Network Working Group
Request for Comments: 3272
Category: Informational

D. Awduche
Movaz Networks
A. Chiu
Celion Networks
A. Elwalid
I. Widjaja
Lucent Technologies
X. Xiao
Redback Networks

May 2002

Принципы организации трафика Internet

Overview and Principles of Internet Traffic Engineering

Статус документа

Этот документ содержит информацию для сообщества Internet и не определяет каких-либо стандартов. Документ может распространяться свободно.

Авторские права

Аннотация

Этот документ описывает принципы организации (построения, формирования) трафика (Traffic Engineering или TE) с сети Internet. Документ служит целям обеспечения более глубокого понимания вопросов организации трафика IP и может служить основой для разработки систем управления трафиком Internet. В документе рассматриваются принципы, архитектура и методология оценки и оптимизации производительности сетей IP.

Оглавление

1.0 Введение

В этом документе рассматриваются принципы организации трафика Internet. Целью этого документа является разработка общих принципов организации трафика Internet и (в тех случаях, когда это приемлемо) обеспечение рекомендаций, руководств и вариантов для развития возможностей организации (построения) трафика Internet и систем поддержки.

Этот документ может помочь сервис-провайдерам при разработке и реализации решений по организации трафика в их сетях. Производители сетевого оборудования и программ также найдут этот документ полезным при разработке механизмов и систем поддержки для среды Internet, обеспечивающих функции организации трафика.

В этом документе приведена терминология, требуемая для понимания и описания концепций построения трафика Internet. Документ включает классификацию известных стилей организации трафика. В контексте документа стиль организации трафика представляет собой абстракцию важных аспектов методологии построения трафика. Стили организации трафика могут рассматриваться с разных точек зрения в зависимости от конкретного контекста и решаемых задач. Сочетание стилей и точек зрения приводит к естественной классификации систем организации трафика.

Хотя максимальная эффективность систем организации трафика обеспечивается при их сквозном применении, данный документ в основном фокусируется на построении трафика в масштабе отдельного домена (т. е., трафика одной автономной системы). Однако из-за преобладания в Internet междоменного (начинающегося в одной AC и завершающегося в другой) трафика данный документ включает обзор аспектов организации междоменного трафика.

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не нужно** (SHALL NOT), **следует** (SHOULD), **не следует** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе должны интерпретироваться в соответствии с RFC 2119.

1.1. Что такое организация трафика Internet?

Организация трафика Internet определяется, как аспект построения сети Internet, связанный с оценкой и оптимизацией производительности действующих сетей IP. Организация трафика включает применение технологических и научных принципов к измерению, описанию, моделированию и управлению трафиком Internet [RFC-2702, AWD2].

Повышение производительности работы сети (на уровне трафика и ресурсов) является основной целью построения трафика Internet. Это достигается путём выполнения ориентированных на трафик требований к производительности при обеспечении экономичного и надёжного использования сетевых ресурсов. Ориентированные на трафик характеристики включают задержку, вариации задержки, частоту потери пакетов и пропускную способность.

Важной целью формирования трафика Internet является обеспечение надёжного функционирования сети [RFC-2702]. Надёжность работы сети может быть достигнута за счёт обеспечения механизмов, повышающих уровень целостности сети, а также её устойчивость. Это ведёт к минимизации уязвимости сети к отказам, связанным с ошибками, сбоями и повреждениями, возникающими в инфраструктуре.

Internet существует для передачи информации от узлов-источников к узлам-приемникам. В соответствии с этой задачей одной из наиболее важных функций Internet является маршрутизация трафика от входных узлов к выходным. Следовательно, наиболее значимой функцией формирования трафика Internet является контроль и оптимизация функций маршрутизации, чтобы направить трафик через сеть наиболее эффективным путём.

В конечном итоге наиболее важной задачей является производительность, доступная для конечных пользователей сетевого сервиса. Это обстоятельство следует принимать во внимание при разработке всех механизмов и правил организации трафика. Характеристики, видимые конечным пользователям, определяются новыми свойствами сети, которые присущи сети в целом. Основной целью сервис-провайдеров является, следовательно, повышение качества сетей с учётом экономических параметров.

Важность отмеченного выше наблюдения в части новых свойств сети состоит в том, что нужно быть особенно осторожными при выборе параметров производительности для оптимизации. Оптимизация не тех параметров может дать локальный позитивный эффект, в то же время оказывая катастрофическое влияние на общие свойства сети и качество предоставляемых конечным пользователям сетевых услуг.

Трудно уловимые, но практически важные результаты систематического применения концепций организации трафика в работающих сетях заключаются в том, что это помогает определить и структурировать цели и приоритеты в терминах повышения качества обслуживания конечных пользователей сетевого сервиса. Применение концепций организации трафика полезно также при оценке и анализе достижения поставленных целей.

Оптимизационные аспекты организации трафика могут быть реализованы за счёт управления возможностями и трафиком. В этом документе управление возможностями включает планирование возможностей, контроль маршрутизации и управление ресурсами. К числу представляющих наибольший интерес ресурсов относятся полоса пропускания каналов, буферная ёмкость и вычислительные ресурсы. В контексте этого документа управление трафиком включает (1) узловые функции контроля трафика (кондиционирование, управление очередями, планирование и т. п.) и (2) другие функции, которые регулируют поток трафика через сеть или распределение сетевых ресурсов между пакетами или потоками трафика.

Оптимизационные задачи организации трафика Internet следует рассматривать, как непрерывный итеративный процесс, а не простое достижение одноразовой цели. Организация трафика также требует концептуальной разработки новых технологий и методологии повышения производительности сети.

Задачи оптимизации в построении трафика Internet могут меняться с течением времени по мере возникновения новых требований, разработки новых технологий и появления новых идей, оказывающих влияние на решение основополагающих задач. Более того, в разных сетях могут возникать разные оптимизационные задачи в зависимости от бизнес-моделей, возможностей сети и эксплуатационных ограничений. Оптимизационные аспекты организации трафика в конечном счёте диктуются управлением сетью, независимо от конкретных задач оптимизации в любой конкретной среде.

Таким образом, оптимизационные аспекты организации трафика могут рассматриваться с точки зрения управления. Аспект управления на арене организации трафика Internet может быть превентивным или реактивным. В превентивном варианте управляющая система организации трафика предпринимает упреждающие действия, чтобы избежать возникновения в будущем нежелательных состояний в сети. Могут также приниматься меры по индуцированию более желательных состояний в будущем. В реактивном варианте система управления корректно и, возможно, адаптивно реагирует на происходящие в сети события.

Управляющие аспекты организации трафика Internet реализуются на разных временных уровнях относительно происходящих в сети событий. Некоторые аспекты управления возможностями (например, планирование) связаны с реакцией только на долговременном уровне (от дней до лет). Появление оптических транспортных сетей с автоматической коммутацией (например, на базе концепции мультипротокольной коммутации длин волн[1]) может привести к существенному сокращению жизненного цикла планирования возможностей за счёт использования полосы в оптических кабелях. Функции контроля маршрутизации работают на промежуточных уровнях временного разрешения (от миллисекунд до дней). Функции уровня обработки пакетов (например, управление скоростью передачи, управления очередями, планирование очередей) работают на уровне очень коротких временных интервалов (от пикосекунд до миллисекунд) при откликах на статистическое поведение трафика в реальном масштабе времени. Управляющая подсистема организации трафика Internet включает функции расширения возможностей, управления маршрутизацией, управления трафиком и управления ресурсами (включая управление политикой обслуживания на элементах сети). При необходимости расширения возможностей в практических целях может оказаться желательной разработка плана развёртывания, что позволит ускорить предоставление требуемой полосы вкупе с минимизацией инсталляционных расходов.

Входные данные системы управления организацией трафика включают переменные состояния сети, переменные политики, а также переменные, связанные с принятием решений.

Одним из основных преимуществ организации трафика Internet является реализация автоматизированного контроля возможностей, который достаточно быстро и экономично адаптируется к существенным изменениям состояния сети при сохранении её стабильной работы.

Другой критически важной функцией организации трафика Internet является оценка производительности сети, которая важна для понимания эффективности методов организации трафика, а также для мониторинга и проверки соответствия планам развития сети. Результаты оценки производительности могут использоваться для обнаружения имеющихся проблем, в качестве руководства по оптимизации сети и для предсказания возможных проблем.

Оценка производительности может быть выполнена разными способами. К наиболее распространенным методам относят аналитические методы, имитацию (моделирование) и эмпирические методы. При использовании аналитических методов или имитации сетевые узлы и каналы могут моделироваться с использованием эксплуатационных параметров сети (таких, как топологии, полосы пропускания, буферной ёмкости, правил обслуживания на узлах - планирование каналов, приоритизация пакетов, управление буферами и т. п.). Аналитические методы могут служить для описания динамических и поведенческих характеристик трафика (размер пиков, статистическое распределение, зависимости).

Оценка производительности в практическом контексте работающей сети может оказаться достаточно сложной. Для упрощения анализа может использоваться множество методов, включаю абстракцию, декомпозицию и аппроксимацию. Например, концептуальные упрощения типа эффективной полосы или эффективного буфера [Elwalid] могут использоваться при аппроксимации поведения узла на уровне пакетов или анализе на уровне соединения. Использование при анализе моделей очередей и схем аппроксимации на основе асимптотического разложения и декомпозиции могут дополнительно упростить ситуацию и сделать результаты анализа более понятными. В частности, новый набор концепций, известный, как сетевые вычисления [CRUZ], и основанный на детерминированных границах, может упростить анализ сети по сравнению со случаем использования классических стохастических методов. При использовании аналитических методов следует осторожно выбирать модели, чтобы они достаточно точно отражали соответствующие характеристики моделируемых сетевых элементов.

Имитация может использоваться для оценки производительности сети или проверки аналитических аппроксимаций. Однако имитация, сама по себе, может потребовать значительных вычислительных ресурсов и не всегда обеспечивает

[1]Multi-protocol Lambda Switching.

желаемые результаты. методика, подходящая для анализа проблем производительности конкретной сети может включать комбинацию аналитических, имитационный и эмпирических методов.

Как правило, концепции и механизмы организации трафика должны быть достаточно конкретны и чётко определены при достаточном уровне гибкости и расширяемости в соответствии с непредвиденными потребностями в будущем.

1.2. Рассматриваемые вопросы

Этот документ посвящён организации внутридоменного трафика, т.е., трафика внутри одной автономной системы Internet. В документе рассматриваются концепции, относящиеся к контроля трафика внутри домена, включая такие вопросы, как маршрутизация, выделение ресурсов на макро- и микроуровне, а также координация решения проблем, которые могут возникать.

В документе приведены описания и характеристики методов, которые уже используются или находятся в процессе разработки. Обсуждаются также способы совместного использования разных методов и сценарии, в которых те или иные методы могут быть полезны.

При рассмотрении различных моделей внутридоменного трафика основное внимание уделяется организации трафика на базе MPLS. Построение трафика на основе манипуляций с метрикой IGP подробно не рассматривается. Этот вопрос может быть детально рассмотрен в документах других рабочих групп.

Хотя документ сконцентрирован на рассмотрении внутридоменного трафика, в разделе 7.0 даётся обзор организации междоменного трафика. Организация междоменного трафика Internet имеет очень важное значение для производительности инфраструктуры Internet в целом.

В документе (там, где это возможно) даются ссылки на связанные с темой документы IETF и других организаций.

1.3 Терминология

В этом параграфе описана терминология, используемая при обсуждении организации трафика Internet. Приведённые определения терминов значимы в пределах данного документа, а в иных документах может использоваться другая трактовка.

Baseline analysis - базовый анализ
Исследование, выполненное в качестве основы для сравнения с реальным поведением сети.

Busy hour - час с наибольшей нагрузкой
Период в 1 час из заданного интервала времени (обычно суток), в течение которого нагрузка на сеть или подсеть является максимальной.

Bottleneck - узкое место (горлышко бутылки)
Элемент сети, для которого входная скорость превышает выходную.

Congestion - насыщение (перегрузка)
Состояние сетевого ресурса, при котором трафик для этого ресурса превышает его выходные возможности в данный момент.

Congestion avoidance - предотвращение насыщения
Модель контроля насыщения, которая пытается предотвратить возникновение перегрузок.

Congestion control - контроль насыщения
Модель обработки перегрузок, которая пытается ослабить уровень уже возникшего насыщения.

Constraint-based routing - принудительная (основанная на ограничениях) маршрутизация
Класс протоколов маршрутизации, принимающих во внимание при определении маршрута заданные атрибуты трафика, сетевые ограничения и правила. Принудительная маршрутизация применима как к агрегатам трафика, так и к потокам. Эта модель маршрутизации является обобщением маршрутизации на основе QoS.

Demand side congestion management - контроль насыщения со стороны потребителя
Схема контроля насыщения, в которой проблема перегрузки решается путём регулирования или кондиционирования предлагаемой нагрузки.

Effective bandwidth - эффективная полоса пропускания
Минимальная полоса, которая может быть выделена для потока или агрегата трафика, чтобы обеспечить приемлемое качество обслуживания для потока или агрегата трафика.

Egress traffic - выходной трафик
Трафик, выходящий из сети или элемента сети.

Hot-spot - горячая точка
Элемент сети или подсистема, находящиеся в состоянии насыщения.

Ingress traffic - входной трафик
Трафик, входящий в сеть или элемент сети.

Inter-domain traffic - междоменный трафик
Трафик, генерируемый в одной автономной системе и принимаемый в другой АС.

Loss network - сеть с потерями
Сеть, которая не обеспечивает адекватной буферизации трафика, в результате чего трафик, обращающийся (входящий) к загруженному ресурсу сети будет отбрасываться, а не помещаться в очередь.

Metric - метрика
Параметр, определённый в терминах стандартных единиц измерения.

Measurement Methodology - методология измерений
Повторяемая методика измерения, применяемая для одной или множества метрик.

Network Survivability - живучесть сети
Способность обеспечивать предписанный уровень QoS для существующего сервиса после возникновения в сети заданного числа отказов.

Offline traffic engineering - внешняя организация трафика
Система организации трафика, находящаяся за пределами сети.

Online traffic engineering - внутренняя организация трафика
Система организации трафика внутри сети, реализуемая обычно на элементах работающей сети или дополнениях к таким элементах.

Performance measures - измерение производительности
Метрика, обеспечивающая количественные или качественные показатели производительности интересующей системы или подсистемы.

Performance management - контроль производительности
Систематическое повышение эффективности для достижения заданных целей в плане повышения производительности.

Performance Metric - метрика производительности
Параметр производительности, определяемый в терминах стандартных единиц измерения.

Provisioning - обеспечение
Процесс выделения или настройки сетевых ресурсов в соответствии с некими запросами.

QoS routing - маршрутизация на основе QoS
Класс систем маршрутизации, в которых выбор пути для потока осуществляется на основе требований по качеству обслуживания (QoS0 этого потока.

Service Level Agreement - соглашение об уровне обслуживания
Соглашение между оператором и заказчиком, гарантирующее заданные уровни производительности и надёжности при определённой стоимости.

Stability - стабильность
Эксплуатационное состояние, в котором сеть не осциллирует в деструктивной манере из одного режима в другой.

Supply side congestion management - контроль насыщения со стороны поставщика
Схема контроля насыщения, обеспечивающая дополнительные сетевые ресурсы для решения возникших и/или намечающихся проблем, связанных с перегрузкой.

Transit traffic - транзитный трафик
Трафик, адресат и отправитель которого находятся за пределами рассматриваемой сети.

Traffic characteristic - характеристики трафика
Описание временных изменений или атрибутов данного потока трафика или агрегата трафика.

Traffic engineering system - система организации трафика
Набор объектов, механизмов и протоколов, совместно используемых для решения задач организации трафика.

Traffic flow - поток трафика
Поток пакетов между двумя оконечными точками, который может быть тем или иным способом охарактеризован. Микропоток имеет более конкретное определение - поток пакетов с одинаковыми значениями адресов и портов отправителя и получателя, а также одинаковыми идентификаторами протоколов.

Traffic intensity - интенсивность (уровень) трафика
Мера нагрузки от трафика относительно ёмкости ресурсов для заданного периода времени. В классических телефонных системах интенсивность трафика измеряется в Эрлангах (Erlang).

Traffic matrix - матрица трафика
Представление потребностей в трафике между набором абстрактных источников и получателей. Абстрактный узел может включать один или множество сетевых элементов.

Traffic monitoring - мониторинг трафика
Процесс наблюдения характеристик трафика в заданной точке сети и сбора данных о трафике для последующего анализа или иных действий.

Traffic trunk - транк трафика
Агрегат потоков трафика, относящихся к одному классу и передаваемых по одному пути. Транк может характеризоваться входным и выходным узлами, а также набором атрибутов, которые определяют требования и параметры поведения.

2.0 Основы

Сеть Internet быстро развилась в коммуникационную инфраструктуру, играющую существенную роль в экономике, образовании и социальной сфере. В то же время на рынке предоставления коммуникационных услуг наблюдается высокая конкуренция и конечные пользователи запрашивают у своих сервис-провайдеров высокое качество обслуживания. Следовательно, оптимизация производительности крупномасштабных сетей IP (особенно, магистральных) становится важнейшей проблемой. Требования к производительности сетей являются многомерными, сложными и, зачастую, противоречивыми, что существенно осложняет решение проблемы организации трафика.

Сеть должна передавать пакеты от входных узлов к выходным эффективно, оперативно и экономично. Следовательно в среде с множеством классов услуг (например, в сетях с поддержкой Diffserv) параметры совместного использования ресурсов должны подобающим образом определяться и настраиваться в соответствии с принятой политикой и моделями обслуживания, чтобы не возникало проблем состязания за ресурсы между разными пакетами, проходящими через сеть. Таким образом, требуется обратить внимание на распределение сетевых ресурсов между потоками трафика одного класса (внутриклассовая борьба за ресурсы) и потоками разных классов (межклассовая борьба за ресурсы).

2.1 Контекст организации трафика Internet

Контекст организации трафика Internet связан со сценариями, где используется построение трафика. Методология организации трафика устанавливает соответствующие правила для решения проблем производительности, возникающих в конкретном контексте. Контекст организации трафика Internet включает указанные ниже части.

1) *Сетевой контекст*, определяющий поле деятельности (в частности, ситуации, где возникают проблемы организации трафика). Сетевой контекст включает структуру сети, правила, характеристики, ограничения, атрибуты качества и критерии оптимизации сетей.
2) *Контекст проблемы*, определяющий общие и частные проблемы, которые решает организация трафика. Данный контекст включает идентификацию, абстракцию связанных с проблемой признаков, представление, формулировку, спецификацию требований к пространству решений и спецификацию желаемых признаков подходящего решения.
3) *Контекст решения*, предлагающий способы решения проблем, идентифицированных в контексте проблемы. Этот контекст включает анализ, оценку вариантов, «рецепт» и решение.
4) Контекст реализации и эксплуатации, в котором используется экземпляр решения. В данный контекст включается планирование, организация и исполнение.

Контекст организации трафика Internet и различные сценарии проблем рассматриваются в остальной части раздела.

2.2 Сетевой контекст

Сети IP варьируются по размеру от небольших кластеров маршрутизаторов, расположенных в одном месте, до тысяч связанных между собой маршрутизаторов, коммутаторов и других компонент, распределенных по всему миру.

Концептуально, на базовом уровне абстракции сеть IP может быть представлена, как распределенная динамическая система, состоящая из (1) набора соединённых между собой ресурсов, обеспечивающих транспортировку трафика IP, с некоторыми ограничениями, (2) системы запросов, предлагающей нагрузку для доставки через сеть и (3) системы откликов, включающей сетевые процессы, протоколы и механизмы, которые способствуют перемещению трафика через сеть [см. также AWD2].

Сетевые элементы и ресурсы могут иметь специфические характеристики, которые могут ограничивать методы обслуживания запросов. Кроме того, сетевые ресурсы могут быть оборудованы механизмами контроля трафика, управляющими способами обслуживания запросов. Механизмы контроля трафика могут, например, использоваться для управления обработкой пакетов в рамках данного ресурса, распределения доступа к ресурсу для разных пакетов и регулировки поведения трафика, проходящего через ресурс. Системы обеспечения и управления конфигурацией могут разрешать манипулирование механизмами контроля трафика внутренним или внешним элементам для осуществления контроля за реакцией элементов сети на внутренние или внешние изменения (события).

Детали предоставления сеть транспортных услуг для пакетов задаются в политике, устанавливаемой сетевыми администраторами и реализуемой с помощью системы управления конфигурацией сети и политики обеспечения. В общем случае типы сервиса, предоставляемого сетью, зависят также от технологий и характеристик сетевых элементов и протоколов, преобладающих моделей обслуживания и способности сетевых администраторов переносить политику в конфигурацию сети.

Современным сетям Internet присущи три важных характеристики: (1) предоставление услуг в реальном масштабе времени, (2) критическая важность и (3) быстро меняющаяся рабочая среда. Динамичность характеристик сетей IP отчасти можно объяснить изменениями спроса, взаимодействиями между различными протоколами и процессами, быстрыми изменениями инфраструктуры, требующие постоянного включения новых технологий и элементов, а также временные и постоянные повреждения в системах.

При передаче пакетов через сеть между ними возникает «конкуренция» за использование ресурсов. Сетевой ресурс рассматривается, как перегруженный (насыщенный), если скорость прибытия пакетов превышает выходную «ёмкость» ресурса в течение некого интервала времени. Перегрузка может приводить к задержке или отбрасыванию некоторых прибывающих пакетов.

Насыщение увеличивает транзитные задержки, вариации задержек, уровень потери пакетов и снижает предсказуемость сетевого сервиса. Очевидно, что насыщение является крайне нежелательным явлением.

Предотвращение перегрузок за разумные средства является одной из задач построения трафика Internet.

Эффективное совместное использование сетевых ресурсов для множества потоков трафика является основным экономическим преимуществом сетей с коммутацией пакетов в целом и Internet, в частности. Одной из основных операционных задач для сетей (особенно для больших сетей IP общего пользования) является повышение эффективности использования ресурсов при минимизации возможных перегрузок (насыщения).

В Internet становится все больше разных классов трафика с разными требованиями по обслуживанию. Появление дифференцированных услуг [RFC-2475] дополнительно обостряет эти требования. Таким образом, пакеты могут группироваться в поведенческие агрегаты, каждому из которых будет присуще определённый набор характеристик поведения или параметров доставки. На практике требования по доставке конкретного множества пакетов могут выражаться явно или неявно. Два наиболее важных параметра доставки трафика связаны с ограничениями возможностей (capacity) и QoS.

Ограничения возможностей могут выражаться статистически через пиковую и среднюю скорость, величину всплесков или тем или иным детерминированным указанием эффективной пропускной способности. Требования QoS могут выражаться в терминах (1) ограничения целостности (например, потери пакетов) и временные ограничения (например, задержки отдельных пакетов и вариации задержки последовательных пакетов одного потока).

2.3 Контекст проблемы

Существует фундаментальная проблема связанная с работой сети, описываемой простой моделью из предыдущего параграфа. В данном параграфе рассматривается контекст проблемы в связи с функцией построения трафика.

Идентификация, абстрагирование, представление и измерение параметров сети, относящихся к организации трафика, являются важными вопросами.

Одна из важных задач связана с явным формулированием проблем, которые пытается решить организация трафика, а также способ идентификации требований к пространству решений и выбора желательного набора параметров для хорошего решения. Важны также вопросы измерения и характеризации эффективности решений.

Другой класс проблем связан с измерением и оценкой относящихся к делу параметров сети. Эффективная организация трафика опирается на правильную оценку предлагаемой нагрузки, а также представления лежащих в основе сетевых технологий и связанных с ресурсами ограничений. Представление топологии в масштабах сети также должно влиять на планирование.

Ещё один класс задач относится к характеристикам состояний сети и способам оценки производительности в разных условиях. Задача оценки производительности делится на две части. Один аспект проблемы связан с оценкой системного уровня производительности сети. Другой аспект относится к оценке производительности на уровне ресурсов и связан с оценкой продуктивности работы отдельных сетевых ресурсов. В этом документе характеристики системного уровня сети будем называть макро-состояниями, а характеристики уровня ресурсов - микро-состояниями. Характеристики системного уровня называют также вторичными (emergent), как отмечено выше. Следовательно, схемы организации трафика, имеющие дело с оптимизацией производительности на системном уровне, будут называться макро-TE, а схему, оптимизирующие на уровне отдельных ресурсов - микро-TE. В некоторых случаях

производительность на системном уровне может быть выведена из характеристик на уровне ресурсов путём применения подходящих правил композиции в зависимости от интересующей метрики производительности.

Другой класс фундаментальных проблем связан со способами эффективной оптимизации сетевой производительности. Оптимизация может перенести решение некоторых конкретных проблем организации трафика на уровень настройки конфигурации сети. Оптимизация может также обеспечить некоторый контроль над управлением ресурсами, маршрутизацией и/или увеличение пропускной способности.

Как отмечено выше, насыщение (перегрузка) является нежелательным явлением в сетях. По этой причине в следующем параграфе рассматриваются вопросы, связанные с перегрузками и их последствиями в контексте организации трафика Internet.

2.3.1 Насыщение и его последствия

Насыщение (перегрузка) является одной из наиболее значимых проблем в контексте работы сетей IP. Насыщение сетевого элемента означает наличие на нем перегрузки в течение некоторого интервала времени. Перегрузка практически всегда приводит к ухудшению обслуживания конечных пользователей. Схемы контроля перегрузок могут включать политику сайтов-потребителей и сайтов-поставщиков. Политика потребителей может ограничивать доступ к перегруженным ресурсам и/или динамически регулировать запросы для преодоления перегрузки. Политика поставщиков может сужать или расширять пропускную способность для более эффективной передачи предлагаемого трафика. Такая политика может также перераспределять сетевые ресурсы путём перенаправления трафика в сетевой инфраструктуре. Перераспределение трафика и ресурсов служат для увеличения «эффективной пропускной способности» для потребителей.

Этот документ акцентирован прежде всего на схемы контроля перегрузок, относящиеся к работе сети, а не к управлению запросами конечных потребителей. Т. е., рассматриваемые в этом документе аспекты контроля перегрузок и предлагаемые решения относятся к элементам сети и действиям сетевых администраторов.

2.4 Контекст решения

Контекст решения по организации трафика Internet включает анализ, оценку вариантов и выбор действия. Обычно контекст решения основан на представлениях о текущем или будущем состоянии сети и последующем принятии решения, которое может включать выбор наиболее предпочтительного из нескольких вариантов действий. Более конкретно, контекст решения требует оценки трафика, определения состояния сети и поиск решения по организации трафика, которое может формулироваться явно или неявно и может также включать управляющие воздействия. Такие воздействия могут включать манипуляции с параметрами маршрутизации, расширение «мощностей» и применение функций управления трафиком.

Ниже приведён список инструментов, применимых в контексте решения по организации трафика Internet.

(1) Набор правил, целей и требований (могут зависеть от контекста) для оценки и оптимизации производительности сети.

(2) Набор инструментов и механизмов для измерения, описания, моделирования и контроля трафика Internet, размещения и распределения сетевых ресурсов, а также контроля за отображением или распределением трафика в инфраструктуре.

(3) Набор ограничений для рабочей среды, сетевых протоколов и организации трафика, как таковой.

(4) Набор качественных и количественных механизмов и методологий для абстрагирования, формулировки и решения задач организации трафика.

(5) Набор параметров администрирования, которыми можно управлять через систему управления конфигурацией (CM[1]). Сама система CM может включать подсистему контроля конфигурации, репозиторий конфигураций, подсистемы учёта и аудита конфигураций.

(6) Набор рекомендаций по оценке производительности сети, её оптимизации и повышению.

Определение характеристик трафика на основе измерений и/или оценок очень полезно при поиске решений по организации трафика. Оценки трафика могут основываться на информации от пользователей, прогнозов и моделей трафика, а также на результатах реальных измерений. Измерения могут проводиться на уровне агрегированного трафика или для отдельных потоков, чтобы получить статистику с разными уровнями детализации. Измерения на уровне потоков или небольших агрегатов трафика могут проводиться на граничных узлах, где трафик входит в сеть или выходит из неё. Измерения на уровне крупных агрегатов трафика могут выполняться в ядре сети, где значительное число потоков может передаваться одновременно.

Для проведения исследований производительности и поддержки планирования существующих и будущих сетей может быть выполнен анализ маршрутизации с целью определения путей выбора протоколов маршрутизации с учётом меняющихся потребностей трафика, а также определение сетевых ресурсов, используемых при маршрутизации трафика через сеть. Анализ маршрутизации должен включать вопросы выбора путей через сеть, распределения трафика через множество доступных маршрутов и мультиплексирования трафика IP через транки (если такие конструкции существуют), а также нижележащую сетевую инфраструктуру. Для анализа маршрутизации требуется модель сетевой топологии. Эта модель может быть получена на основе документов по архитектуре сетей и их устройству, конфигурационных файлов маршрутизаторов, баз данных по маршрутизации, маршрутных таблиц, а также автоматических средств обнаружения и отображения сетевых топологий. Топологическая информация может быть также получена от серверов мониторинга сетей и серверов обеспечения.

Маршрутизация в сетях IP может административно контролироваться на разных уровнях абстракции, включая атрибуты BGP и изменение метрики IGP. Для ориентированных на пути технологий типа MPLS маршрутизация может дополнительно контролироваться путём изменения параметров организации трафика, параметров ресурсов и административных ограничений. В контексте MPLS путь с явной коммутацией по меткам (LSP) можно рассчитать и организовать разными способами: (1) вручную, (2) автоматически (online) с использованием основанного на

[1]Configuration Management.

ограничениях процесса маршрутизации, реализованного на маршрутизаторах с коммутацией по меткам или (3) автоматически (offline) с помощью основанных на ограничениях элементах маршрутизации во внешних системах поддержки организации трафика.

2.4.1 Решение проблемы перегрузок

Минимизация перегрузок является важным аспектом организации трафика Internet. В этом параграфе даётся обзор основных вариантов решения проблемы перегрузок.

Вопросы контроля перегрузок можно разделить на несколько категорий (в работе [YARE95] схемы контроля перегрузок рассмотрены более подробно): (1) время реализации (быстро, средне, медленно), (2) реактивные и превентивные меры контроля и предотвращения перегрузок, (3) контроль на стороне поставщиков и потребителей. Эти аспекты более подробно рассматриваются в последующих параграфах.

(1) Контроль перегрузок по продолжительности решений

- Долгосрочные решения (от недель до месяцев). Планирование сетевых мощностей на достаточно продолжительный срок основывается на оценках или предсказаниях потребностей в трафике и его распределении. Поскольку организация каналов и установка маршрутизаторов требует времени, обновления обычно длятся недели и месяцы, а иногда - годы.

- Среднесрочные решения (дни). Некоторые правила попадают в среднесрочную категорию. Примеры таких правил включают: (1) подстройку параметров IGP и/или BGP для маршрутизации трафика некоторых сегментов сети; (2) организацию и настройку некоторых явных путей с коммутацией по меткам (ER-LSP[1]) в сетях MPLS для отвода транков трафика за пределы перегруженных участков или перевода на предпочтительные пути; (3) перенастройка логической топологии сети с учётом пространственного распределения трафика (например, за счёт использования на базовых уровнях ориентированных на пути технологий типа MPLS LSP, ATM PVC или группы оптических каналов). Многие из таких адаптивных схем основаны на системах измерения, осуществляющих мониторинг распределения трафика и загрузки сетевых ресурсов с передачей информации механизмам организации трафика. Механизмы и средства организации трафика могут быть распределенными или централизованными, а их структура может быть плоской или иерархической. Сравнительные метрики для распределенных и централизованных систем управления известны достаточно хорошо. Централизованные системы могут иметь глобальную область действия и обеспечивать более оптимальные решения. Однако такие системы менее отказоустойчивы по сравнению с распределенными схемами. Более того, используемая централизованной схемой информация может быть устаревшей и не соответствующей реальному состоянию сети. Рекомендации по выбору централизованной или распределенной схемы управления выходят за рамки данного документа. Этот выбор сетевые администраторы должны делать сами с учётом своих реальных потребностей.

- Краткосрочные решения (от пикосекунд до минут). Эта категория включает функции обработки на уровне пакетов и события со сроком действия порядка нескольких периодов кругового обхода. Сюда входят механизмы маршрутизаторов для активного и пассивного управления буферами. Эти механизмы служат для контроля перегрузок и/или информирования о них конечных систем, которые могут в результате адаптивно менять скорость передачи трафика в сеть. Одним из наиболее популярных (особенно для TCP) механизмов является RED[2] [FLJA93], который предотвращает перегрузки путём контроля среднего размера очереди. Во время перегрузки (но до заполнения очереди) схема RED выбирает приходящие пакеты для «маркировки» в соответствии с вероятностным механизмом, принимающим во внимание средний размер очереди. Для маршрутизаторов, не использующий явных уведомлений о перегрузке (ECN[3]) (см., например, [FLOY94]), отмеченные пакеты могут просто отбрасываться для сигнализации перегрузки конечным системам. С другой стороны, если маршрутизатор поддерживает ECN, он может установить поле ECN в заголовке пакета. Были предложены варианты RED с поддержкой разных уровней предпочтения при отбрасывании в средах с множеством классов трафика [RFC-2597] (например, RIO[4] и Weighted RED). Принято считать, что RED предотвращает перегрузки не хуже, чем управление очередями TD[5] (отбрасывание прибывающих пакетов при заполнении очереди). Важно отметить, что RED снижает вероятность глобальной синхронизации и более беспристрастен к разным сессиям TCP. Однако сам по себе RED не может предотвратить перегрузок и не обеспечивает беспристрастности при наличии отправителей, не воспринимающих RED (например, трафик UDP и некоторые реализации с ошибками). Были предложены другие схемы повышения производительности и уровня беспристрастности в присутствии «безответственного» трафика. Некоторые из таких схем были предложены, как теоретические модели и не всегда доступны в существующих коммерческих системах. Двумя такими схемами являются LQD[6] и RND[7] [SLDC98].

(2) Превентивные и реактивные меры

- Реактивный контроль перегрузок (восстановление) заключается в реагировании на возникшие проблемы с перегрузками. Большинство перечисленных выше долгосрочных и среднесрочных мер может быть отнесено к категории реактивных, особенно в тех случаях, когда политика основывается на мониторинге и идентификации возникающих перегрузок, включая соответствующие действия по облегчению ситуации.

- Превентивные (предсказание/предотвращение) меры ставят своей целью предотвращение перегрузок на основе оценок и прогнозов. Некоторые упомянутые выше долгосрочные и среднесрочные меры относятся к категории превентивных. Они не обязательно являются откликом на возникшие проблемы насыщения. Прогнозы потребностей в трафике и распределения нагрузки могут предотвратить возникновение перегрузок в будущем. Схемы, упомянутые в краткосрочных мерах (например, RED и варианты, ECN, LQD,

[1] Explicitly routed label switched path – явно маршрутизируемые пути с коммутацией по меткам.
[2] Random Early Detection - произвольное раннее обнаружение.
[3] Explicit congestion notification.
[4] RED with In and Out .
[5] Tail-Drop - «обрубание хвостов».
[6] Longest Queue Drop - отбрасывание длиннейшей очереди.
[7] Dynamic Soft Partitioning with Random Drop - мягкое динамическое деление с произвольным отбрасыванием.

и RND), могут служить для предотвращения перегрузок за счёт отбрасывания или маркировки пакетов до того, как реальное переполнение очередей заставит отправителей TCP снижать скорость передачи.

(3) Решения на стороне поставщиков и потребителей

- Меры контроля насыщения на стороне поставщика включают расширение эффективной полосы пропускания трафика. Это может быть достигнуто за счёт повышения производительности. Другой вариант может обеспечиваться за счёт оптимизации распределения трафика по сети. Например, может быть обеспечено повышение производительности за счёт увеличения числа каналов и повышения их пропускной способности в соответствии с прогнозами роста трафика, а также оптимизация распределения трафика на основе прогнозов (с учётом бюджетных и других ограничений). Однако, если реальное распределение трафика не соответствует топологии, созданной при планировании производительности (например, ошибки в прогнозах или отсутствие соответствующих возможностей), трафик может отображаться на существующую топологию с использованием ориентированных на пути технологий (например, MPLS LSP или оптические каналы) для изменения логической топологии, а также с помощью иных механизмов перераспределения трафика.

- Меры контроля перегрузок на стороне потребителя также могут способствовать преодолению и предотвращению перегрузок. Например, некоторые из описанных выше краткосрочных мер (RED с вариантами, ECN, LQD, RND) а также правила и механизмы формирования трафика позволяют регулировать уровень загрузки. Одним из инструментов воздействия могут служить и тарифы. Однако сегодня тарифная политика на стороне потребителей практически не используется для контроля перегрузок в Internet.

Для решения проблемы перегрузок в сетях IP может использоваться множество механизмов, работающих во временных интервалах самой разной протяжённости.

2.5 Контекст реализации и эксплуатации

Операционный контекст организации трафика Internet характеризуется постоянными изменениями на разных уровнях абстракции. В контексте реализации требуется эффективное планирование, организация и исполнение. Аспекты планирования могут включать определение набора действий для достижения желаемых целей. Организация включает распределение ответственности за различные компоненты систем организации трафика и координацию действий по достижению целей TE. Исполнение включает оценку и состояния и корректировочные действия в целью поддержки желаемого состояния TE.

3.0 Модель процесса организации трафика

В этом разделе описывается общая модель процесса, которая захватывает практические аспекты верхних уровней организации трафика Internet в операционном контексте. Модель описывается как последовательность действий, которые организатор или, в более общем виде, система организации трафика должна предпринять для оптимизации производительности работающей сети (см. также [RFC-2702, AWD2]). Описываемая модель представляет широкую область действий для большинства методологий организации трафика, хотя детали такой организации могут меняться от сети к сети. Эта модель может также явно или неявно исполняться человеком и/или автоматически.

Модель процесса организации трафика является итеративной [AWD2]. 4 описанных ниже фазы модели процесса повторяются непрерывно.

В первой фазе модели процесса TE определяются имеющие отношение к делу правила управления, которые регулируют работу сети. Эти правила могут зависеть от множества факторов, включая бизнес-модель, структуру расходов, операционные ограничения, критерии оптимизации.

Вторая фаза модели процесса включает механизм обратной связи с возможностью получения результатов измерений операционных параметров сети. Если данные сетевых измерений не доступны, вместо них может использоваться теоретическая модель загрузки, отражающая ожидаемое состояние сети. Такая модель может быть построена на основе оценки или экстраполяции результатов проведённых ранее измерений. Для построения теоретической модели может также применяться математическое моделирования характеристик трафика или иные методы.

Третьей фазой модели процесса является анализ состояния сети и определение нагрузочных характеристик. Анализ производительности может быть проактивным и/или реактивным. Превентивный анализ позволяет идентифицировать потенциальные проблемы, которых ещё нет, но они могут возникнуть в будущем. Реактивный анализ определяет существующие проблемы и причины их возникновения с помощью диагностики, а также оценивает при необходимости подходы к решению проблем. В процессе анализа может использоваться множество количественных и качественных методов, включая анализ на базе моделей и имитацию. Фаза анализа может включать исследование концентрации и распределения трафика по сети или её подсетям, определение нагрузочных характеристик, обнаружение имеющихся или потенциальных «пробок», а также идентификацию сетевых «патологий» типа неэффективного расположения каналов, конструктивных ошибок, конфигурационных проблем. В процессе анализа может быть создана матрица трафика. Результаты анализа сети могут быть описательными или директивными.

Четвёртой фазой модели процесса TE является оптимизация производительности сети. Эта фаза включает процесс решения по выбору и реализации действий из числа имеющихся вариантов. Действия по оптимизации могут включать применение подходящих методов для контроля предлагаемого трафика или распределения трафика по сети. К числу таких действие может относиться и добавление каналов или увеличение пропускной способности имеющихся каналов, развёртывание дополнительного оборудования (маршрутизаторы, коммутаторы), систематическая подстройка параметров, связанных с маршрутизацией (таких, как метрики IGP и атрибуты BGP), и параметров управления трафиком. Оптимизация производительности может также включать инициирование процесса планирования сети с целью улучшения архитектуры и организации сети, расширения ёмкости, выбора технологии и конфигурации сетевых элементов с учётом текущих и перспективных потребностей.

3.1 Компоненты модели организации трафика

Основными компонентами модели процесса организации трафика являются подсистемы управления, моделирования и анализа, оптимизации. Ниже эти компоненты рассмотрены более подробно применительно к модели процесса организации трафика.

3.2 Измерения

Измерения является важной частью функций организации трафика. Операционное состояние сети может быть надёжно определено только путём измерений. Измерения важны также для оптимизации, поскольку они обеспечивают данные для подсистемы управления организацией трафика. Эти данные служат для адаптивной оптимизации сетевой производительности в ответ на событие и побудительные мотивы (как внутренние, так и внешние). Измерения нужны также для определения качества сетевых услуг и оценки эффективности правил организации трафика. Опыт показывает, что измерения наиболее эффективны при их систематическом выполнении.

При разработке системы измерения для поддержки организации трафика в сетях IP следует внимательно рассмотреть ряд вопросов. Почему нужно измерять именно в этом контексте? Какие параметры измеряются? Как следует проводить измерения? Где следует выполнять измерения? Когда следует выполнять измерения? Как часто следует измерять переменные при мониторинге? Какой уровень точности и надёжности измерений желателен? Какой уровень точности и надёжности измерений достижим реально? В какой степени измерительные средства могут оказывать влияние на сетевые компоненты и переменные? Какова приемлемая стоимость измерений? Ответы на эти вопросы определяют инструменты и методологию для любого контекста организации трафика.

Следует также подчеркнуть различие между измерением и оценкой. Измерение обеспечивает исходные данные о состоянии сети и параметрах сетевых элементов. Оценка использует исходные данные для того, чтобы сделать выводы о состоянии контролируемой системы.

Измерения в поддержку TE могут выполняться на разных уровнях абстракции. Например, можно выполнить измерения для определения характеристик на уровне пакетов, потоков, пользователей, агрегатов трафика, компонент или сети в целом.

3.3 Моделирование и анализ

Моделирование и анализ являются важными аспектами организации трафика. Моделирование включает разработку абстрактного или физического представления, отражающего реальные характеристики трафика и атрибуты сети.

Модель сети является абстрактным представлением сети, которое включает имеющие отношение к делу характеристики, атрибуты и функции (например, атрибуты и ограничения узлов сети). Модель сети может упростить анализ и имитацию, которая может применяться для предсказания производительности сети в разных условиях, а также для оценки планов расширения сети.

В общем случае модели организации трафика Internet можно разделить на структурные и поведенческие. Структурные модели фокусируются на сети и её компонентах. Поведенческие модели рассматривают динамику сети и её загрузки. Моделирование организации трафика Internet может быть также формальным и неформальным.

Аккуратность моделирования источников трафика очень полезна для анализа. Разработка моделей поведения источников трафика, согласующихся с эмпирическими данными из работающей сети, является основной темой исследований по организации трафика Internet. Эти модели источников должны быть гибкими и пригодными для анализа. Тема моделирования источников для трафика IP является отдельной темой исследований, которая выходит за рамки данного документе. Однако здесь следует подчеркнуть важность этой темы.

Средства моделирования являются очень полезным инструментом для организации трафика. По причине сложности реалистичного количественного анализа поведения сети некоторые аспекты исследования производительности сети можно эффективно провести лишь с помощью моделирования. Хорошая модель сети может служить для безопасной и неразрушающей имитации и визуализации характеристик сети в разных условиях. Например, моделирование позволяет отметить перегруженные ресурсы и «горячие точки», а также обеспечить рекомендации по возможным решениям проблем сетевой производительности. Хорошую модель можно также использовать для оценки эффективности планируемых решений без вмешательства в работу реальной сети или целесообразности дорогостоящей модернизации. Более того, в процессе планирования сети моделирование позволяет выявить возможные аномалии типа критических точек отказа, которые могут потребовать дополнительного резервирования, или потенциальные «пробки» и «горячие точки», которые могут потребовать дополнительных «мощностей».

Моделирование маршрутизации особенно полезно для больших сетей. Модель поможет идентифицировать запланированные каналы, которые реально не будут использоваться для маршрутизации трафика существующими протоколами маршрутизации. Моделирование можно также применять для анализа на основе сценариев и возмущений или для изучения чувствительности. Результаты моделирования могут послужить основой для первоначальных действий в разных ситуациях. Например, важным приложением моделирования сети является исследование и идентификация наиболее эффективных вариантов развития и расширения сети.

3.4 Оптимизация

Оптимизация производительности сети включает решение сетевых проблем путём их преобразования в концепции, которые позволяет найти и реализовать решение. Оптимизация сети может быть корректировочной или совершенствующей. При коррекционной оптимизации цель заключается в устранении имеющихся или намечающихся проблем. При совершенствовании цель заключается в повышении производительности сети даже если реальных проблем не возникает и не предвидится.

Как было отмечено выше, оптимизация производительности сети - это непрерывный процесс. Итерации этого процесса могут включать оптимизационные субпроцессы, выполняемые в реальном масштабе времени, а также субпроцессы планирования действий. Разница между оптимизацией в реальном масштабе времени и планированием заключается, прежде всего, во временном масштабе действий и возможности их дробления на более мелкие части. Одной из целей выполняемой в реальном масштабе времени оптимизации является контроль отображения и распределения трафика через существующую сетевую инфраструктуру для предотвращения и/или снижения уровня перегрузок, обеспечения

удовлетворительного качества услуг и оптимизации использования ресурсов. Необходимость оптимизации в реальном масштабе времени обусловлена возможностью случайных инцидентов типа повреждения кабеля или резкого роста трафика, которые не зависят от качества организации сети. Такие инциденты могут вызывать перегрузки и другие проблемы в работе сети. Оптимизация в реальном масштабе времени должна решать такие проблемы в кратко- и среднесрочной перспективе (от микросекунд до минут и часов). Примерами такой оптимизации являются управление очередями, настройка метрики IGP/BGP, использование технологий типа MPLS с явными LSP для смены путей некоторых транков [XIAO].

Одной из функций сетевого планирования является инициирование действий по систематической оценке архитектуры, технологий, топологии и пропускной способности сети. При наличии в сети проблем для их незамедлительного устранения следует использовать оптимизацию в реальном масштабе времени. В этом случае решение проблем требуется срочно и выбранный вариант может быть совсем не лучшим. Впоследствии могут потребоваться те или иные действия в части сетевого планирования для исправления или улучшения принятого при устранении проблем решения. Сетевое планирование требуется также при расширении сети, связанном с ростом трафика, и для изменения картины распределения трафика во времени. Как было отмечено выше, результатом сетевого планирования может стать изменение топологии и/или ёмкости сети.

Очевидно, что сетевое планирование и оптимизация в реальном масштабе времени взаимно дополняют друг друга. Хорошо спланированная и организованная сеть упрощает и оптимизацию в реальном масштабе времени, а систематическая оптимизация работы сети в реальном масштабе времени позволяет сосредоточить планирование на долгосрочных стратегических задачах, не разменяваясь на решение сиюминутных проблем. Систематическая оптимизация в реальном масштабе времени также обеспечивает информацию, требуемую для сетевого планирования.

Важным фактором оптимизации в реальном масштабе времени является стабильность работы сети и этот аспект рассматривается в данном документе неоднократно.

4.0 Исторический обзор и свежие разработки

В этом разделе кратко рассматриваются различные модели организации трафика, предложенные и реализованные в телекоммуникационных и компьютерных сетях. Обсуждение не претендует на полноту и предназначено, прежде всего, для освещения подходов к организации трафика в Internet и традиционных телекоммуникационных сетях.

4.1 Организация трафика в классических телефонных сетях

В этом параграфе представлен краткий обзор организации трафика в телефонных сетях, которые часто используются в качестве пути для пользовательского трафика между генерирующим и потребляющим данные узлами. Приведённое здесь описание темы достаточно кратко. Более подробные описания разных стратегий маршрутизации в телефонных сетях даны в книге G. Ash [ASH2].

В ранних телефонных сетях использовалась иерархическая статическая маршрутизация, при которой картина маршрутизации сохранялась постоянной независимо от состояния сети или времени суток. Иерархия служила для адаптации к высокому уровню трафика и повышения уровня надёжности сети за счёт использования дополнительных маршрутов, а также для предотвращения петель за счёт использования строгих правил иерархии. Сеть обычно была недогружена, поскольку данный фиксированный маршрут рассчитывался на передачу пользовательского трафика в часы и дни пиковой загрузки. Иерархическая маршрутизация в телефонных сетях оказалась слишком жёсткой для использования с цифровыми коммутаторами и программами управления, способными реализовать более сложные правила организации трафика.

Динамическая маршрутизация была введена для повышения уровня гибкости, обеспечивающего повышение эффективности работы сети. Это обеспечило значительный экономический эффект [HUSS87]. Динамическая маршрутизация обычно снижает общий уровень потерь на 10 - 20 процентов по сравнению с иерархической статической маршрутизацией. Кроме того, динамическая маршрутизация делает сеть более отказоустойчивой за счёт организации маршрутов на уровне отдельных вызовов и периодического обновления маршрутов.

В телефонных сетях используется три основных типа динамической маршрутизации - по времени, по состоянию (SDR[1]) и по событиям (EDR[2]).

В маршрутизации по времени используются регулярные изменения картины трафика с течением времени (по дням недели и времени суток) для планирования таблиц маршрутизации. В маршрутизации по состояниям таблицы маршрутизации обновляются в соответствии с текущим состоянием сети (например, потребности в трафике, уровень загрузки и т. п.). В маршрутизации по событиям изменения маршрутов инициируются теми или иными событиями (например, перегрузка или блокировка линий) с использованием обучающихся моделей. Методы EDR являются адаптивными в реальном масштабе времени и не требуют наличия глобальной информации о состоянии, как в SDR. Примеры схем EDR включают DAR[3] от компании BT, STR[4] от NTT и STT[5] от AT&T.

Динамическая неиерархическая маршрутизация DNHR[6] является примером динамической маршрутизации, реализованной в телефонной сети компании AT&T в 1980-е годы в качестве отклика на регулярные временные вариации загрузки. Зависящую от времени информацию в терминах нагрузки можно разделить по трём интервалам - сутки, неделя и год. Следовательно, для таблиц маршрутизации используется три предопределённых алгоритма. Алгоритм организации сети работает на длинных интервалах (год), тогда как алгоритм обслуживания запросов используется для недельных интервалов с целью подбора полосы каналов и таблиц маршрутизации для исправления ошибок годичного прогнозирования. Для коротких интервалов используется алгоритм маршрутизации, выполняющий в ограниченных масштабах тонкую настройку в соответствии с распределением трафика по времени суток. Годичные и недельные таблицы рассчитываются автономно (offline). Обычно для таких расчётов требуется расширенный поиск возможных маршрутов. С другой стороны, маршрутизация может потребовать расчётов в реальном масштабе времени

[1]State-dependent routing – маршрутизация в зависимости от состояния.
[2]Event dependent routing - маршрутизация, определяемая событиями.
[3]Dynamic alternate routing – динамическое чередование маршрутов.
[4]State-and-time dependent routing - маршрутизация в зависимости от состояния и времени.
[5]Success-to-the-top.
[6]Dynamic non-hierarchical routing.

(online) для обработки нетривиальных ситуаций (crankback). DNHR использует модель «двух каналов» (two-link), в которой путь может включать не более двух каналов. Алгоритм маршрутизации представляет упорядоченный список путей между исходным и целевым коммутатором. При возникновении перегрузки промежуточный коммутатор (на пути между исходным и целевым) будет передавать исходному коммутатору crankback-сигнал. По таком сигналу этот коммутатор будет выбирать следующий маршрут и так далее, пока имеются дополнительные маршруты.

4.2 Развитие организации трафика в сетях пакетной коммутации

В этом параграфе даётся обзор предшествующих работ в части повышения производительности сетей передачи данных. Практическая оптимизация производительности сетей передачи данных началась с первых дней существования сети ARPANET. Вклад в работы по оптимизации производительности и дифференцированию услуг внесли и другие коммерческие сети типа SNA.

В терминах организации трафика сеть Internet до недавнего времени представляла собой среду с доставкой данных по мере возможностей (best effort service). В частности, сети IP обеспечивали весьма ограниченные возможности управления трафиком для обеспечения дифференцированного управления очередями и планирования услуг для пакетов разных классов.

В терминах управления маршрутами сеть Internet использует распределенные протоколы для внутридоменной маршрутизации. Эти протоколы хорошо масштабируются и достаточно отказоустойчивы. Однако они основаны на простых алгоритмах выбора пути с весьма ограниченной функциональностью в части гибкости процесса выбора пути.

В следующих параграфах рассматривается развитие механизмов практической организации трафика в сетях IP и предшественники этих механизмов.

4.2.1 Адаптивная маршрутизация в ARPANET

На ранних этапах существования ARPANET стала ясной важность адаптивной маршрутизации, когда решения о выборе маршрута принимаются на основе текущего состояния сети [MCQ80]. В ранних моделях маршрутизации по минимальной задержке пакеты направлялись адресатам по пути, который обеспечивал минимальное оценочное время доставки. Каждый узел поддерживал таблицу задержек в сети, которые могли возникнуть при передаче пакета адресату по данному пути. Таблица минимальных задержек периодически рассылалась узлами их соседям. Распространялась также информация о кратчайшем пути (в интервалах пересылки - hop count).

Одним из недостатков такой модели было возникновение «притяжения трафика» в эффективные в данный момент каналы, что создавало перемещающиеся из одной точки сети в другую перегрузки, которые приводили к ненужным осцилляциям и нестабильности сетей.

4.2.2 Динамическая маршрутизация в Internet

Сеть Internet, развившаяся из APARNET, использует алгоритмы динамической маршрутизации с распределенным управлением для определения путей передачи пакетов их адресатам. Алгоритмы маршрутизации являются адаптацией алгоритмов поиска кратчайшего пути, где «стоимость» (протяжённость) определяется метрикой каналов. В качестве метрических характеристик каналов могут использоваться статические или динамические количественные параметры. Основанная на статических параметрах метрика каналов может определяться административными мерами в соответствии с локальными критериями. Динамическая метрика может быть функцией параметров загрузки сети (таких, как время задержки и уровень потери пакетов.

Достаточно быстро пришло понимание того, что статическое задание метрики не является адекватным, поскольку в этом случае отдельные каналы окажутся перегруженными, а другие будут простаивать. Одной из основных причин неадекватности статической метрики явилось то, что метрические параметры зачастую присваивались без рассмотрения матриц трафика в сети. Кроме того, протоколы маршрутизации не учитывали атрибуты трафика и ограничения каналов при выборе маршрутов. Это приводило к концентрации трафика в отдельных частях сетевой инфраструктуры и могло вызывать перегрузки. Даже при выделении метрических параметров с учётом матриц трафика между разными каналами мог возникать дисбаланс загрузки по ряду причин, включая:

- отсутствие возможности размещения ресурсов в местах, оптимальных с точки зрения маршрутизации;
- ошибки в оценках объёмов трафика и его распространения;
- динамический характер матрицы трафика, связанный с временными вариациями, изменением политики BGP и т. п.

Неадекватность системы внутренней маршрутизации Internet послужила одной из причин разработки ориентированных на пути технологий с явными маршрутами и поддержкой маршрутизации с учётом ограничения (типа MPLS).

4.2.3 Маршрутизация ToS

Маршрутизация с учётом типа обслуживания (ToS[1]) предполагает поддержку разных маршрутов к одному получателю, выбираемых в зависимости от значения поля ToS в заголовке пакетов IP [RFC-2474]. Классы ToS можно связать с малой задержкой или высокой пропускной способностью. С каждым каналов связывается множество значений «стоимости» и эти значения применяются при расчёте маршрута для конкретного ToS. Для каждого значения ToS рассчитывается отдельное дерево с кратчайшим путём. Классическая маршрутизация на основе ToS сейчас считается устаревшей, поскольку соответствующее поле в заголовке IP было заменено полем Diffserv. Эффективная организация трафика при классической маршрутизации на основе ToS была сложна, поскольку для каждого класса по-прежнему использовался единственный кратчайший путь, что приводило к аномалиям распределения трафика через сети.

[1]Type-of-Service.

4.2.4 Множество равноценных путей

ECMP[1] является другим методом, который пытается преодолеть неэффективность систем внутренней маршрутизации SPF[2] [RFC-2328]. В классическом алгоритме SPF при наличии двух или более равноценных путей к адресату алгоритм выбирает один из них. В ECMP алгоритм был слегка изменён, чтобы при наличии между парой узлов двух или более равноценных кратчайших путей разрешить распределение трафика между равноценными путями. Распределение трафика обычно выполняется одним из двух способов - (1) циклический перебор путей на уровне пакетов или (2) выбор пути на уровне потоков с использованием хеширования адресов IP получателя и отправителя, а также других полей заголовка IP. Первая модель может приводить к нарушению порядка доставки пакетов, а вторая зависит от числа и распределения потоков. Распределение нагрузки на уровне потоков может приводить к непредсказуемым результатам в корпоративных сетях, где число потоков сравнительно мало и они более однородны (например, хеширование может создавать неоднородности), но в общем случае такое распределение более эффективно в магистральных сетях с большим числом и неоднородностью потоков трафика.

В ECMP «стоимость» канала задаётся статически и ограничения пропускной способности не принимаются во внимание, поэтому ECMP пытается распределять трафик по возможности равномерно, независимо от загрузки каждого из путей. В результате при двух равноценных путях возможно возникновение перегрузок на одном из них. Другим недостатком ECMP является невозможность распределения трафика между неравноценными путями.

4.2.5 Маршрутизация Nimrod

Система Nimrod разработана для обеспечения маршрутизации в неоднородной среде Internet, когда приходится принимать во внимание множество ограничений [RFC-1992]. Важно отметить, что Nimrod представляет собой протокол маршрутизации на основе состояния каналов, который поддерживает ориентированную на пути пересылку пакетов. Протокол использует концепцию отображения сетевой связности и служб на разных уровнях абстракции. Обеспечиваются механизмы ограничения области распространения маршрутной информации.

Система Nimrod не получила широкого распространения в сети общего пользования Internet, многие из ключевых концепций архитектуры Nimrod (такие, как выбор пути на узле-инициаторе) находят применение в последующих разработках систем маршрутизации с учётом ограничений.

4.3 Наложенные сети

В модели с наложением (Overlay) используется сеть виртуальных каналов (ATM, Frame Relay, WDM) для организации виртуальных соединений между маршрутизаторами, расположенными на краях облака виртуальных устройств. В этом режиме соединённые между собой виртуальным каналом два маршрутизатора видят себя соединёнными напрямую, независимо от физического маршрута виртуального канала через сеть ATM, Frame Relay или WDM. Таким образом, модель с наложением отвязывает видимую маршрутизаторами логическую топологию от физической топологии сети ATM, Frame Relay или WDM. Наложенная сеть на базе ATM или Frame Relay позволяет администратору сети или программам управления реализовать концепции организации трафика для оптимизации пути за счёт изменения конфигурации или соединений виртуальных устройств, чтобы менять путь передачи данных при возникновении в сети перегрузок на виртуальных каналах или неоптимальной работе физических соединений. В наложенной сети организация трафика используется также для организации связи между параметрами управления трафиком (например, PCR, SCR, MBS для ATM) технологии виртуальных устройств и реальным трафиком, проходящим через каждое устройство. Эти связи могут создаваться на основе известных или проектных профилях трафика, а также некоторых других факторов.

При наложении сети IP на ATM требуется управление двумя разными сетями с различными технологиями (IP и ATM), что усложняет и удорожает работу. В полносвязной модели с наложением каждый маршрутизатор соединён со всеми другими маршрутизаторами сети и число соединений между маршрутизаторами растет пропорционально квадрату числа маршрутизаторов. Некоторые проблемы модели с наложением рассмотрены в работе [AWD2].

4.4 Маршрутизация с учётом ограничений

Маршрутизация с учётом ограничений основана на расчёте маршрутов через сеть в соответствии с некими наборами требований и ограничений. В самом общем случае такую маршрутизацию можно рассматривать, как способ оптимизации работы сети с минимальными расходами.

Ограничения и требования могут вноситься самой сетью или административными правилами. Ограничения могут включать пропускную способность, число интервалов пересылки, задержки и инструменты политики (такие, как атрибуты класса ресурсов). В число ограничений могут включаться специфические атрибуты доменов, использующих некоторые сетевые технологии, и контексты, вносящие ограничения на выбор маршрутов. Ориентированные на пути технологии типа MPLS сделали маршрутизацию с учётом ограничений возможной и привлекательной для публичных сетей IP.

Концепция основанной на ограничениях маршрутизации в контексте требований MPLS по организации трафика в сетях IP была изначально определена в [RFC-2702].

В отличие от маршрутизации на основе QoS (например, [RFC-2386] и [MA]), которая в общем случае решает вопросы маршрутизации отдельных потоков трафика на основе требований к качеству обслуживания (QoS) с учётом доступности сетевых ресурсов, маршрутизация на основе ограничений применима к агрегатам трафика наравне с потоками, а также может использовать различные типы ограничений, включая ограничения, обусловленные политикой.

4.5 Обзор других проектов IETF, связанных с организацией трафика

В этом параграфе рассматриваются многочисленные действия IETF в части организации трафика Internet. Действия эти направлены в первую очередь на развитие архитектуры IP с целью поддержки новых определений служб, обеспечивающих предпочтительную или дифференцированную обработку для некоторых типов трафика.

[1]Equal Cost Multi-Path - множество равноценных путей.
[2]Shortest Path First - сначала кратчайший путь.

4.5.1 Интегрированные услуги

Рабочая группа IETF Integrated Services подготовила модель интегрированных услуг (Intserv). Эта модель требует резервирования ресурсов (таких, как пропускная способность и буферная ёмкость) для потоков трафика данного класса с целью обеспечения запрашиваемого для этого класса качества обслуживания. Модель интегрированных услуг включает дополнительные компоненты, сверх применяемых в обычной модели best-effort - к ним относятся классификаторы и планировщики пакетов, а также средства контроля допуска. Классификаторы пакетов служат для идентификации потоков, которым предоставляется определённый уровень обслуживания. Планировщик обеспечивает планирование обслуживания различных потоков трафика для выполнения требований QoS. Контроль допуска служит для определения маршрутизаторов, имеющих ресурсы, потребные для восприятия нового потока.

В модели Integrated Services были определены два типа сервиса - гарантированное обслуживание [RFC-2212] и контролируемая нагрузка (Controlled-load) [RFC-2211].

Гарантированный сервис может применяться для приложений, которым требуется ограничение времени доставки пакетов. Для этого типа приложений данные, доставленные после заданного интервала времени, обычно считаются бесполезными. Следовательно, гарантированный сервис был предназначен для обеспечения гарантированных границ задержки доставки пакетов для потока. Это достигается путём контроля задержки в очередях на элементах сети вдоль пути потока данных. Однако модель гарантированного обслуживания не обеспечивает ограничения вариаций задержки (изменений интервала между доставкой последовательных пакетов).

Сервис Controlled-load может применяться для адаптивных приложений, которые устойчивы к некоторой задержке, но чувствительны к перегрузкам в сети. Приложения этого типа обычно удовлетворительно работают в сетях со слабой загрузкой, но их работа существенно ухудшается при росте уровня загрузки в сети. Сервис Controlled-load был разработан для обеспечения обслуживания, близкого к обычному (best-effort) в условиях слабой загрузки сети, независимо от реальных условий в сети. Сервис Controlled-load описывается на качественном уровне в том смысле, что количественные параметры задержки или потерь для него не задаются.

Основной проблемой модели Integrated Services является масштабирование [RFC-2998], особенно в больших публичных сетях IP, где могут одновременно передаваться миллионы микро-потоков.

Важной особенностью модели Integrated Services является явная потребность в механизмах сигнализации для передачи запросов QoS от оконечных систем к маршрутизаторам [RFC-2753]. Протокол RSVP[1] выполняет эти сигнальные функции и является важнейшей компонентой модели Integrated Services. Протокол RSVP рассмотрен ниже.

4.5.2 RSVP

RSVP представляет собой протокол сигнализации для состояний [RFC-2205]. Он поддерживает инициированное получателем резервирование ресурсов как для индивидуальных, так и для групповых потоков данных. RSVP разрабатывался в качестве сигнального протокола для модели интегрированного обслуживания с целью передачи запросов QoS от приложений в сеть для резервирования ресурсов в соответствии с требованиями QoS [RFC-2205].

В RSVP узел-источник или отправитель передаёт получателю сообщение PATH с теми же адресами отправителя и получателя, которые будут использоваться в генерируемом источником потоке данных. Сообщение PATH включает (1) заданную отправителем спецификацию Tspec с характеристиками трафика, (2) шаблон отправителя Template, задающий формат трафика, и (3) необязательную спецификацию Adspec, служащую для поддержки концепции «в один приём с анонсированием» (OPWA[2]) [RFC-2205]. Все промежуточные маршрутизаторы на пути пересылают сообщение PATH на следующий интервал, определяемый протоколом маршрутизации. При получении сообщения PATH адресат отвечает на него сообщением RESV, включающим дескриптор потока, для которого запрашивается резервирование ресурсов. Сообщение RESV передаётся к источнику трафика по пути, обратному относительно пути передачи сообщения PATH. Каждый промежуточный маршрутизатор на пути может принять или отвергнуть резервирование ресурсов, запрошенное в сообщении RESV. Если запрос отвергается, соответствующий маршрутизатор передаёт получателю данных (отправителю сообщения) сообщение об ошибке и сигнальный процесс на этом завершается. Если запрос принят маршрутизатором, для потока выделяется запрошенная пропускная способность и буферная ёмкость, а на маршрутизаторе устанавливается соответствующее состояние для данного потока.

Одной из проблем исходной спецификации RSVP была масштабируемость. Это было связано с тем, что резервирование осуществлялось на уровне микропотоков, поэтому число состояний, поддерживаемых на элементах сети росло пропорционально числу микропотоков. Эти вопросы рассмотрены [RFC-2961].

Позднее протокол RSVP был изменён и расширен для снятия остроты проблемы масштабирования. В результате он стал универсальным протоколом сигнализации в Internet. Например, RSVP был расширен для резервирования ресурсов на уровне агрегатов потоков, организации явных путей с коммутацией по меткам MPLS, а также реализации иных сигнальных функций в Internet. Есть также множество предложений по снижению числа обновлений, которые требуется передавать для поддержки организованных сессий RSVP [RFC-2961].

В разработки, связанные с протоколом RSVP, вовлечено множество рабочих групп IETF, включая исходную группу RSVP, рабочие группы MPLS, Resource Allocation Protocol (протокол выделения ресурсов), Policy Framework.

4.5.3 Дифференцированные услуги

Целью работ по дифференцированным услугам (Diffserv[3]) под эгидой IETF была разработка масштабируемых механизмов разбиения трафика на поведенческие агрегаты и последующая дифференцированная трактовка каждого из таких агрегатов в плане обслуживания (в частности, выделения таких ресурсов, как пропускная способность и буферная ёмкость [RFC-2475]). Одним из основных мотивов разработок Diffserv была потребность в разработке дополнительных механизмов дифференциации трафика в Internet, которые позволят преодолеть проблемы масштабирования, присущие модели Intserv.

Рабочая группа IETF Diffserv определила поле Differentiated Services в заголовке IP (поле DS). Это поле включает 6 битов, которые раньше использовались в заголовках IP, как октет TOS (тип обслуживания). Поле DS служит для

[1] Resource Reservation Protocol - протокол резервирования ресурсов.
[2] One pass with advertising.
[3] Differentiated Services - дифференцированные услуги.

индикации условий пересылки, которые желательно обеспечить для пакета на транзитных узлах [RFC-2474]. Рабочая группа Diffserv также стандартизовала множество групп PHB) groups[1]. За счёт использования PHB можно определить несколько классов обслуживания, для которых применяются дифференцированные параметры классификации, правил, формирования трафика и управления буферами.

Для конечных пользователей сетевых услуг получение дифференцированного обслуживания от ISP[2] может потребовать специального соглашения об обслуживании (SLA[3]) с ISP. SLA может явно или неявно задавать условия кондиционирования трафика (TCA[4]), которые определяют правила классификации, маркировки, отбрасывания и формования трафика.

Пакеты классифицируются на входе в сеть Diffserv и к ним могут также применяться некие правила и средства формирования трафика. При прохождении пакетов через границу между доменами Diffserv поле DS может изменяться в соответствии с соглашением между этими доменами.

Модель дифференцированных услуг позволяет поддерживать ограниченное число классов обслуживания, указываемых полем DS. Основным преимуществом модели Diffserv по сравнению с Intserv является масштабируемость. Ресурсы распределяются на уровне классов трафика и число поддерживаемых состояний, пропорциональное количеству классов, значительно меньше числа потоков данных от приложений.

Из предыдущего обсуждения должно быть понятно, что модель Diffserv управляет трафиком на уровне отдельного интервала пересылки. Модель управления Diffserv включает набор механизмов управления micro-TE. Для обеспечения требуемого качества обслуживания в сетях Diffserv требуются и другие функции построения трафика типа управления пропускной способностью сетей (включая управление маршрутизацией). Концепция PDB[5] была разработана для более эффективного понимания реализации дифференцированных услуг на уровне домена в целом [RFC-3086].

4.5.4 MPLS

MPLS представляет собой схему пересылки пакетов, включающую расширения традиционных протоколов уровня управления IP. MPLS расширяет модель маршрутизации Internet и повышает эффективность управления путями и пересылкой пакетов [RFC-3031].

На входе с домен MPLS маршрутизаторы LSR[6] разбивают пакеты IP по классам эквивалентной пересылки (FEC[7]) на основе множества факторов, включая комбинацию данных из заголовка IP и локальную маршрутную информацию, поддерживаемую LSR. В соответствии с выбранным классом пересылки перед пакетом добавляется метка MPLS. В средах, отличных от ATM/FR, метка имеет размер 32 бита и включает 20-битовое поле метки, 3-битовое экспериментальное поле (ранее известное, как поле класса обслуживания или CoS[8]), 1-битовый индикатор стека меток и 8-битовое поле TTL. В средах ATM и FR метка содержит информацию, представленную в поле VCI/VPI или DLCI. Поддерживающий MPLS маршрутизатор (LSR) проверяет поле метки и, возможно, экспериментальное поле для использования информации при выборе пути пересылки пакета.

LSR принимают решение о пересылке, используя метки из пакетов в качестве индекса для локальной таблицы NHLFE[9]. Далее пакет обрабатывается в соответствии с NHLFE. Входящая метка при этом может быть заменена на исходящую, а пакет может быть скоммутирован в следующий LSR. Процесс коммутации по меткам очень похож на переключение по меткам (VCI/VPI) в сетях ATM. До того, как пакет покинет домен MPLS, метка MPLS может быть удалена. LSP[10] представляет собой путь между входным LSR и выходным LSR, через который проходит помеченный пакет. Маршрут для явного LSP определяется входным узлом (инициатором) LSP. Для организации LSP в MPLS могут применяться сигнальные протоколы типа RSVP или LDP.

MPLS является очень мощной технологией для организации трафика Internet, поскольку она поддерживает явные LSP, позволяющие эффективно задавать основанную на ограничениях маршрутизацию в сетях IP [AWD2]. Требования по организации трафика с использованием MPLS описаны в [RFC-2702]. Расширения RSVP для поддержки явных LSP рассмотрены в [RFC-3209]. Расширение LDP (известное, как CR-LDP) для поддержки явных LSP представлено в [JAM].

4.5.5 Метрики производительности IP

Рабочая группа IETF IPPM[11] подготовила набор стандартных метрик, которые могут применяться для мониторинга качества, производительности и надёжности услуг Internet. Эти метрики могут использоваться операторами, конечными пользователями, независимыми группами тестирования и поставщиками услуг для лучшего понимания производительности и надёжности Internet-компонент «облака», которое они используют или предоставляют в пользование [RFC-2330]. Критерии для метрик производительности, разработанных IPPM WG, описаны в [RFC-2330]. Примерами таких метрик являются потери пакетов на пути в одном направлении [RFC-2680], задержка на пути в одном направлении [RFC-2679], параметры связности между парой узлов [RFC-2678]. К другим метрикам относятся измерения второго порядка для задержек и потери пакетов.

Некоторые из метрик производительности, предложенных IPPM WG, будут полезны при подготовке соглашений SLA. Такие соглашения представляют собой набор параметров обслуживания, согласованных между поставщиком и потребителем услуг, где каждый параметр задаётся комбинацией из некоторого числа метрик (возможно с некоторыми ограничениями).

---

[1] Per-Hop Behavior - поведение на этапах пересылки.
[2] Internet Service Provider - поставщик услуг доступа в Internet.
[3] Service Level Agreement - соглашение об уровне обслуживания.
[4] Traffic Conditioning Agreement - соглашение о кондиционировании трафика.
[5] Per Domain Behavior - поведение на уровне домена.
[6] Label switching router - маршрутизатор с коммутацией по меткам.
[7] Forwarding equivalence clas
[8] Class-of-Service.
[9] Next hop label forwarding entry - запись для следующего интервала пересылки по меткам.
[10] Label Switched Path - путь с коммутацией по меткам.
[11] IP Performance Metrics - метрики производительности IP.

4.5.6 Измерение потока

Рабочая группа IETF RTFM[1] подготовила архитектурный документ, определяющий метод задания потоков трафика и множество компонент для измерения параметров потока (измерители, считыватели результатов, управление) [RFC-2722]. Система измерения потоков позволяет измерять и анализировать потоки сетевого трафика с различными целями. Как отмечено в RFC 2722, система измерения потоков может быть очень полезна для (1) изучения поведения существующих сетей, (2) планирования развития и расширения сетей, (3) числового выражения производительности, (4) проверки качества сетевых услуг, (5) описания работы в сети для пользователей.

Система измерения потоков включает измерители, считыватели и средства управления (менеджеры). Измерители наблюдают потоки пакетов через точку измерения, классифицируют их по группам, собирают те или иные данные (например, число пакетов или байтов для каждой группы) и сохраняют результаты в таблице потоков. Группы могут представлять пользовательские приложения, хосты, сети, группы сетей и т. п. Считыватели собирают данные от измерителей и представляют их в доступной для анализа форме. Средства управления отвечают за настройку и контроль за работой измерителей и считывателей. Получаемые измерителем от менеджера инструкции включают спецификации потоков, параметры управления измерителем и методы отбора данных. Считыватели получают инструкции об адресах измерителей для сбора данных, частоте считывания и типах потоков, для которых собирается информация.

4.5.7 Контроль перегрузки на конечных точках

В [RFC-3124] предложен набор механизмов контроля насыщения, которые могут применяться в транспортных протоколах. Целью документа являлась также разработка механизмов для унификации контроля насыщения среди группы конечных точек с активными индивидуальными (unicast) соединениями (congestion group). Менеджер перегрузок непрерывно контролирует состояние пути для каждой из контролируемых им групп. Полученную информацию менеджер использует для передачи планировщикам инструкций по распределению пропускной способности в периоды возникновения перегрузок для данной группы.

4.6 Обзор действий ITU по части организации трафика

В этом параграфе приводится обзор предшествующих работ ITU-T по организации трафика в традиционных телекоммуникационных сетях.

Рекомендации ITU-T E.600 [ITU-E600], E.701 [ITU-E701] и E.801 [ITU-E801] посвящены вопросам организации трафика в традиционных телекоммуникационных сетях. E.600 определяет терминологию для описания концепций организации трафика, E.701 определяет эталонные соединения, уровень сервиса (GOS[2]) и параметры трафика ISDN. В E.701 используется концепция эталонного соединения для идентификации типичных случаев различных типов соединений без описания специфики из реальной физической реализации. Как указано в E.600, «соединение представляет собой связывание ресурсов, обеспечиваемых средства коммуникаций между двумя или более устройствами, включёнными в телекоммуникационную сеть или подключёнными к ней." E.600 также определяет «ресурс, как любое множество физически или концептуально идентифицируемых элементов телекоммуникационной сети, использование которых может быть однозначно определено» [ITU-E600]. Соединения могут быть разнотипными, поскольку число и тип используемых в них ресурсов может меняться.

Обычно путь соединения включает разные сегменты сетей. Например, соединение может быть местным, междугородным и международным. Эталонные соединения служат для прояснения и спецификации параметров производительности на различных интерфейсах между разными сетевыми доменами. Каждый домен может включать множество сетей сервис-провайдеров.

Эталонные соединения обеспечивают основу для определения параметров уровней обслуживания (GoS), относящихся к организации трафика в модели ITU-T. Как определено в E.600, «GoS использует множество переменных организации трафика, которые позволяют обеспечить контроль адекватности группы ресурсов в конкретных условиях». Такими переменными GoS могут быть потери, тональные сигналы, задержки и т. п. Они важны для организации и эксплуатации сети, а также для спецификации производительности компонент.

В модели ITU уровень сервиса (GoS) отличается от качества обслуживания (QoS). QoS представляет производительность с точки зрения пользователя телекоммуникационных услуг и выражает уровень удовлетворённости качеством обслуживания. Параметры QoS связаны с аспектами производительности, наблюдаемыми в точках доступа к сервису и на сетевых интерфейсах, а не внутри сети. С другой стороны, GoS представляет набор связанных с сетью параметров, характеризующих адекватность группы ресурсов в конкретных условиях. Для эффективного обслуживания пользователей сети значения параметров GoS и QoS должны быть согласованы и параметры GoS обычно оказывают основное влияние на параметры QoS.

В соответствии с E.600 набор параметров GoS должен выбираться и определяться на «сквозной» основе (end-to-end basis) для всех основных категорий сервиса, обеспечиваемых сетью, чтобы помочь поставщикам услуг в повышении эффективности сети. На основе выбранного набора эталонных соединений выделяются значения для выбранных параметров GoS в условиях нормальной и высокой загрузки сети. Затем эти «сквозные» значения GoS распределяются по отдельным компонентам ресурсов эталонных соединений.

4.7 Распределение информационного содержимого

В сети Internet преобладают взаимодействия «клиент-сервер», в частности для трафика Web (в будущем возможно доминирование более изощрённых медиа-серверов). Расположение и производительность основных информационных серверов оказывают существенное влияние на картину трафика в Internet, а также на приемлемость качества обслуживания с точки зрения конечных пользователей.

Было разработано множество методов распределения и балансирования нагрузки для повышения производительности за счёт использования «реплик» информационных серверов. Эти методы могут обеспечивать пространственное распределение трафика и улучшать динамические характеристики Internet за счёт динамического распределения «зеркал» с учётом размещения клиентов, исходных серверов, относительной загрузки серверов, относительной

[1] Real Time Flow Measurement - измерение трафика в реальном масштабе времени.
[2] Grade of Service.

загрузки разных сетей и их фрагментов. Такой процесс привязки распределенных серверов к клиентам называют Traffic Directing («режиссура» трафика). Процесс работает на уровне приложений.

Схемам Traffic Directing, распределяющим серверы по разным, географически распределенным местам, могут потребоваться опытные данные о производительности сети для принятия более эффективных решений. В будущем могут потребоваться сетевые измерительные системы для получения таких данных. Набор измеряемых параметров пока не определён.

При возникновении в сети перегрузок системам управления трафиком (Traffic Directing и Traffic Engineering) следует действовать согласованно. Этот вопрос требует дальнейшего изучения.

Вопросы размещения и репликации информационных серверов (в частности, серверов web) имеют важное значение для организации трафика Internet, поскольку эти серверы обеспечивают значительный вклад в трафик Internet.

5.0 Классификация систем организации трафика

В этом разделе кратко рассмотрена систематизация моделей организации трафика. Систематизация может быть организована на основе стилей организации трафика и представлений, как показано в списке:

- в зависимости от времени, состояния или событий;
- автономный и интерактивный расчёт;
- централизованные и распределенные;
- локальная и глобальная информация;
- предписывающие и описывающие;
- открытые и замкнутые;
- тактика и стратегия.

Эти системы классификации более подробно рассматриваются в последующих параграфах.

5.1 Организация трафика в зависимости от времени, состояний и событий

Методологии организации трафика можно классифицировать по управляющим факторам - время, состояния или события. Все рассматриваемые в этом документе схемы TE являются динамическими. Статическая организация трафика не предусматривает использования методологии или алгоритмов построения трафика.

В TE по времени используются сохранённые данные о периодических вариациях трафика (например, по времени суток) для предварительного программирования планов маршрутизации и других управляющих механизмов TE. В дополнение к этому могут учитываться подписки пользователей и проекции трафика. Заранее программируемые планы маршрутизации обычно меняются достаточно редко (например, посуточно). Алгоритмы управления по времени не пытаются адаптироваться к случайным изменениям сетевого трафика или сменам условий в сети. Примером такого алгоритма может служить глобальный централизованный оптимизатор, в котором входной информацией является матрица требований QoS для разных классов трафика, как описано в [MR99].

TE в зависимости от состояний адаптирует планы маршрутизации на основе текущего состояния сети. Текущее состояние обеспечивает дополнительные данные о вариациях реального трафика (например, отклонения от регулярных вариаций) которые невозможно предсказать с использованием усреднённых временных зависимостей. Примером организации трафика по состояниям для сравнительно длинных интервалов времени является маршрутизация на основе ограничений. Примером систем организации трафика по состоянием для более коротких интервалов могут служить алгоритмы балансировки нагрузки, описанные в [MATE].

Состояние сети может включать такие параметры, как уровень загрузки, задержки и частота потери пакетов и т. п. Эти параметры могут быть получены разными путями. Например, каждый маршрутизатор может периодически или по какому-то событию рассылать эти данные другим маршрутизаторам. Другим решением для маршрутизатора с адаптивной TE является отправка зондирующих пакетов для сбора информации о состоянии пути. Ещё одним вариантом является использование системы управления для сбора требуемых данных с элементов сети.

Оперативность и точность сбора и распространения данных о состоянии имеет важное значение для адаптивных систем TE, поскольку состояние сети меняется. Определяемые состоянием алгоритмы могут применяться для повышения уровня эффективности и отказоустойчивости сети. Алгоритмы на организации трафика по времени более подходят для предсказуемых вариаций трафика. С другой стороны, алгоритмы организации трафика на базе состояний более подходят для адаптации к преобладающим состояниям сети.

Методы TE на основе событий могут также применяться для выбора пути в TE. Эти методы отличаются от методов TE по времени и состоянию манерой выбора пути. Алгоритмы организации трафика являются адаптивными и распределенными по своей природе и обычно используют модели с обучением для поиска пути TE через сеть. Методы TE по состоянию обычно используют лавинную рассылку ALB[1] для выбора пути TE, а основанным на событиях методам TE не нужна рассылка ALB. Вместо этого обычно определяется доступная пропускная способность с использованием обучающихся моделей, как в методе STT[2]. Лавинная рассылка ALB может быть ресурсоёмкой, поскольку ей требуется полоса для рассылки LSA и процессорная мощность для обработки LSA, что может ограничивать размер области действия/автономной системы (AS[3]). На основании моделирования можно предположить применение методов TE на основе событий для снижения издержек ALB без потери пропускной способности сети [ASH3].

[1]Available-link-bandwidth - доступная полоса канала.
[2]Success-to-the-top - успех прежде всего.
[3]Autonomous system.

5.2 Автономный и интерактивный расчёт

Для организации трафика требуется расчёт планов маршрутизации. Расчёт может выполняться в автономном (offline) или интерактивном (online) режиме. Автономно вычисления могут выполняться для случаев, когда планы маршрутизации не требуется выполнять в реальном масштабе времени. Например, планы маршрутизации для прогноза, можно рассчитывать автономно. Обычно расчёты этого типа используются также при расширенном поиске в многомерном пространстве решений.

Интерактивные расчёты нужны в тех случаях, когда планы маршрутизации должны адаптироваться к меняющимся условиям в сети с помощью зависящего от состояний алгоритма. В отличие от автономных расчётов (которые могут быть ресурсоёмкими) интерактивные расчёты должны быть достаточно простыми и быстрыми, обеспечивая выбор маршрутов, хорошо настраиваемое распределение ресурсов и распределение нагрузки.

5.3 Централизованное и распределенное управление

При централизованном управлении имеется центральный орган, определяющий планы маршрутизации и, возможно, другие параметры управления TE от имени каждого маршрутизатора. Этот орган периодически собирает от всех маршрутизаторов данные о состоянии сети и возвращает им маршрутную информацию. Цикл обновления маршрутизации является критически важным параметром производительности управляемой сети. Для централизованного управления могут потребоваться значительные вычислительные ресурсы и пропускная способность каналов управления.

При распределенном управлении выбор маршрутов происходит автономно на каждом маршрутизаторе на основе представления этого маршрутизатора о состоянии сети. Данные о состоянии сети могут быть получены с использованием зондирования или приняты от других маршрутизаторов в анонсах состояний каналов. Информация о состоянии сети может распространяться даже в исключительных ситуациях.

5.4 Локальная и глобальная информация

Для алгоритмов организации трафика может потребоваться локальная или глобальная информация о состоянии сети.

К локальной информации относится состояние части домена. Примерами могут служить пропускная способность и уровень потери пакетов на конкретном пути. Локальной информации о состоянии может оказаться достаточно для некоторых экземпляров TE с распределенным управлением.

Глобальная информация говорит о состоянии домена организации трафика в целом. Примерами такой информации служат глобальная матрица трафика и данные о загрузке каждого канала в домене. Глобальная информация обычно нужна централизованным системам управления, но в некоторых случаях её применяют и распределенные системы TE.

5.5 Предписания и описания

Системы TE можно также разделить на предписывающие и описывающие.

Предписывающая организация трафика оценивает возможные варианты и рекомендует направление действий. Предписывающую организацию трафика можно дополнительно разделить на корректировку и совершенствование. В корректирующей TE предписывается направление действий для устранения существующих или предполагаемых аномалий. TE для совершенствования предписывают направления действий для оценки и повышения производительности сети даже при отсутствии очевидных аномалий.

Описательная организация трафика, с другой стороны, характеризует состояние сети и оценивает влияние разных политик без рекомендаций каких-либо направлений действия.

5.6 Открытые и замкнутые циклы

Системами организации трафика с разомкнутым циклом называют те, в которых не используется обратной связи с данными о текущем состоянии сети. Однако управляющие воздействия могут использовать локальную информацию в целях учёта.

Системы с замкнутым циклом используют обратную связь с данными о состоянии сети. Данные обратной связи могут быть результатами прошлых или текущих измерений.

5.7 Стратегия и тактика

Целью тактической организации трафика является решение специфических проблем производительности (типа устранения «горячих точек»), которые возникают в сети на тактическом уровне без учёта стратегических потребностей. Без надлежащего планирования и идеи тактическая организация трафика становится сиюминутной.

Стратегическая организация трафика рассматривает вопросы TE с более организованных и систематизированных позиций, принимая во внимание среднесрочные и долгосрочные последствия тех или иных правил и действий.

6.0 Рекомендации по организации трафика Internet

В этом разделе даны рекомендации верхнего уровня для организации трафика в сети Internet. Эти рекомендации представлены в терминах общего назначения.

Рекомендации описывают возможности, требуемые для решения проблем в организации трафика и достижения целей организации трафика. В широком смысле эти рекомендации можно разделить на функциональные и нефункциональные.

Функциональные рекомендации по организации трафика Internet описывают функции, которые следует выполнять системам организации трафика. Эти функции нужны для достижения целей организации трафика путём решения проблем в его организации.

Нефункциональные рекомендации по организации трафика Internet относятся к атрибутам качества и характеристикам состояний системы организации трафика. Такие рекомендации могут включать противоречивые утверждения, а количественное выражение в некоторых случаях может быть затруднительно.

6.1 Базовые нефункциональные рекомендации

Базовые нефункциональные рекомендации по организации трафика Internet включают удобство, автоматизацию, масштабируемость, стабильность, гибкость, наблюдаемость, простоту, эффективность, надёжность, корректность, ремонтопригодность, расширяемость, совместимость и безопасность. В конкретном контексте часть рекомендаций может быть очень важна в то время, как другие окажутся малозначимыми и необязательными. Следовательно, на этапе разработки системы организации трафика (или её компонент) может потребоваться расстановка приоритетов с учётом реальных условий работы.

Ниже кратко рассматриваются некоторые аспекты нефункциональных рекомендация по организации трафика Internet.

Удобство (Usability) - аспект организации трафика, связанный с человеческим фактором. Удобство характеризует простоту развёртывания и обслуживания системы организации трафика. В общем случае желательно иметь систему TE, которую можно быстро развернуть в существующей сети. Желательно также иметь систему TE, которая проста в эксплуатации и поддержке.

Автоматизация (Automation) - при любой возможности система должна автоматически выполнять функции организации трафика, минимизируя участие человека в управлении и анализе действующих сетей. Автоматизация особо важна для больших масштабируемых сетей общего пользования, поскольку привлечение людей к процессам организации трафика значительно повышает эксплуатационные расходы и риск возникновения проблем, связанных с человеческими ошибками. Автоматизация может потребовать встраивания обратной связи и интеллектуальных функций в некоторые компоненты систем организации трафика.

Масштабируемость (Scalability) - современные сети общего пользования растут очень быстро в размерах и по объёмам трафика. Следовательно, система TE должна быть масштабируемой, чтобы оставаться применимой и для выросшей сети. В частности, системы TE должны сохранять функциональность при увеличении сети в части количества маршрутизаторов и каналов, а также при росте объёмов сетевого трафика. Системы TE должны иметь масштабируемую архитектуру и не оказывать негативного влияния на функции и процессы в элементах сети, а также потреблять слишком много ресурсов при сборе и распространении информации о состояниях или выполнении операций по управлению.

Стабильность (Stability) - очень важная характеристика систем организации трафика, реагирующих на изменения состояний сети. Методологии организации трафика в зависимости от состояния обычно должны находить компромисс между стабильностью и чувствительностью. При наличии конфликтующих требований настоятельно рекомендуется делать выбор в пользу стабильности (особенно в магистральных IP-сетях общего пользования).

Гибкость (Flexibility) - система TE должна быть достаточно гибкой, чтобы политику оптимизации можно было менять. В частности, системам TE следует обеспечивать достаточно число конфигурационных опций, позволяющих администратору настроить TE для работы в конкретной среде. Желательно также иметь интерактивную и автономную подсистемы TE, которые можно включать и отключать независимо. Системы TE используемые в сетях с множеством классов должны также включать опцию для поддержки оценки и оптимизации производительности по классам.

Наблюдаемость (Visibility) - в системе TE должны обеспечиваться механизмы сбора статистики из сети и анализа этой статистики для контроля эффективности работы сети. Статистические данные - матрицы трафика, загрузка каналов, задержки, потеря пакетов и другие параметры производительности, которые могут быть определены по измеренным данным - могут служить индикаторами преобладающих в сети условий. Другим примером данных о состоянии, которые следует контролировать, является информация об имеющихся маршрутах (в контексте MPLS информация о маршрутах LSP) и т. п.

Простота (Simplicity) - в общем случае система TE должна быть максимально простой. Важнейшим свойством систем TE является простота использования (т.е., ясный, удобный и интуитивно понятный пользовательский интерфейс). Простота пользовательского интерфейса совсем не означает использования в системе TE примитивных алгоритмов. При использовании сложных алгоритмов и внутренних структур эти сложности должны быть скрыты от сетевого администратора простым пользовательским интерфейсом.

Взаимодействие (Interoperability) - по возможности системы организации трафика и их компоненты следует разрабатывать с использованием интерфейсов на базе открытых стандартов для обеспечения взаимодействия с другими системами и компонентами.

Безопасность (Security) - критически важный вопрос для систем организации трафика. Такие системы обычно контролируют функциональны аспекты работы сетей для обеспечения производительной работы. Следовательно, должны быть приняты адекватные меры защиты от уязвимостей, которые могут быть связаны с нарушениями безопасности и другими недостатками в системах организации трафика.

В оставшейся части раздела приведены функциональные рекомендации верхнего уровня по организации трафика.

6.2 Рекомендации по маршрутизации

Управление маршрутизацией является важным аспектом организации трафика Internet. Маршрутизация влияет на многие важные параметры производительности сетей, включая пропускную способность, задержку, распределение нагрузки. В общем случае очень трудно обеспечить высокое качество обслуживания в распределенной сети без эффективного контроля маршрутизации. Желательно иметь систему маршрутизации, которая при выборе маршрутов принимает во внимание характеристики трафика и имеющиеся в сети ограничения, обеспечивая при этом достаточную стабильность.

Традиционные протоколы внутренней маршрутизации по кратчайшему пути (SPF[1]) основаны на алгоритмах выбора кратчайшего пути через сеть и им присущи весьма ограниченные возможности управления в части организации трафика [RFC-2702, AWD2]. Ограничения этих протоколов рассмотрены ниже.

1. Общеизвестно, что протоколы SPF при выборе маршрута не принимают во внимание ограничения сетей и характеристики трафика. Например, по причине использования протоколами IGP кратчайших путей (на основе административно заданной метрики) пересылки трафика не обеспечивается возможность распределения нагрузки между неравноценными путями. Использование кратчайшего пути для пересылки трафика позволяет экономить ресурсы, но может вызывать ряд проблем: 1) если трафик от источника к адресату превышает пропускную способность канала на кратчайшем пути, на канале (и, следовательно, кратчайшем пути) возникает перегрузка, а более длинные пути между данной парой узлов при этом могут быть не загруженными; 2) кратчайшие пути для разных отправителей могут совпадать на некоторых каналах (если общий трафик от всех источников превысит возможности такого канала, возникает перегрузка). Проблемы могут также возникать в результате изменения трафика с течением времени, если маршрутная конфигурация не изменится достаточно быстро. Это приводит к тому, что топология сети и маршрутная конфигурация со временем становятся не оптимальными, что может приводить к возникновению постоянной перегрузки.

2. Поддержка множества равноценных путей (ECMP[2]) в SPF IGP позволяет распределять трафик по нескольким равноценным путям между парами узлов. Однако ECMP пытается разделить трафик поровну между такими путями. В общем случае ECMP не поддерживает настраиваемого распределения нагрузки между равноценными путями. В результате загрузка путей может существенно различаться, поскольку пути могут использоваться также для трафика других узлов. В конечном итоге это может приводить к перегрузкам отдельных каналов.

3. Изменение метрики IGP для управления маршрутизацией трафика оказывает влияние на всю сеть. В результате могут возникать непредвиденные или нежелательные изменения картины трафика. Недавние работы, упомянутые в разделе 8.0 могут обеспечить более эффективный контроль [FT00, FT01].

Перечисленные ограничения требуют новых возможностей для повышения эффективности функций маршрутизации в сетях IP. Некоторые из таких возможностей уже были описаны где-либо и кратко упоминаются ниже.

Маршрутизация на базе ограничений желательна для развития архитектуры маршрутизации IP-сетей, особенно в сложных опорных сетях IP с комплексной топологией [RFC-2702]. В маршрутизации на основе ограничений маршруты рассчитываются с учётом выполнения требований, связанных с ограничениями. Ограничения могут применяться для пропускной способности, числа интервалов, задержек или средств административного управления правилами типа классов атрибутов для ресурсов [RFC-2702, RFC-2386]. Это позволяет выбирать маршруты, соответствующие заданному набору требований, на которые могут воздействовать сетевые или административные ограничения. Основанная на ограничениях маршрутизация хорошо работает с основанными на путях технологиями, которые поддерживают явную маршрутизацию (например, MPLS).

Основанная на ограничениях маршрутизация может также применяться в качестве способа перераспределения трафика в инфраструктуру (даже для трафика best effort). Например, при аккуратном определении и настройке требований к пропускной способности при выборе пути и атрибутов резервирования полосы на сетевых каналах можно избежать или существенно снизить перегрузку, связанную с неравномерным распределением трафика. Это позволяет повысить уровни производительности и эффективности сети.

Требуется внести множество изменений в обычные протоколы IGP на основе состояния канала (такие, как OSPF и IS-IS), чтобы позволить им распространение дополнительной информации о состояниях, требуемой для маршрутизации на основе ограничений. Такие расширения для протокола OSPF были описаны в [KATZ], а для IS-IS - в [SMIT]. Важно отметить, что такие усовершенствования требуют распространения дополнительной информации в анонсах состояний каналов. В частности, дополнительно к основным данным о состоянии канала от усовершенствованного IGP требуется распространять топологическую информацию, требуемую для маршрутизации на основе ограничений. Такая дополнительная топологическая информация может включать атрибуты канала типа доступной для резервирования пропускной способности и атрибутов класса ресурсов (административно управляемое свойство канала). Концепция атрибутов класса ресурсов определена в [RFC-2702]. Дополнительная топологическая информация передаётся в новых TLV и суб-TLV протокола IS-IS или в Opaque LSA протокола OSPF [SMIT, KATZ].

Усовершенствованный протокол IGP может применять лавинную рассылку информации чаще обычного IGP. Это связано с тем, что даже при отсутствии изменений в топологии смена резервирования полосы или сходства каналов могут инициировать лавинную рассылку IGP. Обычно требуется поиск компромисса между своевременностью рассылки информации и частотой лавинной рассылки, чтобы предотвратить чрезмерный расход пропускной способности и вычислительных ресурсов, а также избежать нестабильности.

В системах TE желательно также обеспечивать для подсистемы маршрутизации возможность настраиваемого распределения нагрузки между несколькими (равноценными или неравноценными) путями. Такая возможность позволяет администратору более гибко управлять распределением трафика через сеть. Это может оказаться весьма полезным для предотвращения или ослабления перегрузок в некоторых ситуациях. Примеры приведены в [XIAO].

Системам маршрутизации также следует поддерживать возможность контроля маршрутов для некого подмножества трафика без воздействия на маршрут остального трафика, если ресурсы это позволяют. Такая возможность обеспечивает более тонкий контроль за распределением трафика через сеть. Например, возможность переноса трафика некой пары отправитель-получатель с одного пути на другой без воздействия на остальной трафик позволяет перенести поток данных на маршрут с достаточным количеством ресурсов. Ориентированные на пути технологии (типа MPLS) изначально поддерживают такую возможность, как показано в [AWD2].

Кроме того, подсистемам маршрутизации следует поддерживать возможность выбора разных путей для различных классов трафика (или различных агрегатов поведения), если сеть поддерживает множество классов обслуживания (агрегатов поведения).

[1]Shortest path first - сначала кратчайший путь.
[2]Equal-Cost Multi-Path.

6.3 Рекомендации по отображению трафика

Отображением трафика является перенос определённой части потока данных на заранее подготовленный путь с целью выполнения неких требований. Таким образом, когда основанная на ограничениях маршрутизация имеет дело с выбором пути, отображение трафика имеет дело со связыванием трафика с определёнными путями, которые могут выбираться с использованием маршрутизации на основе ограничений или иными способами. Отображение трафика может использовать механизмы, зависящие от времени или состояния, как описано в параграфе 5.1.

Важной особенностью функции отображения трафика является способность организовать множество путей между данными отправителем и получателем, а также возможность распределять трафик между парой узлом по некому множеству путей в соответствии с теми или иными правилами. Условием использования такой схемы является наличие гибких механизмов деления трафика и его распределения по разным «параллельным» путям. Эти требования были отмечены в [RFC-2702]. При распределении трафика по множеству параллельных путей рекомендуется принимать специальные меры по сохранению порядка доставки пакетов, относящихся к одному приложению (или микропотоку).

Как правило, механизмам отображения трафика следует направлять трафик в сетевую инфраструктуру так, чтобы снижался уровень перегрузок. Если суммарный трафик не может быть воспринят полностью или функции маршрутизации и отображения не могут достаточно быстро реагировать на изменение условий в сети, система отображения трафика может работать на основе механизмов краткосрочного контроля перегрузок (например, управление очередями, планировщики пакетов и т. п.) для снижения уровня перегрузки. Таким образом, механизмы отображения трафика следует использовать в дополнение к существующим механизмам контроля насыщения. В работающих сетях обычно желательно отображать трафик в инфраструктуру так, чтобы минимизировалась конкуренция за ресурсы внутри классов и между классами.

При использовании механизмов отображения трафика с динамической обратной связью (например, MATE) следует принимать меры по обеспечению стабильности сети.

6.4 Рекомендации по измерению

Измерения играют важную роль для организации трафика, обсуждаемой в этом документе. Следует обеспечивать механизмы измерения и сбора статистики из сети для поддержки функций организации трафика. Для анализа собранной статистики могут потребоваться дополнительные средства. Работа механизмов анализа не должна негативно влиять на точность и целостность собранных данных. Механизмы сбора статистических данных должны быть масштабируемыми с учётом возможного расширения сети.

Статистику трафика можно классифицировать по временному масштабу на краткосрочную и долговременную. Долговременная статистика очень полезна для организации трафика. Такая статистика может отражать периодичность сетевой нагрузки (по часам, дням недели, неделям и т. п.), а также выявлять тенденции изменения трафика. Аспекты собранной статистики могут также отражать характеристики разных классов обслуживания в сетях, поддерживающих множество таких классов. Анализ долговременной статистики **может** давать вторичные статистические данные типа характеристик в часы пиковой загрузки, картин роста трафика, наличие постоянных перегрузок, «горячие точки» и несбалансированность каналов, вызванные маршрутными аномалиями, и т. п..

Следует поддерживать механизмы создания матриц трафика для краткосрочной и долговременной статистики. В мультисервисных сетях IP такие матрицы могут создаваться для разных классов трафика. Каждый элемент матрицы трафика представляет статистику трафика между парой абстрактных узлов, которые, в свою очередь, могут представлять отдельные маршрутизаторы, их множества или сайты в VPN.

Собранная статистика трафика должна обеспечивать разумные и надёжные индикаторы текущего состояния сети в краткосрочной ретроспективе. Некоторые элементы краткосрочной статистики могут отражать уровень загрузки или насыщения каналов. Примерами индикаторов перегрузки могут служить значительные задержки пакетов, высокий уровень потерь или значительная загрузка ресурсов. Примерами механизмов распространения статистических данных могут служить SNMP, зонды, FTP, анонсы состояния каналов IGP и т. п.

6.5 Жизнестойкость сети

Жизнестойкостью сети называют способность сохранять работоспособность при возникновении отказов. Это может быть достигнуто за счёт быстрого устранения неисправностей и обеспечения требуемого QoS для существующих служб после восстановления. Жизнестойкость сетей становится все более важной для сообщества Internet по причине роста уровня критически важного и передаваемого в реальном масштабе времени трафика, а также других данных с высоким приоритетом, которые передаются через Internet. Вопросы жизнестойкости можно решать на уровне оборудования, повышая уровень надёжности сетевых компонент, а также на сетевом уровне за счёт реализации избыточности в архитектуре, структуре и работе сети. Рекомендуется реализовать концепции отказоустойчивости и жизнестойкости в архитектуре, устройстве и работе систем организации трафика, служащих для управления сетями IP (особенно публичными). Поскольку в зависимости от контекста может требоваться разный уровень живучести, механизмы обеспечения жизнестойкости сетей должны быть гибкими для адаптации к разным условиям.

Защита от отказов и средства восстановления стали доступными на многих уровнях в качестве сетевых технологий и продолжают совершенствоваться. В основании многоуровневого стека современные оптические сети поддерживают кольца и многосвязные топологии для быстрого восстановления на уровне длин волн, а также традиционные механизмы защиты. На уровне SONET/SDH средства обеспечения жизнестойкости включают автоматическое переключение (APS[1]), а также кольцевые и многосвязные топологии с функциями «самолечения». Подобная функциональность обеспечивается и технологиями канального уровня типа ATM (в общем случае в более медленным восстановлением). На уровне IP традиционно применяется изменение маршрутов для восстановления при отказах каналов или узлов. Перемаршрутизация на уровне IP выполняется по истечении времени схождения маршрутов, которое может составлять секунды и даже минуты. Некоторые разработки в контексте MPLS обеспечивают восстановление на уровне IP до завершения схождения маршрутов [SHAR].

[1]Automatic Protection Switching - автоматическое защитное переключение.

Для поддержки расширенных требований по живучести могут применяться ориентированные на пути технологии типа MPLS, которые повышают жизнестойкость сетей IP потенциально экономически эффективными способами. Преимущества ориентированных на пути технологий типа MPLS для восстановления IP становятся более очевидными в тех случаях, когда требуется защита и восстановление на основе классов.

Недавно был предложен общий стек протоколов управляющего плана для MPLS и оптических транспортных сетей, названный Multi-protocol Lambda Switching [AWD1]. Новая парадигма мультипротокольной коммутации по длинам вол н будет поддерживать более изощрённые возможности восстановления на оптическом уровне за счёт многосвязности для архитектуры с передачей трафика IP по сетям с мультиплексированием по длине волны (WDM).

Другим важным аспектом многоуровневого обеспечения жизнестойкости является то, что на разных уровнях технологии защиты и восстановления обеспечивают различные временные гарантии и разную гранулярность пропускной способности (от уровня пакетов до уровня длин волн). Возможности защиты и восстановления зависят также от классов обслуживания в разных моделях работы сетей.

Влияние перебоев в обслуживании заметно различается для различных классов сервиса в зависимости от продолжительности перебоев. Продолжительность отказов может составлять от миллисекунд (слабо влияет на обслуживание) до секунд (возможен сброс соединений IP-телефонии и тайм-ауты в сессиях ориентированных на соединения транзакций) и даже минут и часов (с потенциально большим влиянием на общество и бизнес).

Координация различных средств защиты и восстановления на множестве уровней для согласованного обеспечения жизнестойкости сети за разумную цену является достаточно сложной задачей. Защита и восстановление на разных уровнях могут быть желаемыми не во всех случаях, поскольку сети на разных уровнях могут относиться к разным административным доменам.

Ниже представлены некоторые общие рекомендации в части координации защиты и восстановления.

- Средства защиты и восстановления из разных уровней следует координировать всякий раз, когда это возможно и желательно для обеспечения живучести сети гибким и экономически эффективным путём. Минимизация дублирования функций на разных уровнях является одним из вариантов координации. Распространение сигналов тревоги и других индикаторов на вышележащие уровни также можно организовать координированно. Установление временного порядка таймеров в триггерах восстановления на разных уровнях является ещё одним из вариантов координации многоуровневой защиты и восстановления.
- Резервные возможности верхних уровней зачастую на нижних уровнях представляются рабочим трафиком. Размещение функций защиты и восстановления на множестве уровней может повысить уровень отказоустойчивости и резервирования, но оно не должно приводить к существенному снижению эффективности использования сетевых ресурсов.
- В общем случае желательно реализовать схемы защиты и восстановления, обеспечивающие эффективное использование пропускной способности.
- Уведомления об отказах должны передаваться через сеть своевременно и надёжно.
- Сигналы тревоги и иные индикаторы и средства мониторинга следует размещать на подходящих уровнях.

6.5.1 Живучесть сетей на базе MPLS

MPLS представляет собой развивающуюся технологию, которая расширяет возможности сетей IP в плане функциональности, свойств и услуг. Поскольку технология MPLS ориентирована на пути, она потенциально может обеспечить более быструю и предсказуемую защиту и восстановление по сравнению с традиционной поэтапной маршрутизацией IP. В этом параграфе рассмотрены некоторые базовые аспекты и рекомендации для сетей MPLS в части защиты и восстановления. Более полное рассмотрение вопросов восстановления на базе MPLS проведено в [SHAR].

Варианты защиты для сетей MPLS можно разделить на защиту каналов, узлов, путей и сегментов.

- Защита канала. Целью защиты канала является защита LSP от отказов на данном канале. При использовании защиты канала защитный путь или резервный (вторичный) LSP отсоединяется от пути рабочего (основного) LSP на конкретном канале, который требуется защитить. При отказе на защищённом канале рабочий LSP переключается на защитный в «голове» (head-end) отказавшего канала. Такой способ «локального ремонта» обеспечивает очень быстрое восстановление. Этот метод может хорошо подходить для ситуаций, когда некоторые элементы данного пути менее надёжны по сравнению с другими.
- Защита узла. Целью защиты узла является защита LSP в случае отказа на данном узле. При использовании защиты узла защитный LSP отсоединяется от рабочего LSP на данном защищаемом узле. Вторичный путь отсоединяется также от основного пути на всех каналах, связанных с защищаемым узлом. При отказе узла трафик переключается с рабочего LSP на резервный через систему защиты LSP на восходящем маршрутизаторе LSR, непосредственно подключённом к отказавшему узлу.
- Защита пути. Целью защиты пути LSP является защита LSP от отказов в любой из точек этого пути. При использовании защиты пути защитный LSP полностью отсоединяется от рабочего LSP. Преимущество защиты пути заключается в том, что резервный путь защищает рабочий LSP от отказов на любом из узлов и каналов на пути, за исключением отказов, которые могут возникать на входном или выходном LSR, а также для связанных отказов, которые могут воздействовать одновременно на рабочий и резервный путь. Кроме того, благодаря сквозному характеру защиты пути, она может быть более эффективной в плане использования ресурсов по сравнению с защитой узлов или каналов. Однако защита пути в общем случае срабатывает медленней защиты узла или канала.
- Защита сегмента. Домен MPLS может быть разделен на множество областей защиты, отказ в каждой из которых не распространяется за её пределы. В тех случаях, когда LSP проходит множество областей защиты, используемого в каждой области механизма достаточно для защиты лежащего в этой области сегмента LSP.

Защита сегментов в общем случае работает быстрее защиты пути, поскольку восстановление обычно выполняется ближе к точке отказа.

6.5.2 Варианты защиты

Другой вопрос связан с вариантами защиты. При описании этих вариантов используются обозначения m:n, где m указывает число защитных LSP, применяемых для защиты n работающих LSP. Возможные варианты приведены ниже.

- **1:1** - один работающий путь LSP защищается одним защитным/восстановительным LSP;
- **1:n** - один защитный LSP используется для защиты/восстановления n рабочих LSP;
- **n:1** - для одного рабочего LSP защита/восстановление обеспечиваются n защитных LSP, возможно с настраиваемым распределением нагрузки между ними. При использовании более одного защитного LSP может оказаться желательным распределение трафика между такими LSP в случаях отказа на рабочем LSP для выполнения требований к транку трафика, связанному с рабочим LSP. Особенно полезно это может оказаться в тех случаях, когда нет возможности выбрать один защитный путь, который будет соответствовать требованиям к пропускной способности для основного LSP.
- **1+1** - трафик передаётся одновременно по рабочему и защитному LSP. В этом случае входной маршрутизатор LSR выбирает один из двух LSP на основе локального процесса обеспечения целостности трафика, который сравнивает трафик обоих путей LSP и обнаруживает несоответствия. Широкое распространение этого варианта в сетях IP маловероятно по причине неэффективности использования ресурсов. Однако доступность и дешевизна пропускной способности могут сделать этот вариант подходящим и эффективным решением для сетей IP.

6.6 Организация трафика в средах Diffserv

В этом параграфе приведён обзор функций организации трафика и рекомендации, применимые для сетей IP с поддержкой дифференцированных услуг (Diffserv[1]) [RFC-2475].

Растущие требования по поддержке множества классов трафика (таких, как best effort и mission critical data) в сети Internet заставляет сети IP дифференцировать трафик в соответствии с некоторыми критериями и предоставлять преимущества некоторым типам трафика. Множество потоков можно объединить в небольшое число поведенческих агрегатов на базе тех или иных критериев в терминах общих параметров производительности, вероятности потери пакетов, задержек и их вариаций или полей заголовков в пакетах IP.

По мере развития Diffserv и развёртывания этой технологии в работающих сетях роль организации трафика становится критически важной для обеспечения контрактов SLA с данным классом обслуживания модели Diffserv. Классы обслуживания (CoS[2]) могут поддерживаться в среде Diffserv путём конкатенации PHB[3] на пути маршрутизации использования механизмов обеспечения услуг и подобающей настройки краевой функциональности (классификация трафика, его маркировка, формование и применение правил). PHB представляет собой модель поведения при пересылке, которая применяется к пакетам на узлах DS (узлы, поддерживающие Diffserv). Нужное поведение обеспечивается за счёт управления буферами и механизмами планирования обработки пакетов. В этом контексте относящимися к некому классу пакетами являются те пакеты, которые принадлежат к соответствующему агрегату упорядочения.

Организация трафика может служить дополнением к механизмам Diffserv для повышения эффективности использования сетевых ресурсов, но в общем случае не является обязательным элементом. При использовании организации трафика она может применяться интегрировано для всех классов обслуживания [RFC-3270] или независимо для каждого класса. Первый вариант служит для повышения эффективности распределения доступных сетевых ресурсов для агрегата трафика (см. [RFC-3270], где подробно описана организация трафика на уровне агрегатов). Второй вариант рассматривается ниже, поскольку он специфичен для сред Diffserv с так называемым построением трафика, понимающим Diffserv [DIFF_TE].

Для некоторых сетей Diffserv может оказаться желательным контроль производительности для некоторых классов трафика путём установки неких соотношений между вносимой каждым из классов долей трафика и размером сетевых ресурсов, выделяемых или обеспечиваемых для этого класса. Такие связи между запросами ресурсов и их выделением можно установить с использованием комбинаций, включающих, например, (1) механизмы организации трафика по классам, которые задают желаемую связь объёма трафика данного класса с выделяемыми этому классу ресурсами, и (2) механизмы, которые динамически регулируют выделяемые для данного класса ресурсы в соответствии с объёмом трафика данного класса.

Кроме того, может оказаться желательным ограничить влияние высокоприоритетного трафика на трафик со сравнительно низким приоритетом. Это может быть достигнуто, например, за счёт контроля доли высокоприоритетного трафика, который маршрутизируется через данный канал. Другим способом является увеличение пропускной способности каналов до такой степени, которая позволит даже для трафика с низким приоритетом обеспечить нужное качество обслуживания. Когда соотношения трафика разных классов сильно меняются от маршрутизатора к маршрутизатору, для управления трафиком могут потребоваться дополнительные механизмы сверх традиционных протоколов маршрутизации IGP и построения трафика для разных классов. Вместо этого может потребоваться построение трафика и контроль маршрутизации непосредственно по классам обслуживания. Одним из путей решения этой задачи в доменах, поддерживающих одновременно MPLS и Diffserv, будет определение специфичных для класса LSP и отображение трафика каждого класса в один или множество соответствующих классу обслуживания LSP. После этого LSP, соответствующий данному классу обслуживания, можно маршрутизировать и защищать/восстанавливать по правилам, определяемым классом обслуживания.

Организация трафика по классам может потребовать распространения неких параметров классов. Отметим, что общепринято использование для некоторых классов общих агрегатных ограничений (например, требование к

[1]Differentiated Services.
[2]Class of service.
[3]Per-hop behavior - модель поведения на этапе (пересылки).

максимальной пропускной способности) без задания ограничений для каждого отдельного класса. Такие классы в этом случае можно группировать в class-type и распространять параметры per-class-type, что позволит повысить уровень масштабируемости. В рамках одного class-type можно организовать более эффективное использование пропускной способности. class-type представляет собой набор классов, удовлетворяющих двум условиям:

1) Классы в одном class-type имеют общие агрегатные требования для обеспечения нужного уровня производительности.

2) В рамках class-type не задаётся требований, которые должны исполняться для отдельных классов. Следует обратить внимание, что в рамках одного class-type возможно, тем не менее, реализовать некие правила, обеспечивающие отдельным классам преимущественный доступ к пропускной способности за счёт использования приоритета преемственности.

Примером class-type может служить класс с малыми потерями (low-loss class-type) включающий агрегаты упорядочивания (Ordering Aggregate) на базе AF1 и AF2 одновременно. Для такого class-type можно реализовать некие правила, в соответствии с которыми можно предоставить более высокий приоритет преемственности для трафика AF1 по сравнению с AF2 или наоборот.

Подробное описание требований по организации трафика с учётом Diffserv приведено в [DIFF-TE].

6.7 Управляемость сетей

Польза от автономного (и интерактивного) построения трафика может оказаться ограниченной, если не обеспечивается эффективный контроль сети для реализации решений в части TE и достижения желаемых целей в плане производительности. Прирост «мощности» является достаточно грубым способом оценки решений по организации трафика. Тем не менее этот простой подход может обеспечивать преимущества, если пропускной способности достаточно много и стоит она недорого или пропускная способность соответствует текущей и ожидаемой сетевой нагрузке. Однако пропускная способность может оказаться дефицитной и дорогой, а рост нагрузки может не соответствовать увеличению «мощности» сети. Корректировка административного «веса» и других параметров, связанных с протоколами маршрутизации, обеспечивает более тонкий контроль для сети, но сложнее в использовании и не обеспечивает достаточной точности по причине маршрутных взаимодействий в сети. В некоторых условиях более тонкое и гибкое решение может быть достигнуто путём отображения трафика на маршруты или за счёт выбора и организации маршрутов, которые могут быть целесообразными и полезными.

Механизмы управления могут быть ручными (например, настройка конфигурации), автоматизированными (например, использование сценариев) или автоматическими (например, системы управления на основе правил). Для сетей большого масштаба автоматизация механизмов зачастую просто необходима. Взаимодействие оборудования разных производителей можно обеспечить за счёт разработки и развёртывания стандартизованных систем управления (например, стандартных MIB) и правил (PIB) для поддержки функций управления, требуемых в целях организации трафика (таких, как распределение нагрузки и защита/восстановление).

Функции сетевого управления должны быть безопасными, надёжными и стабильными, поскольку зачастую им приходиться работать в условиях возникновения неполадок в сети (например, при возникновении перегрузок или атак).

7.0 Междоменное взаимодействие

Организация трафика между доменами концентрируется на оптимизации производительности для трафика, который начинается в одном административном домене и завершается в другом.

Обмен трафиком между автономными системами в Internet происходит на основе протоколов внешней маршрутизации. В настоящее время стандартным протоколом этого типа для сети Internet является BGP [BGP4]. Этот протокол поддерживает множество атрибутов и возможностей (например, фильтрацию маршрутов), которые могут применяться для организации трафика между доменами. Более конкретно, BGP позволяет контролировать обмен трафиком и маршрутной информацией между автономными системами (AS[1]) в Internet. BGP включает процесс последовательного принятия решений, в котором рассчитывается уровень предпочтения для разных маршрутов в заданную сеть. Ниже указаны два фундаментальных аспекта организации междоменного трафика с использованием BGP.

- Распространение маршрутов (Route Redistribution) - контроль импорта и экспорта маршрутов между AS, а также контроль обмена маршрутами между BGP и другими протоколами внутри AS.

- Выбор лучшего пути - определение лучшего из множества путей в данную сеть. Выбор лучшего пути выполняется процессом принятия решений BGP на основе упорядоченной процедуры, принимающей во внимание множество разных аспектов. В конечном счёте выбор лучшего пути в BGP сводится к выбору предпочтительной точки выхода из AS в направлении сети адресата. На процесс выбора пути BGP может оказывать влияние воздействие на атрибуты, связанные с процессом принятия решений BGP. К таким атрибутам относятся NEXT-HOP, WEIGHT (фирменный атрибут Cisco, реализованный ещё рядом производителей), LOCAL-PREFERENCE, AS-PATH, ROUTE-ORIGIN, MULTI-EXIT-DESCRIMINATOR (MED), IGP METRIC и др.

С помощью route-map обеспечивается гибкий в реализации комплекс правил BGP, основанных на заданных в конфигурации логических условиях. В частности route-map можно применять для контроля правил импорта и экспорта на входящих и исходящих маршрутах, правил обмена маршрутами между BGP и другими протоколами, а также для воздействия на процесс выбора лучшего пути за счёт воздействия на связанные с процессом принятия решений BGP атрибутами. С помощью route-map, атрибутов BGP, списков доступа (access-list) и атрибутов Community можно создавать сложные логические выражения для различных типов правил.

При рассмотрении возможных стратегий междоменного построения трафика с BGP следует принимать во внимание, что точка выхода исходящего трафика может быть выбрана, тогда как точка входа внешнего трафика, получаемая от партнёра по EBGP, обычно не может быть изменена без специальных действий, выполняемых вместе с этим партнёром. Следовательно, каждой сети для реализации стратегии TE нужно обеспечить эффективную доставку трафика, исходящего от заказчика, в одну из точек соединения с партнёром. Политики TE чаще всего основаны на

[1]Autonomous System.

стратегии поиска «ближайшего выхода», когда междоменный трафик направляется ближайшему внешнему партнёру в направлении целевой автономной системы. Большинство методов изменения точки входа трафика из сети партнёра EBGP (несогласованные с партнёром. анонсы, добавление AS, передача MED) либо не эффективны, либо не приемлемы для партнерского сообщества.

Построение междоменного трафика с BGP, как правило, достаточно эффективно, однако реализуется обычно методом проб и ошибок. Какой-то систематической модели организации междоменного трафика пока не разработано.

Организация междоменного трафика в современной архитектуре Internet более сложна по сравнению с TE внутри домена. Причины этого являются и техническими и административными. Технический аспект связан с тем, что информация о топологии и состоянии каналов, полезная для более эффективного отображения трафика, не передаётся по протоколу BGP из одного домена в другой по причинам, связанным со стабильностью и масштабируемостью. В общем случае решение, которое хорошо работает в одном домене, может оказаться не подходящим для другого домена. Кроме того, обычно для любого домена не желательно влияние других доменов на организацию и маршрутизацию его внутреннего трафика.

Туннели MPLS TE (явные LSP) могут повышать уровень гибкости при выборе точек выхода для междоменной маршрутизации. Для этих целей могут применяться концепции абсолютной и относительной метрики. Идея заключается в том, что если определить атрибуты BGP таким образом, чтобы процесс принятия решения для выбора точки выхода междоменного трафика зависел от метрики IGP, то для некого междоменного трафика, направленного в данную партнерскую сеть, может быть сделана предпочтительной конкретная точка выхода путём организации туннеля TE между делающим выбор маршрутизатором и партнерской точкой с присвоением туннелю TE метрики, которая будет меньше «стоимости» IGP для других точек соединения с партнёрами. Если партнёр принимает и обрабатывает атрибуты MED, можно использовать подобную схему на основе туннели MPLS TE, когда та или иная точка выхода делается более предпочтительной путём установки для MED «стоимости» IGP, которая изменяется туннельной метрикой.

Подобно организации внутридоменного трафика для междоменного TE наилучшие результаты могут быть достигнуты при построении матрицы, отражающей объем трафика из одной автономной системы в другую.

В общем случае перераспределение междоменного трафика требует координации между взаимодействующими партнёрами. Политика экспорта в одном домене, приводящая к перераспределению нагрузки на партнерские точки в другом домене, будет оказывать существенное влияние на локальную матрицу трафика в этом домене. Это, в свою очередь, повлияет на TE внутри домена за счёт изменения пространственного распределения трафика. Следовательно, координация изменений в политике между партнёрами. очень важна, поскольку такие изменения могут существенно менять междоменный трафик. В некоторых случаях такая координация может оказаться достаточно сложной задачей в силу технических и иных причин.

Важно разобраться, как MPLS или похожие технологии можно расширить для обеспечения возможности выбора путей через границы доменов с учётом ограничений.

8.0 Обзор применения TE в сетях IP

В этом разделе рассмотрены некоторые современные методы организации трафика из практики IP-сетей. Рассматриваются, прежде всего, аспекты, связанные с управлением маршрутизацией в работающих сетях. Задача заключается в рассмотрении широко используемых методов. Обсуждаются далеко не все вопросы.

В настоящее время сервис-провайдеры используют множество механизмов организации трафика, описанных в этом документе, для оптимизации производительности своих IP-сетей. Эти методы включают планирование ёмкости на перспективу, управление маршрутизацией с использованием метрики IGP и MPLS для решения среднесрочных задач, а также механизмы управления трафиком для решения краткосрочных задач.

Когда сервис-провайдер хочет построить сеть IP или расширить ёмкость имеющейся сети, планирование эффективной ёмкости является важной компонентой процесса. Такой план может принимать во внимание множество аспектов - расположение новых узлов (если они есть), имеющаяся и предполагаемая картина трафика, стоимость, ёмкость каналов, топология, маршрутизация, живучесть.

Оптимизация производительности рабочей сети обычно представляет собой непрерывный процесс, в котором статистика трафика, параметры производительности, индикация неполадок постоянно собираются в сети. Собранные эмпирические данные анализируются и используются различными механизмами организации трафика. Инструменты анализа могут использоваться и в процессах TE, позволяя рассматривать различные варианты до изменения конфигурации работающей сети.

Традиционно внутридоменная организация трафика с IGP выполняется путём увеличения значений метрики OSPF или IS-IS для перегруженных каналов, пока трафик не будет в достаточно степени переведён на другие каналы. Этой модели присущи некоторые ограничения, описанные в параграфе 6.2. Недавно были предложены некоторые другие модели и инструменты организации внутридоменного трафика [RR94][FT00][FT01][WANG]. Эти модели и средства принимают на входе матрицу трафика, топологию и задачи увеличения производительности сети, а на выходе дают значения метрики для каналов, которые могут сопровождаться некими неравноценными отношениями для маршрутизаторов в отдельных ECMP. Эти новые модели открывают новые возможности для более систематизированной организации трафика внутри доменов с IGP.

Модель с перекрытием (IP over ATM или IP over Frame relay) является другим распространенным практическим вариантом [AWD2]. Методы IP over ATM уже не представляются предпочтительными по причине недавних успехов MPLS и расширения возможностей маршрутизирующего оборудования.

Развёртывание MPLS для приложений организации трафика началось в сетях некоторых сервис-провайдеров. Одним из работающих вариантов является использование MPLS вместе с IGP (IS-IS-TE или OSPF-TE) с поддержкой расширений для организации трафика совместно с маршрутизацией на базе ограничений для явного расчёта маршрутов и сигнальными протоколами (например, RSVP-TE или CRLDP) для организации LSP.

В современном контексте организации трафика MPLS сетевые администраторы могут задать и настроить атрибуты каналов и ресурсные ограничения типа максимальной резервируемой пропускной способности или атрибутов класса

ресурсов для каналов (интерфейсов) в домене MPLS. Протокол на базе состояний каналов с поддержкой расширений TE (IS-IS-TE или OSPF-TE) будет служить для распространения информации о топологии сети а атрибутах каналов всем маршрутизаторам в области маршрутизации. Сетевые администраторы также указывают все LSP, которые начинаются на каждом маршрутизаторе. Для каждого LSP администраторы указывают целевой узел и атрибуты LSP, которые показывают требования, проверяемые при выборе пути. После этого каждый маршрутизатор будет использовать локальный процесс маршрутизации на основе ограничений при расчёте явных путей для всех начинающихся на нем LSP. Далее используется сигнальный протокол для организации LSP. За счёт выделения соответствующей пропускной способности для каналов и LSP обычно удаётся предотвратить или ослабить перегрузки, вызываемые неравномерным распределением трафика.

Используемые для организации трафика атрибуты пропускной способности LSP могут периодически меняться. Базовый подход заключается в том, что выделенная для LSP пропускная способность должна как-то соотноситься с потребностями в пропускной способности для трафика, который реально передаётся через этот LSP. Атрибут трафика для LSP может изменяться с учётом роста трафика и постоянных его смен. Если в результате какой-то неожиданности в сети возникает перегрузка, имеющиеся LSP можно перемаршрутизировать для облегчения ситуации или администратор может организовать новые LSP для переноса части трафика в эти пути. Резервируемая пропускная способность на перегруженных каналах также может быть уменьшена для форсирования перемаршрутизации некоторых LSP.

В домене MPLS матрицу трафика можно оценить по результатам мониторинга трафика в LSP. Такая статистика может использоваться с разными целями, включая планирование и оптимизацию сети. Текущий опыт показывает, что развёрнутые сети MPLS из сотен маршрутизаторов и тысяч LSP вполне реальны. Отметим в заключение, что опыт развёртывания показывает высокую эффективность модели MPLS при организации трафика в сетях IP [XIAO].

Как было отмечено ранее в разделе 7.0, обычно нет прямого контроля за распределением входящего трафика. Следовательно, основной современной междоменной TE является оптимизация распределения исходящего трафика между имеющимися междоменными соединениями. При работе в глобальной сети поддержка способности работы в качестве региональной сети с сохранением преимуществ глобальной сети также становится важной задачей.

Междоменная TE с BGP обычно начинается с размещения множества точек соединения с внешними партнёрами. в местах с высокой плотностью партнёров, находящихся вблизи точек генерации/потребления трафика в собственной сети, а также обеспечивающих невысокую стоимость размещения. В каждом регионе планеты обычно имеется несколько точек обмена трафиком, где организуются соединения между крупнейшими сетями региона. Некоторые проблемы выбора места, связанные с междоменной маршрутизацией, рассмотрены в [AWD5].

После того, как места соединений определены и нужные для этого устройства развёрнуты, принимается решение как лучше обрабатывать получаемые от партнёра маршруты и как отправлять этому партнёру маршруты из своих сетей. Одним из способов организации исходящего трафика для сети с множеством партнёров EBGP является построение иерархии партнёров. Как правило, для всех партнёров устанавливается одно значение Local Preference и для пересылки трафика будут выбираться кратчайшие пути AS. Затем, путём замены входной метрики MED[1] метрикой BGP для маршрутов от разных партнёров, формируется иерархия. Например, для всех партнёров устанавливается Local Preference = 200, предпочтительным внутренним (private) партнёрам присваивается метрика BGP 50, остальным внутренним партнёрам - метрика BGP 100, а внешним (public) - метрика BGP 600. В качестве «предпочтительных» можно выбрать партнёров, имеющих достаточную ёмкость, чья база пользователей больше по сравнению с другими, соединения с которыми дешевле или для которых проще увеличить имеющуюся ёмкость. В сетях с невысокой загрузкой на периметре это работает достаточно хорошо. Эту же концепцию можно использовать для сетей с высокой загрузкой периметра, создав дополнительные уровни градации метрики BGP для более тонкого выбора точек выхода трафика в партнерские сети, с которыми имеется два соединения.

Если изменения ограничиваются заменой входной метрики MED на метрику BGP, изменяются лишь точки выхода для маршрутов с одинаковой длиной AS-Path (процесс принятия решений BGP рассматривает сначала Local Preference, затем длину AS-Path и только потом метрику BGP). Рассмотрим в качестве примера сеть с двумя точками выхода - A и B. Через каждого из партнёров проходит по 40% маршрутов Internet, а остальные 20% связаны с маршрутами пользователей, имеющих двойное подключение к A и B. Предположим, что для обоих партнёров установлено Local Preference = 200 и метрика BGP 100. Если канал к партнёру A окажется перегруженным, увеличение для него метрики BGP при сохранении Local Preference = 200 обеспечит для 20% общего числа маршрутов предпочтительное использование точки выхода через партнёра B. Описанный выше вариант будет применяться в ситуациях, когда все точки выхода к данному партнёру перегружены и трафик нужно целиком отвести от данного партнёра.

Когда перегружена только одна из множества имеющихся точек выхода к данному партнёру, нет необходимости полностью уводить трафик от этого партнёра - достаточно лишь отвести его с перегруженного устройства. Это можно реализовать с помощью пассивных метрик IGP, фильтрации AS-path или фильтрации префиксов.

Время от времени могут требоваться более радикальные изменения - например при наличии «проблемного партнёра», с которым трудно проводить обновления или связность с его сетью стоит слишком дорого. В таких случаях значение Local Preference для этого партнёра можно уменьшить по сравнению с предпочтениями для других партнёров. Это будет эффективно снижать уровень трафика, передаваемого данному партнёру (наличие транзитных партнёров не рассматривается). Такие изменения будут влиять на большой объем трафика и к ним следует обращаться лишь в тех случаях, когда другие методы не дали желаемого результата.

Хотя это не вызывает проблем в региональных сетях, распространение партнерских маршрутов через сеть следует принимать во внимание, когда сеть участвует в партнерских отношениях глобального характера. Иногда в этом контексте на выбор политики BGP могут оказывать влияние деловые аспекты. Например, с точки зрения бизнеса может оказаться неосторожностью работа в глобальной сети и предоставление доступа к глобальной базе клиентов для небольшой сети из отдельной страны. Однако в целях обеспечения своим клиентам качественного сервиса в отдельном регионе хорошая связность может стать необходимостью и для небольших региональных (национальных) сетей. Это может быть достигнуто за счёт создания групп (community) по периметру сети и распространения маршрутов с тегами нужных групп через глобальную сеть. Маршруты от локальных партнёров не будут распространяться в глобальную сеть, тогда как маршрутам от более крупных партнёров может быть разрешено свободное

[1] Multi-exit-discriminator metric - метрика выбора из множества выходов или метрика BGP. В документе используются оба термина.

распространение через всю глобальную сеть. Путём реализации гибкой стратегии для групп (community) могут быть реализованы преимущества использования одного глобального номера AC (ASN[1]) при сохранении преимуществ действующих региональных сетей. Другим вариантом может служить использование разных ASN в разных регионах в результате чего будет увеличиваться длина AS path для маршрутов, анонсируемых этим сервис-провайдером.

9.0 Заключение

В этом документе описаны принципы организации трафика в сети Internet. Документ включает обзор некоторых базовых вопросов, связанных с организацией трафика в сетях IP. Описан контекст TE, модели процессов TE и классификация стилей TE. Представлен также краткий исторический обзор предшествующих работ, связанных с организацией трафика. Приведён также обзор современных методов TE, используемых в действующих сетях. В дополнение к этому документ задаёт набор базовых требований, рекомендаций и вариантов для организации трафика Internet.

10.0 Вопросы безопасности

Этот документ не порождает новых вопросов, связанных с безопасностью.

11.0 Благодарности

Авторы благодарят Jim Boyle за предложения в раздел рекомендаций, Francois Le Faucheur за предложения по аспектам Diffserv, Blaine Christian за предложения по измерениям, Gerald Ash за предложения по маршрутизации в телефонных сетях и текст о методах TE на основе событий, Steven Wright за предложения по управляемости сетей и Jonathan Aufderheide за предложения по TE между доменами с использованием BGP. Отдельная благодарность Randy Bush за предложения по классификации TE с учётом тактических и стратегических методов. Параграф «Обзор действий ITU по части организации трафика» был подготовлен на основе материала Waisum Lai. Полезные отклики и ссылки на связанные материалы были получены от J. Noel Chiappa. На финальном этапе работы были получены дополнительные комментарии от Glenn Grotefeld. В заключение авторы хотят поблагодарить Ed Kern, сопредседателя TEWG за поддержку и комментарии.

12.0 Литература

[ASH2] J. Ash, Dynamic Routing in Telecommunications Networks, McGraw Hill, 1998.

[ASH3] Ash, J., "TE & QoS Methods for IP-, ATM-, & TDM-Based Networks", Work in Progress, March 2001.

[AWD1] D. Awduche and Y. Rekhter, "Multiprocotol Lambda Switching: Combining MPLS Traffic Engineering Control with Optical Crossconnects", IEEE Communications Magazine, March 2001.

[AWD2] D. Awduche, "MPLS and Traffic Engineering in IP Networks", IEEE Communications Magazine, Dec. 1999.

[AWD5] D. Awduche et al, "An Approach to Optimal Peering Between Autonomous Systems in the Internet", International Conference on Computer Communications and Networks (ICCCN'98), Oct. 1998.

[CRUZ] R. L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis", IEEE Transactions on Information Theory, vol. 37, pp. 132-141, 1991.

[DIFF-TE] Le Faucheur, F., Nadeau, T., Tatham, M., Telkamp, T., Cooper, D., Boyle, J., Lai, W., Fang, L., Ash, J., Hicks, P., Chui, A., Townsend, W. and D. Skalecki, "Requirements for support of Diff-Serv-aware MPLS Traffic Engineering", Work in Progress[2], May 2001.

[ELW95] A. Elwalid, D. Mitra and R.H. Wentworth, "A New Approach for Allocating Buffers and Bandwidth to Heterogeneous, Regulated Traffic in an ATM Node", IEEE IEEE Journal on Selected Areas in Communications, 13:6, pp. 1115-1127, Aug. 1995.

[FGLR] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, and J. Rexford, "NetScope: Traffic Engineering for IP Networks", IEEE Network Magazine, 2000.

[FLJA93] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, Vol. 1 Nov. 4., p. 387-413, Aug. 1993.

[FLOY94] S. Floyd, "TCP and Explicit Congestion Notification", ACM Computer Communication Review, V. 24, No. 5, p. 10-23, Oct. 1994.

[FT00] B. Fortz and M. Thorup, "Internet Traffic Engineering by Optimizing OSPF Weights", IEEE INFOCOM 2000, Mar. 2000.

[FT01] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS Weights in a Changing World", www.research.att.com/~mthorup/PAPERS/papers.html.

[HUSS87] B.R. Hurley, C.J.R. Seidl and W.F. Sewel, "A Survey of Dynamic Routing Methods for Circuit-Switched Traffic", IEEE Communication Magazine, Sep. 1987.

[ITU-E600] ITU-T Recommendation E.600, "Terms and Definitions of Traffic Engineering", Mar. 1993.

[ITU-E701] ITU-T Recommendation E.701, "Reference Connections for Traffic Engineering", Oct. 1993.

[ITU-E801] ITU-T Recommendation E.801, "Framework for Service Quality Agreement", Oct. 1996.

[JAM] Jamoussi, B., Editior, Andersson, L., Collon, R. and R. Dantu, "Constraint-Based LSP Setup using LDP", RFC 3212, January 2002.

[1]AS Number.
[2]Работа опубликована в RFC 3564. *Прим. перев.*

[KATZ] Katz, D., Yeung, D. and K. Kompella, "Traffic Engineering Extensions to OSPF", Work in Progress[1], February 2001.

[LNO96] T. Lakshman, A. Neidhardt, and T. Ott, "The Drop from Front Strategy in TCP over ATM and its Interworking with other Control Features", Proc. INFOCOM'96, p. 1242-1250, 1996.

[MA] Q. Ma, "Quality of Service Routing in Integrated Services Networks", PhD Dissertation, CMU-CS-98-138, CMU, 1998.

[MATE] A. Elwalid, C. Jin, S. Low, and I. Widjaja, "MATE: MPLS Adaptive Traffic Engineering", Proc. INFOCOM'01, Apr. 2001.

[MCQ80] J.M. McQuillan, I. Richer, and E.C. Rosen, "The New Routing Algorithm for the ARPANET", IEEE. Trans. On Communications, vol. 28, no. 5, pp. 711-719, May 1980.

[MR99] D. Mitra and K.G. Ramakrishnan, "A Case Study of Multiservice, Multipriority Traffic Engineering Design for Data Networks", Proc. Globecom'99, Dec 1999.

[RFC-1458] Braudes, R. and S. Zabele, "Requirements for Multicast Protocols", RFC 1458, May 1993.

[RFC-1771] Rekhter, Y. and T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771[2], March 1995.

[RFC-1812] Baker, F., "Requirements for IP Version 4 Routers", STD 4, RFC 1812, June 1995.

[RFC-1992] Castineyra, I., Chiappa, N. and M. Steenstrup, "The Nimrod Routing Architecture", RFC 1992, August 1996.

[RFC-1997] Chandra, R., Traina, P. and T. Li, "BGP Community Attributes", RFC 1997, August 1996.

[RFC-1998] Chen, E. and T. Bates, "An Application of the BGP Community Attribute in Multi-home Routing", RFC 1998, August 1996.

[RFC-2205] Braden, R., Zhang, L., Berson, S., Herzog, S. and S. Jamin, "Resource Reservation Protocol (RSVP) - Version 1 Functional Specification", RFC 2205, September 1997.

[RFC-2211] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, September 1997.

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

[RFC-2215] Shenker, S. and J. Wroclawski, "General Characterization Parameters for Integrated Service Network Elements", RFC 2215, September 1997.

[RFC-2216] Shenker, S. and J. Wroclawski, "Network Element Service Specification Template", RFC 2216, September 1997.

[RFC-2328] Moy, J., "OSPF Version 2", STD 54, RFC 2328, July 1997.

[RFC-2330] Paxson, V., Almes, G., Mahdavi, J. and M. Mathis, "Framework for IP Performance Metrics", RFC 2330, May 1998.

[RFC-2386] Crawley, E., Nair, R., Rajagopalan, B. and H. Sandick, "A Framework for QoS-based Routing in the Internet", RFC 2386, August 1998.

[RFC-2474] 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, December 1998.

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

[RFC-2597] Heinanen, J., Baker, F., Weiss, W. and J. Wroclawski, "Assured Forwarding PHB Group", RFC 2597, June 1999.

[RFC-2678] Mahdavi, J. and V. Paxson, "IPPM Metrics for Measuring Connectivity", RFC 2678, September 1999.

[RFC-2679] Almes, G., Kalidindi, S. and M. Zekauskas, "A One-way Delay Metric for IPPM", RFC 2679, September 1999.

[RFC-2680] Almes, G., Kalidindi, S. and M. Zekauskas, "A One-way Packet Loss Metric for IPPM", RFC 2680, September 1999.

[RFC-2702] Awduche, D., Malcolm, J., Agogbua, J., O'Dell, M. and J. McManus, "Requirements for Traffic Engineering over MPLS", RFC 2702, September 1999.

[RFC-2722] Brownlee, N., Mills, C. and G. Ruth, "Traffic Flow Measurement: Architecture", RFC 2722, October 1999.

[RFC-2753] Yavatkar, R., Pendarakis, D. and R. Guerin, "A Framework for Policy-based Admission Control", RFC 2753, January 2000.

[RFC-2961] Berger, L., Gan, D., Swallow, G., Pan, P., Tommasi, F. and S. Molendini, "RSVP Refresh Overhead Reduction Extensions", RFC 2961, April 2000.

[RFC-2998] 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, November 2000.

[RFC-3031] Rosen, E., Viswanathan, A. and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, January 2001.

[RFC-3086] Nichols, K. and B. Carpenter, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", RFC 3086, April 2001.

[RFC-3124] Balakrishnan, H. and S. Seshan, "The Congestion Manager", RFC 3124, June 2001.

[1]Работа опубликована в RFC 3630. *Прим. перев.*
[2]Документ заменён RFC 4271. *Прим. перев.*

[RFC-3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V. and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", RFC 3209, December 2001.

[RFC-3210] Awduche, D., Hannan, A. and X. Xiao, "Applicability Statement for Extensions to RSVP for LSP-Tunnels", RFC 3210, December 2001.

[RFC-3213] Ash, J., Girish, M., Gray, E., Jamoussi, B. and G. Wright, "Applicability Statement for CR-LDP", RFC 3213, January 2002.

[RFC-3270] Le Faucheur, F., Wu, L., Davie, B., Davari, S., Vaahanen, P., Krishnan, R., Cheval, P. and J. Heinanen, "Multi-Protocol Label Switching (MPLS) Support of Differentiated Services", RFC 3270, April 2002.

[RR94] M.A. Rodrigues and K.G. Ramakrishnan, "Optimal Routing in Shortest Path Networks", ITS'94, Rio de Janeiro, Brazil.

[SHAR] Sharma, V., Crane, B., Owens, K., Huang, C., Hellstrand, F., Weil, J., Anderson, L., Jamoussi, B., Cain, B., Civanlar, S. and A. Chui, "Framework for MPLS Based Recovery", Work in Progress.

[SLDC98] B. Suter, T. Lakshman, D. Stiliadis, and A. Choudhury, "Design Considerations for Supporting TCP with Per-flow Queueing", Proc. INFOCOM'98, p. 299-306, 1998.

[SMIT] Smit, H. and T. Li, "IS-IS extensions for Traffic Engineering", Work in Progress[1].

[WANG] Y. Wang, Z. Wang, L. Zhang, "Internet traffic engineering without full mesh overlaying", Proceedings of INFOCOM'2001, April 2001.

[XIAO] X. Xiao, A. Hannan, B. Bailey, L. Ni, "Traffic Engineering with MPLS in the Internet", IEEE Network magazine, Mar. 2000.

[YARE95] C. Yang and A. Reddy, "A Taxonomy for Congestion Control Algorithms in Packet Switching Networks", IEEE Network Magazine, p. 34-45, 1995.

13.0 Адреса авторов

Daniel O. Awduche
Movaz Networks
7926 Jones Branch Drive, Suite 615
McLean, VA 22102
Phone: 703-298-5291
EMail: awduche@movaz.com

Angela Chiu
Celion Networks
1 Sheila Dr., Suite 2
Tinton Falls, NJ 07724
Phone: 732-747-9987
EMail: angela.chiu@celion.com

Anwar Elwalid
Lucent Technologies
Murray Hill, NJ 07974
Phone: 908 582-7589
EMail: anwar@lucent.com

Indra Widjaja
Bell Labs, Lucent Technologies
600 Mountain Avenue
Murray Hill, NJ 07974
Phone: 908 582-0435
EMail: iwidjaja@research.bell-labs.com

XiPeng Xiao
Redback Networks
300 Holger Way
San Jose, CA 95134
Phone: 408-750-5217
EMail: xipeng@redback.com

Перевод на русский язык
Николай Малых

nmalykh@protokols.ru

14.0 Полное заявление авторских прав

Подтверждение

Финансирование функций RFC Editor обеспечивается Internet Society.

[1]Работа опубликована в RFC 3784. *Прим. перев.*