.rfc-editor.org/info/rfc8571>.

[RFC8655] Finn, N., Thubert, P., Varga, B., and J. Farkas, "Deterministic Networking Architecture", RFC 8655, DOI 10.17487/RFC8655, October 2019, <https://www.rfc-editor.org/info/rfc8655>.

[RFC8664] Sivabalan, S., Filsfils, C., Tantsura, J., Henderickx, W., and J. Hardwick, "Path Computation Element Communication Protocol (PCEP) Extensions for Segment Routing", RFC 8664, DOI 10.17487/RFC8664, December 2019, <https://www.rfc-editor.org/info/rfc8664>.

[RFC8684] Ford, A., Raiciu, C., Handley, M., Bonaventure, O., and C. Paasch, "TCP Extensions for Multipath Operation with Multiple Addresses", RFC 8684, DOI 10.17487/RFC8684, March 2020, <https://www.rfc-editor.org/info/rfc8684>.

[RFC8685] Zhang, F., Zhao, Q., Gonzalez de Dios, O., Casellas, R., and D. King, "Path Computation Element Communication Protocol (PCEP) Extensions for the Hierarchical Path Computation Element (H-PCE) Architecture", RFC 8685, DOI 10.17487/RFC8685, December 2019, <https://www.rfc-editor.org/info/rfc8685>.

[RFC8795] Liu, X., Bryskin, I., Beeram, V., Saad, T., Shah, H., and O. Gonzalez de Dios, "YANG Data Model for Traffic Engineering (TE) Topologies", RFC 8795, DOI 10.17487/RFC8795, August 2020, <https://www.rfc-editor.org/info/rfc8795>.

[RFC8803] Bonaventure, O., Ed., Boucadair, M., Ed., Gundavelli, S., Seo, S., and B. Hesmans, "0-RTT TCP Convert Protocol", RFC 8803, DOI 10.17487/RFC8803, July 2020, <https://www.rfc-editor.org/info/rfc8803>.

[RFC8896] Randriamasy, S., Yang, R., Wu, Q., Deng, L., and N. Schwan, "Application-Layer Traffic Optimization (ALTO) Cost Calendar", RFC 8896, DOI 10.17487/RFC8896, November 2020, <https://www.rfc-editor.org/info/rfc8896>.

[RFC8938] Varga, B., Ed., Farkas, J., Berger, L., Malis, A., and S. Bryant, "Deterministic Networking (DetNet) Data Plane Framework", RFC 8938, DOI 10.17487/RFC8938, November 2020, <https://www.rfc-editor.org/info/rfc8938>.

[RFC8955] Loibl, C., Hares, S., Raszuk, R., McPherson, D., and M. Bacher, "Dissemination of Flow Specification Rules", RFC 8955, DOI 10.17487/RFC8955, December 2020, <https://www.rfc-editor.org/info/rfc8955>.

[RFC8972] Mirsky, G., Min, X., Nydell, H., Foote, R., Masputra, A., and E. Ruffini, "Simple Two-Way Active Measurement Protocol Optional Extensions", RFC 8972, DOI 10.17487/RFC8972, January 2021, <https://www.rfc-editor.org/info/rfc8972>.

[RFC9000] Iyengar, J., Ed. and M. Thomson, Ed., "QUIC: A UDP-Based Multiplexed and Secure Transport", RFC 9000, DOI 10.17487/RFC9000, May 2021, <https://www.rfc-editor.org/info/rfc9000>.

[RFC9023] Varga, B., Ed., Farkas, J., Malis, A., and S. Bryant, "Deterministic Networking (DetNet) Data Plane: IP over IEEE 802.1 Time-Sensitive Networking (TSN)", RFC 9023, DOI 10.17487/RFC9023, June 2021, <https://www.rfc-editor.org/info/rfc9023>.

[RFC9040] Touch, J., Welzl, M., and S. Islam, "TCP Control Block Interdependence", RFC 9040, DOI 10.17487/RFC9040, July 2021, <https://www.rfc-editor.org/info/rfc9040>.

[RFC9113] Thomson, M., Ed. and C. Benfield, Ed., "HTTP/2", RFC 9113, DOI 10.17487/RFC9113, June 2022, <https://www.rfc-editor.org/info/rfc9113>.

[RFC9256] Filsfils, C., Talaulikar, K., Ed., Voyer, D., Bogdanov, A., and P. Mattes, "Segment Routing Policy Architecture", RFC 9256, DOI 10.17487/RFC9256, July 2022, <https://www.rfc-editor.org/info/rfc9256>.

[RFC9262] Eckert, T., Ed., Menth, M., and G. Cauchie, "Tree Engineering for Bit Index Explicit Replication (BIER-TE)", RFC 9262, DOI 10.17487/RFC9262, October 2022, <https://www.rfc-editor.org/info/rfc9262>.

[RFC9298] Schinazi, D., "Proxying UDP in HTTP", RFC 9298, DOI 10.17487/RFC9298, August 2022, <https://www.rfc-editor.org/info/rfc9298>.

[RFC9315] Clemm, A., Ciavaglia, L., Granville, L. Z., and J. Tantsura, "Intent-Based Networking - Concepts and Definitions", RFC 9315, DOI 10.17487/RFC9315, October 2022, <https://www.rfc-editor.org/info/rfc9315>.

[RFC9332] De Schepper, K., Briscoe, B., Ed., and G. White, "Dual-Queue Coupled Active Queue Management (AQM) for Low Latency, Low Loss, and Scalable Throughput (L4S)", RFC 9332, DOI 10.17487/RFC9332, January 2023, <https://www.rfc-editor.org/info/rfc9332>.

[RFC9350] Psenak, P., Ed., Hegde, S., Filsfils, C., Talaulikar, K., and A. Gulko, "IGP Flexible Algorithm", RFC 9350, DOI 10.17487/RFC9350, February 2023, <https://www.rfc-editor.org/info/rfc9350>.

[RFC9439] Wu, Q., Yang, Y., Lee, Y., Dhody, D., Randriamasy, S., and L. Contreras, "Application-Layer Traffic Optimization (ALTO) Performance Cost Metrics", RFC 9439, DOI 10.17487/RFC9439, August 2023, <https://www.rfc-editor.org/info/rfc9439>.

[RFC9502] Britto, W., Hegde, S., Kaneriya, P., Shetty, R., Bonica, R., and P. Psenak, "IGP Flexible Algorithm in IP Networks", RFC 9502, DOI 10.17487/RFC9502, November 2023, <https://www.rfc-editor.org/info/rfc9502>.

[RFC9552] Talaulikar, K., Ed., "Distribution of Link-State and Traffic Engineering Information Using BGP", RFC 9552, DOI 10.17487/RFC9552, December 2023, <https://www.rfc-editor.org/info/rfc9552>.

[RR94] Rodrigues, M. and K.G. Ramakrishnan, "Optimal routing in shortest-path data networks", Bell Labs Technical Journal, Volume 6, Issue 1, Pages 117-138, DOI 10.1002/bltj.2267, August 2002, <https://onlinelibrary.wiley.com/doi/abs/10.1002/bltj.2267>.

[SLDC98] Suter, B., Lakshman, T.V., Stiliadis, D., and A.K. Choudhury, "Design considerations for supporting TCP with per-flow queueing", Proceedings IEEE INFOCOM '98, DOI 10.1109/INFCOM.1998.659666, April 1998, <https://ieeexplore.ieee.org/document/659666>.

[SR-TE-POLICY] Previdi, S., Filsfils, C., Talaulikar, K., Ed., Mattes, P., and D. Jain, "Advertising Segment Routing Policies in BGP", Work in Progress, Internet-Draft, draft-ietf-idr-segment-routing-te-policy-26, 23 October 2023, <https://datatracker.ietf.org/doc/html/draft-ietf-idr-segment-routing-te-policy-26>.

[SR-TI-LFA] Bashandy, A., Litkowski, S., Filsfils, C., Francois, P., Decraene, B., and D. Voyer, "Topology Independent Fast Reroute using Segment Routing", Work in Progress, Internet-Draft, draft-ietf-rtgwg-segment-routing-ti-lfa-13, 16 January 2024, <https://datatracker.ietf.org/doc/html/draft-ietf-rtgwg-segment-routing-ti-lfa-13>.

[TE-QoS-ROUTING] Ash, G., "Traffic Engineering & QoS Methods for IP-, ATM-, & Based Multiservice Networks", Work in Progress, Internet-Draft, draft-ietf-tewg-qos-routing-04, October 2001, <https://datatracker.ietf.org/doc/html/draft-ietf-tewg-qos-routing-04>.

[WANG] Wang, Y., Wang, Z., and L. Zhang, "Internet traffic engineering without full mesh overlaying", Proceedings IEEE INFOCOM 2001, DOI 10.1109/INFCOM.2001.916782, April 2001, <https://ieeexplore.ieee.org/document/916782>.

[XIAO] Xiao, X., Hannan, A., Bailey, B., and L. Ni, "Traffic Engineering with MPLS in the Internet", IEEE Network, Volume 14, Issue 2, Pages 28-33, DOI 10.1109/65.826369, March 2000, <https://courses.cs.washington.edu/courses/cse561/02au/papers/xiao-mpls-net00.pdf>.

[YARE95] Yang, C. and A. Reddy, "A Taxonomy for Congestion Control Algorithms in Packet Switching Networks", IEEE Network, Pages 34-45, DOI 10.1109/65.397042, August 1995, <https://ieeexplore.ieee.org/document/397042>.

Приложение A. Отличия от RFC 3272

Изменения, внесённые в документ по сравнению с [RFC3272], существенны и их нелегко обобщить по параграфам . Материал в этом документе был перемещён, часть текста удалена и добавлен новый текст. Принятый здесь подход заключается в перечислении содержимого [RFC3272] и данного документа с соответствующим указанием места размещения текста и его источника.

A.1. RFC 3272

- Раздел 1.0 (Введение) отредактирован в разделе 1.
  - Параграф 1.1 (Что такое Internet TE?) отредактирован в параграфе 1.1.
  - Параграф 1.2 (Область действия) перенесён в параграф 1.3.
  - Параграф 1.3 (Терминология) перенесён в параграф 1.4 с удалением некоторых устаревших терминов и незначительным редактированием.
- Раздел 2.0 (Основы) сохранен как раздел 2 с удалением части текста.
  - Параграф 2.1 (Контекст Internet TE) сохранен как параграф 2.1.
  - Параграф 2.2 (Контекст сети) переписан как параграф 2.2.
  - Параграф 2.3 (Контекст задачи) переписан как параграф 2.3.
    - Параграф 2.3.1 (Перегрузки и их последствия) сохранен как параграф 2.3.1.
  - Параграф 2.4 (Контекст решения) отредактирован как параграф 2.4.
    - Параграф 2.4.1 (Борьба с перегрузками) переформатирован как параграф 2.4.1.
  - Параграф 2.5 (Контекст реализации и применения) сохранен как параграф 2.5.
- Раздел 3.0 (Модель процесса TE) сохранен как раздел 3.
  - Параграф 3.1 (Компоненты модели процесса TE) сохранен как параграф 3.1.
  - Параграф 3.2 (Измерения) объединён с параграфом 3.1.
  - Параграф 3.3 (Моделирование, анализ и имитация) объединён с параграфом 3.1.
  - Параграф 3.4 (Оптимизация) объединён с параграфом 3.1.
- Раздел 4.0 (Исторический обзор и дальнейшее развитие) сохранен как раздел 5 с удалением исторических аспектов.
  - Параграф 4.1 (TE в классических телефонных сетях) удалён.
  - Параграф 4.2 (Эволюция TE в Internet) удалён.
  - Параграф 4.3 (Модель наложения) удалён.
  - Параграф 4.4 (Маршрутизация на основе ограничений) сохранен как параграф 5.1.3.1, но перемещён в параграф 5.1.
  - Параграф 4.5 (Обзор других проектов IETF, связанных с TE) сохранен как параграф 5.1 с добавлением новых подпараграфов.
    - Параграф 4.5.1 (Интегрированные услуги) сохранен как параграф 5.1.1.1.
    - Параграф 4.5.2 (RSVP) сохранен как параграф 5.1.3.2 с некоторым редактированием.
    - Параграф 4.5.3 (Дифференцированные услуги) сохранен как параграф 5.1.1.2.
    - Параграф 4.5.4 (MPLS) сохранен как параграф 5.1.3.3.
    - Параграф 4.5.5 (Показатели производительности IP) сохранен как параграф 5.1.3.6.
    - Параграф 4.5.6 (Измерение потоков) сохранен как параграф 5.1.3.7 с некоторым переформатированием.
    - Параграф 4.5.7 (Контроль перегрузок конечных точек) сохранен как параграф 5.1.3.8.
  - Параграф 4.6 (Обзор действий ITU, связанных с TE) удалён.
  - Параграф 4.7 (Распространение содержимого) сохранен как параграф 5.2.
- Раздел 5.0 (Таксономия систем TE) сохранен как раздел 4.
  - Параграф 5.1 (Управление в зависимости от времени и состояния) сохранен как параграф 4.1.
  - Параграф 5.2 (Автономные и интерактивные системы) сохранен как параграф 4.2.

- Параграф 5.3 (Централизованные и распределенные системы) сохранен как параграф 4.3 с дополнениями.
- Параграф 5.4 (Локальные и глобальные сведения) сохранен как параграф 4.4.
- Параграф 5.5 (Предписывающие о описательные системы) сохранен как параграф 4.5 с дополнениями.
- Параграф 5.6 (Системы с открытым и закрытым контуром) сохранен как параграф 4.6.
- Параграф 5.7 (Тактические и стратегические системы) сохранен как параграф 4.7.

- Раздел 6.0 (Рекомендации для Internet TE) сохранен как раздел 6.
  - Параграф 6.1 (Базовые нефункциональные рекомендации) сохранен как параграф 6.1.
  - Параграф 6.2 (Рекомендации по маршрутизации) сохранен как параграф 6.2 с редактированием.
  - Параграф 6.3 (Рекомендации по отображению трафика) сохранен как параграф 6.3.
  - Параграф 6.4 (Рекомендации по измерениям) сохранен как параграф 6.4.
  - Параграф 6.5 (Живучесть сети) сохранен как параграф 6.6.
    - Параграф 6.5.1 (Живучесть сети на основе MPLS) сохранен как параграф 6.6.1.
    - Параграф 6.5.2 (Варианты защиты) сохранен как параграф 6.6.2.
  - Параграф 6.6 (TE в среде Diffserv) сохранен как параграф 6.8 с редактированием.
  - Параграф 6.7 (Управляемость сети) сохранен как параграф 6.9.
- Раздел 7.0 (Междоменное взаимодействие) сохранен как раздел 7.
- Раздел 8.0 (Обзор современной практики TE в работающих сетях IP) сохранен как раздел 8.
- Раздел 9.0 (Заключение) удалён.
- Раздел 10.0 (Вопросы безопасности) сохранен как раздел 9 с существенным добавлением текста.

A.2. Этот документ

- Раздел 1 основан на разделе 1 из [RFC3272].
  - Параграф 1.1 основан на параграфе 1.1 из [RFC3272].
  - Параграф 1.2 написан заново.
  - Параграф 1.3 основан на параграфе 1.2 из [RFC3272].
  - Параграф 1.4 основан на параграфе 1.3 из [RFC3272].
- Раздел 2 основан на разделе 2 из [RFC3272].
  - Параграф 2.1 основан на параграфе 2.1 из [RFC3272].
  - Параграф 2.2 основан на параграфе 2.2 из [RFC3272].
  - Параграф 2.3 основан на параграфе 2.3 из [RFC3272].
    - Параграф 2.3.1 основан на параграфе 2.3.1 из [RFC3272].
  - Параграф 2.4 основан на параграфе 2.4 из [RFC3272].
    - Параграф 2.4.1 основан на параграфе 2.4.1 из [RFC3272].
  - Параграф 2.5 основан на параграфе 2.5 из [RFC3272].
- Раздел 3 основан на разделе 3 из [RFC3272].
  - Параграф 3.1 основан на параграфах 3.1, 3.2, 3.3, 3.4 из [RFC3272].
- Раздел 4 основан на разделе 5 из [RFC3272].
  - Параграф 4.1 основан на параграфе 5.1 из [RFC3272].
  - Параграф 4.2 основан на параграфе 5.2 из [RFC3272].
  - Параграф 4.3 основан на параграфе 5.3 из [RFC3272].
    - Параграф 4.3.1 написан заново.
    - Параграф 4.3.2 написан заново.
  - Параграф 4.4 основан на параграфе 5.4 из [RFC3272].
  - Параграф 4.5 основан на параграфе 5.5 из [RFC3272].
    - Параграф 4.5.1 написан заново.
  - Параграф 4.6 основан на параграфе 5.6 из [RFC3272].
  - Параграф 4.7 основан на параграфе 5.7 из [RFC3272].
- Раздел 5 основан на разделе 4 из [RFC3272].
  - Параграф 5.1 основан на параграфе 4.5 из [RFC3272].

- Параграф 5.1.1.1 основан на параграфе 4.5.1 из [RFC3272].
- Параграф 5.1.1.2 основан на параграфе 4.5.3 из [RFC3272].
- Параграф 5.1.1.3 написан заново.
- Параграф 5.1.1.4 написан заново.
- Параграф 5.1.1.5 написан заново.
- Параграф 5.1.2.1 написан заново.
- Параграф 5.1.2.2 написан заново.
- Параграф 5.1.2.3 написан заново.
- Параграф 5.1.3.1 основан на параграфе 4.4 из [RFC3272].
  - Параграф 5.1.3.1.1 написан заново.
- Параграф 5.1.3.2 основан на параграфе 4.5.2 из [RFC3272].
- Параграф 5.1.3.3 основан на параграфе 4.5.4 из [RFC3272].
- Параграф 5.1.3.4 написан заново.
- Параграф 5.1.3.5 написан заново.
- Параграф 5.1.3.6 основан на параграфе 4.5.5 из [RFC3272].
- Параграф 5.1.3.7 основан на параграфе 4.5.6 из [RFC3272].
- Параграф 5.1.3.8 основан на параграфе 4.5.7 из [RFC3272].
- Параграф 5.1.3.9 написан заново.
- Параграф 5.1.3.10 написан заново.
- Параграф 5.1.3.11 написан заново.
- Параграф 5.1.3.12 написан заново.
- Параграф 5.1.3.13 написан заново.
- Параграф 5.1.3.14 написан заново.
- Параграф 5.1.3.15 написан заново.

- Параграф 5.2 основан на параграфе 4.7 из [RFC3272].

- Раздел 6 основан на разделе 6 из [RFC3272].
  - Параграф 6.1 основан на параграфе 6.1 из [RFC3272].
  - Параграф 6.2 основан на параграфе 6.2 из [RFC3272].
  - Параграф 6.3 основан на параграфе 6.3 из [RFC3272].
  - Параграф 6.4 основан на параграфе 6.4 из [RFC3272].
  - Параграф 6.5 написан заново.
  - Параграф 6.6 основан на параграфе 6.5 из [RFC3272].
    - Параграф 6.6.1 основан на параграфе 6.5.1 из [RFC3272].
    - Параграф 6.6.2 основан на параграфе 6.5.2 из [RFC3272].
  - Параграф 6.7 написан заново.
  - Параграф 6.8 основан на параграфе 6.6 из [RFC3272].
  - Параграф 6.9 основан на параграфе 6.7 из [RFC3272].
- Раздел 7 основан на разделе 7 из [RFC3272].
- Раздел 8 основан на разделе 8 из [RFC3272].
- Раздел 9 основан на разделе 10 из [RFC3272].

Благодарности

Значительная часть текста этого документа заимствована из [RFC3272]. Редактор и создатели этого документа выражают свою признательность всем, кто принимал участие в работе. Несмотря на существенную переработку текста, авторы исходного документа, указанные ниже, должны считаться соавторами этой работы.

Daniel O. Awduche
Movaz Networks

Angela Chiu
Celion Networks

Anwar Elwalid
Lucent Technologies

Indra Widjaja
Bell Labs, Lucent Technologies

XiPeng Xiao
Redback Networks

Ниже приведён раздел благодарностей из [RFC3272]. Всем, кто помог в создании этого документа, нужно поблагодарить за вклад в новый документ.

> Авторы благодарны Jim Boyle за вклад в раздел рекомендаций, Francois Le Faucheur - за вклад об аспектах Diffserv, Blaine Christian - за вклад по измерениям, Gerald Ash - за вклад по маршрутизации в телефонных сетях и текст о методах TE по событиям, Steven Wright - за вклад по управляемости сети, Jonathan Aufderheide - за вклад по междоменному TE с BGP. Отдельная благодарность Randy Bush за предложение таксономии TE на основе сравнения тактических и стратегических методов. Параграф, описывающий действия ITU, связанные с TE был создан на основе предложений Waisum Lai. Полезные отклики и ссылки на соответствующие материалы были предоставлены J. Noel Chiappa. Дополнительные комментарии предоставил Glenn Grotefeld во время рабочих встреч процесса last call. Кроме того, авторы благодарны Ed Kern, сопредседателю TEWG, за комментарии и поддержку.

Ранние черновые варианты этого документа были подготовлены командой разработчиков RFC3272bis в рамках рабочей группы TEAS. Полный список членов команды приведён ниже.

Acee Lindem
Adrian Farrel
Aijun Wang
Daniele Ceccarelli
Dieter Beller
Jeff Tantsura
Julien Meuric
Liu Hua
Loa Andersson
Luis Miguel Contreras
Martin Horneffer
Tarek Saad
Xufeng Liu

Этот документ содержит исправление оригинального текста, отмеченное в информации об ошибке [Err309] от Jean-Michel Grimaldi.

Редактор документа благодарит также Dhruv Dhody, Gyan Mishra, Joel Halpern, Dave Taht, John Scudder, Rich Salz, Behcet Sarikaya, Bob Briscoe, Erik Kline, Jim Guichard, Martin Duke, Roman Danyliw за рецензии.

Эта работа частично поддерживалась Европейской комиссией в рамках гранта Horizon 2020 по соглашению 101015857 Secured autonomic traffic management for a Tera of SDN flows (Teraflow).

Участники работы

Ниже указаны люди, внёсшие существенный вклад в этот документ.

Gert Grammel
Email: ggrammel@juniper.net

Loa Andersson
Email: loa@pi.nu

Xufeng Liu
Email: xufeng.liu.ietf@gmail.com

Lou Berger
Email: lberger@labn.net

Jeff Tantsura
Email: jefftant.ietf@gmail.com

Daniel King
Email: daniel@olddog.co.uk

Boris Hassanov
Email: bhassanov@yandex-team.ru

Kiran Makhijani
Email: kiranm@futurewei.com

Dhruv Dhody
Email: dhruv.ietf@gmail.com

Mohamed Boucadair
Email: mohamed.boucadair@orange.com

Адрес автора

Adrian Farrel (editor)
Old Dog Consulting
Email: adrian@olddog.co.uk

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru