Network Working Group
Request for Comments: 2475
Category: Informational

S. Blake
Torrent Networking Technologies
D. Black
EMC Corporation
M. Carlson
Sun Microsystems
E. Davies
Nortel UK
Z. Wang
Bell Labs Lucent Technologies
W. Weiss
Lucent Technologies
December 1998

# Архитектура дифференцированного обслуживания (Diffserv)

**An Architecture for Differentiated Services**

**Статус документа**

Этот документ содержит информацию для сообщества Internet. Документ не задаёт каких-либо стандартов Internet. Допускается свободное распространение документа.



**Аннотация**

В этом документе определена архитектура для реализации масштабируемой дифференциации услуг в Internet. Эта архитектура обеспечивает масштабирование за счёт агрегирования состояний классификации трафика, которые передаются путём маркировки пакетов IP с использованием поля DS [DSFIELD]. Пакеты классифицируются и маркируются для обеспечения определённой обработки на узлах каждого этапа на пути доставки пакета. Изощрённые методы классификации, маркировки, реализации политики, формовки трафика требуются только на границах сетей или хостах. Сетевые ресурсы выделяются для потоков трафика правилами предоставления услуг, которые определяют маркировку и кондиционирование трафика на входе в сеть с поддержкой дифференцированных услуг, а также пересылку трафика внутри такой сети. На основе этого можно построить широкий спектр сетевых услуг.

# Оглавление

# *1. Введение*

## 1.1 Обзор

В этом документе определена архитектура для реализации масштабируемой дифференциации услуг в Internet. «Услуга» определяет некоторые значимые характеристики передачи пакетов в одном направлении по одному или множеству путей внутри сети. Эти характеристики могут быть заданы количественно или статистически в части пропускной способности, задержки, вариаций задержки и/или потери пакетов или могут быть определены с форме некоторого относительного уровня приоритета доступа к сетевым ресурсам. Дифференциация услуг желательна для аккомодации разнородных требований приложений и ожиданий пользователей, а также позволяет дифференцировать стоимость услуг Internet.

Архитектура включает множество функциональных элементов, реализованных на узлах сети, включая небольшой набор аспектов пересылки на каждом этапе, функции классификации трафика и функции кондиционирования трафика, в том числе измерение, маркировку, формование и применение правил. Эта архитектура обеспечивает масштабируемость за счёт реализации комплексных функций классификации и кондиционирования исключительно на граничных узлах сетей и выполнения определённых операций на каждом этапе пересылки применительно к агрегатам трафика, маркированным с использованием поля DS в заголовках IPv4 или IPv6 [DSFIELD]. Задаётся поведение на каждом этапе[1] для того, чтобы обеспечить разумную гранулярность распределения буферов и полосы пропускания на каждом узле между конкурирующими потоками трафика. Потоки для каждого приложения и состояние пересылки для каждого заказчика не требуется поддерживать в ядре сети. Поддерживаются различия между:

- услугами, обеспечиваемыми для каждого агрегата трафика;
- функциями кондиционирования и поведением на каждом этапе, используемыми для реализации услуг;
- значением поля DS (код DS), используемым для маркировки пакетов с целью задания поведения на этапах;
- механизмами реализации конкретных узлов, которые обеспечивают режим обработки на отдельных этапах.

Правила предоставления услуг и кондиционирования трафика отделены от пересылки внутри сети, чтобы обеспечить возможность реализации широкого набора вариантов поведения для услуг и пространство для его расширения.

Эта архитектура обеспечивает дифференциацию услуг только для одного направления потока трафика и, следовательно, является асимметричной. Разработка дополнительной симметричной архитектуры является предметом исследования и выходит за пределы данного документа. Более подробно о таких исследованиях можно узнать из работы [EXPLICIT].

В параграфе 1.2 определены основные термины, используемые в документе. В параграфе 1.3 перечислены требования, порождаемые этой архитектурой, а параграф 1.4 содержит краткое сравнение с другими моделями дифференциации услуг. В разделе 2 приводится детальное рассмотрение компонент архитектуры. Раздел 3 содержит рекомендации по спецификации поведения на каждом этапе. В разделе 4 обсуждаются вопросы взаимодействия с узлами и сетями, которые не реализуют дифференциации услуг в соответствии с этим документом и [DSFIELD]. В разделе 5 рассматриваются вопросы доставки групповых услуг, а раздел 6 посвящён вопросам безопасности и рассмотрению туннелей.

## 1.2 Терминология

В этом параграфе содержится концептуальный обзор терминов, используемых в документе. Некоторые из этих терминов более подробно рассмотрены в последующих разделах.

***Behavior Aggregate (BA)***
Агрегат режима (поведения) DS.

***BA classifier***
Классификатор, выбирающий пакеты исключительно на основе значения поля DS.

***Boundary link - граничное соединение***
Канал, соединяющий краевые узлы двух доменов.

***Classifier - классификатор***
Элемент, выделяющий пакеты на основе содержимого заголовков в соответствии с заданными правилами.

***DS behavior aggregate - агрегат режима DS***
Набор пакетов с одинаковым кодом DS, проходящих по каналу в определённом направлении.

***DS boundary node - граничный узел DS***
Узел DS, соединяющий домен DS с узлом другого домена DS или узлом домена, не поддерживающего DS.

***DS-capable***
Способность реализации дифференцированных услуг в соответствии с описываемой архитектурой; обычно этим термином обозначают домены, состоящие из узлов, поддерживающих DS.

***DS codepoint - код DS***
Значение компоненты DSCP поля DS, используемое для выбора PHB.

***DS-compliant - поддерживающий DS***
Поддерживающий функции дифференцированного обслуживания и поведение, определённое в [DSFIELD], данном документе и других документах, посвящённых дифференцированному обслуживанию; обычно этот термин используют применительно к узлу или устройству.

***DS domain - домен DS***
Поддерживающий DS домен - связное множество узлов, использующих общий набор правил обслуживания и определений PHB.

[1]Per-hop behavior.

***DS egress node - выходной узел DS***
Граничный узел DS, который обслуживает трафик, выходящий из домена DS.

***DS ingress node - входной узел DS***
Граничный узел DS, который обслуживает трафик, входящий в домен DS.

***DS interior node - внутренний узел DS***
Узел DS, не являющийся граничным.

***DS field - поле DS***
Октет TOS в заголовке IPv4 или октет Traffic Class в заголовке IPv6, интерпретируемый в соответствии с определением [DSFIELD]. Биты поля DSCP определяют код DS, а остальные биты в настоящее время не используются.

***DS node - узел DS***
Поддерживающий DS узел.

***DS region - зона DS***
Связное множество доменов DS, которые могут предлагать дифференцированное обслуживание на путях, проходящих через эти домены.

***Downstream DS domain - нисходящий домен DS***
Домен DS обрабатывающий нисходящий поток трафика на граничном соединении.

***Dropper - отбрасыватель***
Устройство, выполняющее отбрасывание пакетов.

***Dropping - отбрасывание***
Процесс уничтожения пакетов на основе заданных правил (политики).

***Legacy node - унаследованный узел***
Узел, реализующий IPv4 Precedence в соответствии с [RFC791,RFC1812], но не поддерживающий DS.

***Marker - маркировщик***
Устройство, выполняющее маркирование пакетов.

***Marking - маркировка***
Процесс установки кода DS в пакетах в соответствии с заданными правилами; используются также термины pre-marking (предварительная маркировка), re-marking (перемаркировка).

***Mechanism - механизм***
Определённый алгоритм или операция (например, дисциплина очереди), поддерживаемый на узле для реализации одного или нескольких режимов PHB.

***Meter - измеритель***
Устройство, выполняющее измерения.

***Metering - измерение***
Процесс измерения зависящих от времени свойств (например, скорости) определённого потока трафика, выбранного классификатором. Моментальное значение может использоваться для управления маркировкой, формовкой или отбрасыванием пакетов и/или для целей учёта и измерения.

***Microflow - микропоток***
Один экземпляр потока пакетов между приложениями, который идентифицируется адресами и номерами портов отправителя и получателя, а также идентификатором протокола.

***MF Classifier - классификатор MF***
Многокомпонентный (MF[1]) классификатор, который выбирает пакеты на основе содержимого некого произвольного числа полей заголовков; обычно набор полей включает ту или иную комбинацию адресов и номеров портов отправителя и получателя, поле DS, идентификатор протокола.

***Per-Hop-Behavior (PHB) - режим поэтапной обработки***
Наблюдаемое извне поведение пересылки пакетов узлом DS по отношению агрегату BA.

***PHB group - группа PHB***
Набор из одного или множества PHB, которые могут выполняться только совместно вследствие общих ограничений, применяемых ко всем PHB данной группы (например, управление очередями). Группа PHB является базовым элементом архитектуры, позволяющим совместно реализовать несколько связанных режимов (например, четыре профиля отбрасывания пакетов). Одиночный PHB является частным случаем группы PHB.

***Policing - реализация политики***
Процесс отбрасывания пакетов в потоке трафика в зависимости от состояния соответствующего измерителя, реализующего профиль трафика.

***Pre-mark - предварительная маркировка***
Установка кода DS для пакета до его входа в нисходящий домен DS.

***Provider DS domain - DS-домен провайдера***
Поддерживающий DS поставщик услуг для домена отправителя.

***Re-mark - перемаркировка***
Изменение кода DS для пакета, обычно выполняемое маркировщиком в соответствии с TCA.

***Service - сервис, обслуживание***
Интегральная обработка трафика определённого подмножества заказчиков (сквозная или в домене DS).

***Service Level Agreement (SLA) - соглашение об уровне обслуживания***
Сервисный контракт между заказчиком и сервис-провайдером, задающий услуги по пересылке пакетов, предоставляемые заказчику. Заказчиком может служить конечный пользователь (домен-отправитель) или другой домен DS (восходящий[2] домен). SLA может включать правила кондиционирования трафика, которые полностью или частично задают TCA.

***Service Provisioning Policy - правила обслуживания***
Правила, определяющие настройку кондиционеров трафика на граничных узлах DS и отображение потоков трафика на агрегаты режимов DS для обеспечения набора услуг.

***Shaper - формовщик***
Устройство, выполняющее формовку трафика.

***Shaping - формовка***
Процесс внесения задержки для пакетов в потоке трафика с целью реализации определённого профиля трафика.

[1]Multi-field - по множеству полей.
[2]Upstream domain.

***Source domain - домен-отправитель***

Домен, содержащий узлы, генерирующие трафик, который получает определённый тип сервиса.

***Traffic conditioner - кондиционер трафика***

Элемент, выполняющий функции кондиционирования трафика, который может включать измерители, маркировщики, отбрасыватели и формовщики. Кондиционеры трафика обычно используются только на граничных узлах DS. Кондиционер может выполнять перемаркировку потоков трафика, отбрасывать пакеты или формовать трафик для изменения временных характеристик в соответствии с профилем трафика.

***Traffic conditioning - кондиционирование трафика***

Функции управления, выполняемые для реализации правил, указанных TCA, включая измерение, маркировку, формовку и реализацию политики.

***Traffic Conditioning Agreement (TCA) - соглашение о кондиционировании трафика***

Соглашение между задающими трафик правилами и любым соответствующим профилем трафика, а также правилами измерения, маркировки, отбрасывания и/или формовки, которые применимы к потоку трафика, выбранному классификатором. TCA включает все правила кондиционирования трафика, явно заданные в SLA, вместе со всеми правилами, неявно следующими из имеющих отношение к делу требований и/или правил обслуживания домена DS.

***Traffic profile - профиль трафика***

Описание временных характеристик потока трафика (таких, как скорость и пиковая скорость).

***Traffic stream - поток трафика***

Административно значимый набор из одного или множества микропотоков, проходящих через сегмент пути. Поток трафика может состоять из множества активных микропотоков, выбранных отдельным классификатором.

***Upstream DS domain - восходящий домен DS***

Домен DS обрабатывающий восходящий поток трафика на граничном канале.

## 1.3 Требования

История Internet представляет собой постоянный рост числа хостов, появление новых приложений, рост ёмкости сетевой инфраструктуры. В обозримом будущем предполагается продолжение такого роста. Масштабируемая архитектура дифференциации обслуживания должна обеспечивать возможность адаптации к продолжающемуся росту.

Ниже перечислены идентифицированные требования, решаемые данной архитектурой:

- следует обеспечивать аккомодацию к широкому спектру услуг и правил их предоставления в масштабах группы сетей или в сквозном режиме;
- следует разрешать «отвязывание» сервиса от конкретного приложения;
- следует работать с существующими приложениями, не требуя изменения API или программ хостов (предполагается развёртывание нужных классификаторов, маркировщиков и других средств кондиционирования трафика);
- следует отвязать функции кондиционирования трафика и предоставления услуг от режимов пересылки, реализованных в узлах ядра сети;
- следует обеспечить независимость от поэтапной (hop-by-hop) сигнализации приложений;
- следует требовать только минимальный набор режимов пересылки, чтобы сложность его реализации не доминировала над стоимостью сетевых устройств и не возникало «пробок» в скоростных системах будущего;
- следует избегать поддержки состояний для отдельных микропотоков и отдельных заказчиков в узлах ядра сети;
- с ядре сети следует использовать только агрегированные классификационные состояния;
- в ядре сети следует использовать простые реализации классификаторов пакетов (классификаторов BA);
- следует поддерживать разумное взаимодействие с узлами, не поддерживающими DS;
- следует поддерживать постепенное развёртывание.

## 1.4 Сравнение с другими моделями

Описываемая в этом документе архитектура дифференцированного обслуживания может противоречить другим моделям дифференциации услуг. Мы классифицируем эти альтернативные модели по нескольким категориям - маркировка относительного приоритета, маркировка сервиса, коммутация меток, интегрированные службы/RSVP и статическая классификация для каждого этапа пересылки.

Примерами маркировки относительного приоритета могут служить маркировка IPv4 Precedence [RFC791], 802.5 Token Ring [TR] и принятая по умолчанию интерпретация классов трафика 802.1p [802.1p]. В этой модели приложения, хосты или узлы-посредники выбирают относительный уровень приоритета или «предпочтения» для пакета (например, приоритет по задержке или отбрасыванию), а сетевые узлы на пути передачи применяют соответствующий значению приоритета в поле заголовка режим пересылки. Наша архитектура может задавать роль и важность граничных узлов и кондиционеров трафика, благодаря этому наша модель обеспечивает более общий подход к поэтапному режиму пересылки, нежели относительный приоритет для задержки или отбрасывания.

Примером модели маркировки услуг является IPv4 TOS в соответствии с определением [RFC1349]. В этом примере каждый пакет маркируется запросом «типа обслуживания», который принимать значения «минимизировать задержку», «максимизировать пропускную способность», «максимизировать надёжность» или «минимизировать стоимость». Узлы сети могут выбирать путь маршрутизации или режим пересылки, который подходит для запрошенного типа обслуживания. Эта модель имеет тонкие отличия от нашей архитектуры. Отметим, что мы не описываем использование поля DS в качестве входной информации для выбора маршрута. Маркировка TOS, определённая в [RFC1349], является весьма общей и не охватывает весь диапазон возможной семантики обслуживания. Более того, с каждым пакетом связывается запрос на обслуживание, а семантика некоторых услуг может зависеть от

агрегированного режима пересылки последовательности пакетов. Модель маркировки услуг достаточно сложно приспособить к будущему росту числа и диапазона услуг (поскольку пространство кодов достаточно мало) и эта модель включает настройку конфигурации ассоциаций «TOS->режим пересылки» на каждом узле ядра сети. Стандартизация маркировки услуг предполагает стандартизацию предложений услуг, которая выходит за пределы компетентности IETF. Отметим, что при распределении пространства кодов DS, часть кодов зарезервирована для обеспечения возможности локального управления кодами, позволяющего провайдерам поддерживать локальную семантику маркировки услуг [DSFIELD].

Примеры модели коммутации меток (или виртуальных устройств) включают Frame Relay, ATM и MPLS [FRELAY, ATM]. В этой модели поддерживается состояние пересылки и управления трафиком или QoS для потоков трафика на каждом интервале пути через сеть. Агрегаты трафика различной гранулярности ассоциируются с путём коммутации меток на входном узле, а пакеты/ячейки в каждом пути коммутации меток маркируются меткой пересылки, которая используется для поиска следующего интервала, а также режима пересылки и заменяется на каждом этапе пересылки. Эта модель обеспечивает тонкую гранулярность распределения ресурсов для потоков трафика, поскольку значения меток не имеют глобальной значимости и важны только для одного канала. Следовательно, можно резервировать ресурсы для агрегирования пакетов/ячеек, принятых на канале с определённой меткой, а семантика коммутации меток определяет выбор следующего интервала, позволяя передавать потоки трафика через специально созданные для них сетевые пути. Такое улучшение гранулярности связано с расходами на выполнение дополнительных требований по управлению и настройке для организации и поддержке путей коммутации меток. В дополнение к этому количество состояний пересылки, поддерживаемых на каждом узле, растёт пропорционально числу краевых узлов сети в лучшем случае (в предположении путей коммутации меток multipoint-to-point) и пропорционально квадрату числа краевых узлов в худшем случае, когда обеспечиваются пути коммутации меток «от края до края» с гарантированными ресурсами.

Модель интегрированных услуг/RSVP[1] основана на традиционной пересылке дейтаграмм по умолчанию, но позволяет отправителям и получателям обмениваться сигнальными сообщениями, которые позволяют организовать дополнительную классификацию пакетов и состояние пересылки на каждом узле между ними [RFC1633, RSVP]. В отсутствии агрегирования состояний их число на каждом узле растёт пропорционально числу одновременных резервирований, которое для скоростных каналов может быть очень большим. Эта модель также требует от приложений поддержки протокола RSVP. Для агрегирования состояний интегрированных услуг/RSVP в ядре сети могут использоваться механизмы дифференциации обслуживания [Bernet].

Вариант модели интегрированных услуг/RSVP позволяет избавиться от требования поэтапной[2] сигнализации за счёт использования только «статической» классификации и политики пересылки, реализуемых на каждом узле пути через сеть. Политика изменяется административными методами, а не в ответ на изменение моментального состояния микропотоков в сети. Требования к состояниям для этого варианта потенциально жёстче, нежели при использовании RSVP, особенно для магистральных узлов, поскольку число статических правил, которые могут применяться на узле в течение времени, может быть больше числа активных сеансов «отправитель-получатель», которые одновременно организуют резервирование состояния на узле. Хотя поддержка большого числа правил классификации и пересылки может может требовать вполне приемлемых вычислительных ресурсов, издержки на управление, связанные с организацией и поддержкой этих правил на каждом узле магистральной сети, через которую проходит трафик, могут быть достаточно велики.

Хотя мы противопоставили альтернативные модели и дифференциацию обслуживания, следует отметить, что эти методы могут использоваться для расширения режимов и семантики дифференцированного обслуживания в коммутируемой инфраструктуре канального уровня (например, ЛВС 802.1p, магистрали Frame Relay/ATM), соединяющей узлы DS, а также в том случае, когда в качестве внутридоменной технологии может использоваться MPLS. Ограничения, вносимые конкретной технологией канального уровня в отдельной зоне домена DS (или в сети, обеспечивающей доступ к доменам DS), могут загрублять дифференциацию обслуживания. В зависимости от отображения PHB на различные услуги канального уровня и способа распределения пакетов по ограниченному набору классов приоритета (или виртуальных устройств различных категорий и производительности), поддерживаться могут все или часть PHB (некоторые могут оказаться нежелательными).

# *2. Архитектурная модель дифференциации услуг*

Архитектура дифференцированного обслуживания основана на простой модели, где трафик классифицируется на входе в сеть и, возможно, кондиционируется на границах сети; классифицированный трафик относится к тому или иному агрегату режима обработки. Каждый агрегат идентифицируется одним кодом DS. Внутри ядра сети пакеты пересылаются в соответствии с заданным для этапа режимом, связанным с кодом DS. В этом разделе обсуждаются компоненты зоны дифференцированного обслуживания, функции классификации и кондиционирования трафика, а также вопросы дифференцированного обслуживания, реализуемого через комбинацию кондиционирования трафика и пересылки на основе PHB.

## 2.1 Домен дифференцированных услуг

Домен DS представляет собой связное множество узлов DS, которые работают с общим набором правил предоставления услуг, и набором групп PHB, реализованных на каждом узле. Домен DS имеет хорошо определённые границы, состоящие из граничных узлов DS, которые классифицируют и, возможно, кондиционируют входящий трафик для обеспечения проходящим через домен пакетам корректной маркировки для выбора PHB из групп PHB, поддерживаемых в домене. Узлы в домене DS выбирают режим пересылки для пакетов на основе кода DS, отображая этот код на один из поддерживаемых режимов PHB с использованием рекомендованного отображения код->PHB или локально управляемого отображения [DSFIELD]. Включение не поддерживающих DS узлов в домен DS может приводить к непредсказуемому изменению производительности и осложнять возможность выполнения сервисных соглашений (SLA).

Домен DS состоит из одной или множества сетей с общим администрированием (например, внутренняя сеть организации или сеть ISP). Администрация домена отвечает за обеспечение достаточных ресурсов и/или их резервирование для обеспечения SLA, предлагаемых доменом.

[1] Integrated Services/RSVP.
[2] Hop-by-hop.

### *2.1.1 Граничные и внутренние узлы DS*

Домен DS состоит из граничных и внутренних узлов DS. Граничные узлы DS соединяют домен DS с другими доменами (DS или не-DS), тогда, как внутренние узлы DS соединяют между собой только внутренние или граничные узлы DS в одном домене DS.

Как граничные, так и внутренние узлы DS должны быть способны применять подходящий режим PHB к пакетам на основе кода DS; в противном случае поведение может стать непредсказуемым. В дополнение от граничных узлов DS может требоваться выполнение функций кондиционирования трафика в соответствии с соглашением о кондиционировании (TCA) между доменом DS и доменами-партнерами, с которыми он соединён (см. параграф 2.3.3).

Внутренние узлы могут выполнять ограниченный набор функций кондиционирования трафика, таких, как перемаркировка кодов DS. Внутренние узлы могут также выполнять более сложные функции классификации и кондиционирования трафика, аналогичные функциям граничных узлов DS (см. параграф 2.3.4.4).

Хост в сети, содержащей домен DS, может действовать как граничный узел DS для трафика приложений, работающих на этом хосте; мы говорим, следовательно, что этот хост находится в домене DS. Если хост не функционирует, как граничный узел, тогда топологически ближайший к хосту узел DS, действует, как граничный узел DS для трафика этого хоста.

### *2.1.2 Входные и выходные узлы DS*

Граничные узлы DS действуют как входные и выходные узлы DS для различных направлений трафика. Трафик входит в домен DS через входной узел DS и покидает домен DS через выходной узел DS. Входной узел DS отвечает за то, что входящий в DS трафик соответствует всем TCA между данным доменом и другими доменами, к которым подключён входной узел. Выходной узел DS может выполнять функции кондиционирования для трафика, пересылаемого в непосредственно подключённые партнерские домены в зависимости от условий TCA между двумя доменами. Отметим, что граничный узел DS может функционировать как внутренний узел DS для некоторого множества интерфейсов.

## 2.2 Зона дифференцированного обслуживания

Зона дифференцированного обслуживания (зона DS[1]) представляет собой набор из одного или множества доменов DS. Зоны DS способны поддерживать дифференциацию обслуживания на пути, проходящем через домены зоны.

Домены DS в зоне DS могут поддерживать различные группы PHB внутри себя и разные отображения код->PHB. Однако для того, чтобы разрешить услуги, которые могут предоставляться через несколько доменов, партнерские домены DS должны организовать партнерские SLA, которые определяют (явно или неявно) TCA, задающие кондиционирование трафика из одного домена DS в другой на границе между двумя доменами DS.

Возможно, что несколько доменов DS внутри зоны DS будут использовать общий набор правил обслуживания и поддерживать общий набор групп PHB и отображений кодов - это избавляет от необходимости кондиционирования трафика между этими доменами DS.

## 2.3 Классификация и кондиционирование трафика

Дифференцированные услуги предоставляются через границу домена DS путём организации SLA между восходящей (upstream) сетью и нисходящим (downstream) доменом DS. SLA может задавать правила классификации и перемаркировки трафика, а также профили трафика и действия для потоков трафика, который относится или не относится к профилю (см. параграф 2.3.2). TCA между доменами организуется (явно или неявно) на основе таких SLA.

Правила классификации трафика задают подмножество трафика, для которого может предоставляться дифференцированное обслуживание путём кондиционирования и/или повторного отображения (путём перемаркировки кодов DS) на один или множество агрегатов поведения внутри домена DS.

Кондиционирование трафика включает измерение, формование, применение политики и/или перемаркировку для того, чтобы входящий в домен DS трафик соответствовал правилам, заданным в TCA, согласно политике предоставления услуг этого домена. Кондиционирование трафика требуется в зависимости от предлагаемых услуг и может представлять собой набор операций от простой перемаркировки кодов до выполнения сложных операций формовки и исполнения политики. Детали кондиционирования трафика согласуются между сетями и выходят за пределы этого документа.

### *2.3.1 Классификаторы*

Классификаторы пакетов выбирают пакеты в потоке трафика на основе значений некоторых полей заголовков. Мы определяем два типа классификаторов. Классификатор BA[2] разделяет пакеты исключительно на основе кода DS. Классификатор MF[3] выделяет пакеты на основе комбинации полей заголовка, включая такие поля, как адреса отправителя и получателя, поле DS, идентификатор протокола, номера портов отправителя и получателя, а также других параметров (например, входной интерфейс).

Классификаторы служат для «направления» пакетов, соответствующих определённому правилу, элементу кондиционирования трафика для последующей обработки. Классификаторы должны настраиваться с помощью той или иной процедуры управления в соответствии с подходящим TCA.

Классификаторам следует проверять подлинность информации, используемой для выбора пакетов (см. раздел 6).

Отметим, что при фрагментации «восходящего» пакета, классификаторы MF, которые проверяют содержимое заголовков транспортного уровня, могут некорректно классифицировать фрагменты, следующие за первым. Возможным решением этой проблемы служит поддержка информации о состоянии фрагментации. Однако такое решение не является достаточно общим, поскольку в восходящем потоке может меняться порядок доставки и даже пути маршрутизации. Правила обработки фрагментов выходят за пределы этого документа.

[1] В оригинале - DS Region. *Прим. перев.*
[2] Behavior Aggregate - агрегат поведения.
[3] Multi-Field - по множеству полей.

## *2.3.2 Профили трафика*

Профиль трафика задаёт временные параметры потока трафика, выбранного классификатором. Это обеспечивает правила для определения соответствия конкретного пакета данному профилю. Например, профиль на основе token bucket[1] может иметь вид:

```
codepoint=X, use token-bucket r, b
```

Такой профиль показывает, что все пакеты с кодом DS = X, следует сравнивать с уровнем потока r и размером выбросов трафика b. В этом примере профилю не соответствуют пакеты из потока трафика, принимаемые в те моменты, когда в корзине не остаётся маркеров. Концепцию можно расширить, используя более двух уровней. Например, можно определить и реализовать профиль с множеством уровней маркирования.

К пакетам, соответствующим и не соответствующим профилю, могут применяться различные операции или включаться те или иные механизмы учёта. Соответствующим профилю пакетам может разрешаться вход в домен DS без дальнейшего кондиционирования или, наоборот, их код DS может быть изменён. Последнее происходит в тех случаях, когда для DS изначально установлено отличное от принятого по умолчанию значение [DSFIELD], или пакеты входят в домен DS, использующий другую группу PHB или иное отображение код->PHB для этого потока трафика. Пакеты, не соответствующие профилю, могут помещаться в очередь, пока они не войдут в профиль (формовка), отбрасываться (политика), маркироваться новым кодом (перемаркировка) или пересылаться без изменений, но с включением некой процедуры учёта. Не соответствующие профилю пакеты могут отображаться в один или несколько агрегатов поведения, которые являются «подчинёнными» в плане скорости пересылки по отношению к BA, куда отображаются соответствующие профилю пакеты.

Отметим, что профиль трафика является необязательной частью TCA и его использование зависит от специфики предоставляемого сервиса и политики предоставления услуг в домене.

## *2.3.3 Кондиционеры трафика*

Кондиционеры трафика могут включать измерители, маркировщики, формовщики и отбрасыватели пакетов. Поток трафика выбирается классификатором, который направляет пакеты логическому экземпляру кондиционера пакетов. Измеритель используется (когда это применимо) для сравнения потока трафика с профилем. Состояние измерителя по отношению к конкретному пакету (например, соответствие или несоответствие профилю) может использоваться для маркировки, отбрасывания или формовки.

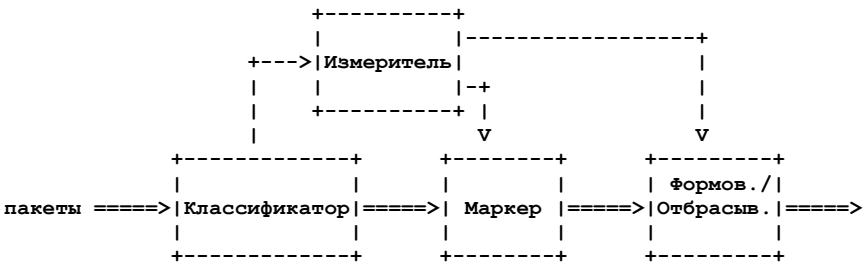


*Рисунок 1. Логическое представление классификатора пакетов и кондиционера трафика.*

Когда пакеты выходят из кондиционера трафика граничного узла DS, для кода DS в каждом пакете должно быть установлено подходящее значение.

Рисунок 1 показывает блок-схему классификатора и кондиционера трафика. Отметим, что кондиционер трафика может не включать все 4 элемента. Например, в тех случаях, когда профили трафика не используются, пакеты могут проходить только через классификатор и измеритель.

### 2.3.3.1 Измерители

Измерители трафика сравнивают временные параметры потока пакетов, выбранных классификатором, с профилем трафика, заданным в TCA. Измеритель передаёт информацию о состоянии другим функциям кондиционирования для включения соответствующих операций по отношению к пакетам, которые соответствуют и не соответствуют профилю.

### 2.3.3.2 Маркировщики

Маркировщики пакетов устанавливают значения поля DS в заголовках пакетов, добавляя маркированные пакеты к соответствующему агрегату поведения DS. Маркировщик можно настроить на маркирование всех пакетов одним значением кода или маркировку каждого пакета с использованием одного из множества кодов, служащих для выбора PHB в группе PHB в зависимости от результата измерения. Изменение кода маркировщиком называют перемаркировкой.

### 2.3.3.3 Формовщики

Формовщики задерживают некоторые или все пакеты в потоке трафика для приведения этого потока в соответствие с профилем трафика. Формовщик обычно использует буфер конечного размера и пакеты могут отбрасываться после исчерпания буферной ёмкости.

### 2.3.3.4 Отбрасыватели

Отбрасыватели сбрасывают все или часть пакетов в потоке трафика для приведения потока в соответствие с профилем трафика. Этот процесс называют применением политики[2] к потоку. Отметим, что отбрасыватель может быть реализован, как частный случай формовщика, путём установки нулевого (или достаточно малого) значения для размера буфера задержки.

[1]Корзина с маркерами. При получении пакета ему выделяется маркер из корзины фиксированного начального размера. Скорость пополнения маркеров в корзине постоянна. Пакеты соответствуют профилю, пока для них в корзине имеются маркеры. *Прим. перев.*
[2]Policing.

### *2.3.4 Расположение кондиционеров трафика и классификаторов MF*

Кондиционеры трафика обычно размещаются в DS на входных и выходных граничных узлах, но могут также размещаться на узлах внутри домена DS или в доменах, не поддерживающих DS.

#### 2.3.4.1 Внутри исходного домена

Определим исходный домен, как домен, содержащий узел (или узлы), порождающий трафик, получаемый определённой службой. Источники трафика и промежуточные узлы внутри исходного домена могут выполнять функции классификации и кондиционирования. Трафик, проходящий из исходного домена через границу, может быть промаркирован непосредственными источниками или промежуточными узлами до того, как пакеты покинут исходный домен. Эта операция называется предварительной маркировкой.

Рассмотрим пример компании, в соответствии с политикой которой пакетам от CEO[1] следует присваивать высший приоритет. Хост CEO может маркировать поле DS всех исходящих пакетов значением DS, показывающим «высший приоритет». Кроме того, первый маршрутизатор, с которым соединён хост CEO, может классифицировать трафик и маркировать пакеты CEO корректным кодом DS. Такой трафик с высоким приоритетом также может быть кондиционирован неподалёку от источника так, чтобы ограничить уровень трафика с высшим приоритетом, пересылаемого от конкретного источника.

Маркировка пакетов вблизи источника трафика обеспечивает некоторые преимущества. Во-первых, источник трафика может принимать во внимание предпочтения приложений при решении вопроса о том, какие пакеты должны обрабатываться в первую очередь. Кроме того, классификация пакетов выполняется значительно проще, пока эти пакеты не были агрегированы с пакетами из других источников, поскольку число правил классификации, которые требуется использовать, в таких случаях существенно меньше.

Поскольку маркировка пакетов может осуществляться на множестве узлов, исходный домен DS отвечает за то, что агрегированный трафик в направлении домена DS его провайдера соответствует TCA. Дополнительные механизмы распределения ресурсов типа брокеров полосы или RSVP могут использоваться для динамического выделения ресурсов отдельному агрегату поведения DS в сети провайдера [2BIT, Bernet]. Граничному узлу исходного домена следует также осуществлять мониторинг соответствия TCA, этот узел может при необходимости выполнять перемаркировку и формовку или применять для пакетов правила политики.

#### 2.3.4.2 На границе домена DS

Потоки трафика могут классифицироваться, маркироваться или кондиционироваться иными способами на конце граничного соединения (выходной узел DS восходящего домена или входной узел DS нисходящего домена). В SLA между доменами следует указывать, какой из доменов отвечает за отображение потоков трафика на агрегаты поведения DS и кондиционирование таких агрегатов в соответствии с подходящим TCA. Однако входной узел DS должен предполагать, что входящий трафик может не соответствовать TCA, и должен быть готов к понуждению исполнения TCA в соответствии с локальной политикой.

Когда пакеты маркируются заранее и кондиционируются в восходящем домене, потенциально снижается число правил классификации и кондиционирования трафика в нисходящем домене DS. В таких обстоятельствах от нисходящего домена DS может потребоваться только перемаркировка или исполнение правил для входящих агрегатов поведения с целью исполнения TCA. Однако более изощрённые типы сервиса, которые зависят от пути или источника, могут требовать классификации MF на входных узлах нисходящего домена DS.

Если входной узел DS подключён к восходящему домену, который не поддерживает DS, этот узел должен быть способен выполнить для входящего трафика все функции кондиционирования трафика.

#### 2.3.4.3 В доменах, не поддерживающих DS

Источники трафика или промежуточные узлы в доменах, не поддерживающих DS, могут реализовать кондиционирование заранее помеченного трафика до того, как он попадёт на вход нисходящего домена DS. Таким способом может быть скрыта локальная политика для классификации и маркировки.

#### 2.3.4.4 На внутренних узлах DS

Хотя базовая архитектура предполагает, что комплексные функции классификации и кондиционирования трафика локализованы только на входных и выходных граничных узлах, реализация этих функций внутри сети не запрещается. Например, на трансокеанских каналах могут быть реализованы более жёсткие правила доступа, требующие классификации MF, и кондиционирование на восходящем узле канала. Такое решение может ограничивать масштабирование по причине потенциально большого числа правил классификации и кондиционирования, которые нужно поддерживать.

## 2.4 Поведение на этапе

Поведение на этапе (PHB[2]) представляет собой описание наблюдаемого извне режима пересылки узла DS, применяемое к конкретному агрегату поведения DS. «Режим пересылки» представляет собой базовую концепцию в этом контексте. Например, в ситуации, когда только один агрегат поведения занимает канал, наблюдаемый режим пересылки (т. е., потери, задержки и их вариации) будет зачастую зависеть только от относительной загрузки канала (в предположении, что режим основан на снижающей объем работы дисциплине планирования). Существенные различия в поведении наблюдаются для тех случаев, когда множество агрегатов поведения полностью используют ресурсы буферизации и полосы на узле. PHB представляет собой способ, посредством которого узел выделяет ресурсы для агрегатов поведения, и располагается поверх базовых поэтапных механизмов распределения ресурсов, которые могут быть полезны для поддержки дифференцированного обслуживания.

Примером наиболее простого PHB является тот случай, когда гарантируется выделение минимальной полосы X% от суммарной полосы канала (в течение некого разумного интервала) для агрегата поведения. Этот PHB можно достаточно просто измерить в широком диапазоне условий конкуренции трафика. Несколько более сложный вариант

[1] Chief Executive Officer - исполнительный директор.
[2] Per-hop behavior.

PHB будет гарантировать минимальное выделение X% от полосы канала с пропорциональным беспристрастным распределением любой избыточной полосы. В общем случае PHB может зависеть от некоторых ограничений на характеристики трафика связанного агрегата поведения или характеристик других агрегатов поведения.

PHB могут задаваться в терминах приоритета их ресурсов (например, буфер, полоса) по отношению к другим PHB или в терминах их относительных значениях наблюдаемых характеристик трафика (например, задержка, потери). PHB могут использоваться в качестве базовых элементов при распределении ресурсов, для согласованности их следует задавать как группы PHB. Группы PHB обычно имеют общие ограничения, применяемые к каждому PHB в данной группе (такие, как планирование очередей пакетов или правила буферизации). Отношения между PHB в группе могут выражаться в терминах абсолютного или относительного приоритета (например, приоритет отбрасывания может задаваться детерминированным или стохастическим порогом), но это не требуется (например, совместное использование N одинаковых каналов). Отдельный изолированный PHB является частным случаем группы PHB.

PHB реализуются на узлах посредством тех или иных механизмов управления буферизацией и планирования очередей пакетов. PHB определяются в терминах характеристик поведения, относящихся к правилам обслуживания, а не в терминах конкретных механизмов реализации. В общем случае для реализации конкретной группы PHB может подходить широкий спектр механизмов реализации. Более того, очевидно, что на узле может быть реализовано более одной группы PHB, используемой в домене. Группы PHB следует определять так, чтобы было возможно подходящее распределение ресурсов между группами; могут реализоваться встроенные механизмы одновременной поддержки двух и более групп. Определение группы PHB должно показывать возможные конфликты с документированными ранее группами PHB, которые могут помешать одновременной работе.

Как описано в документе [DSFIELD], PHB выбирается на узле путём отображения кода DS в принятом пакете. Стандартизованные PHB имеют рекомендованные значения кода. Однако общее пространство кодов больше, нежели пространство, доступное для рекомендованных значений стандартизованных PHB и [DSFIELD] оставляет возможность для локально настраиваемых отображений. Таблица отображений код->PHB может содержать отображения как 1->1, так и N->1. Все коды должны быть отображены на тот или иной PHB; в отсутствии какой-либо локальной политики коды, которые не отображаются на стандартизованные PHB, в соответствии со спецификацией PHB следует отображать на Default PHB.

## 2.5 Распределение сетевых ресурсов

При реализации, настройке, работе и администрировании поддерживаемых групп PHB на узлах домена DS следует эффективно разделять ресурсы этих узлов и межузловых каналов между агрегатами поведения в соответствии с политикой обслуживания в этом домене. Кондиционеры трафика могут дополнительно контролировать использование этих ресурсов с помощью исполнения TCA и, возможно, с использованием обратной связи с узлами и кондиционерами трафика в домене. Хотя широкий спектр функций может быть развернут без использования сложных средств кондиционирования (например, на основе только статических правил маркировки), такие функции, как исполнение правил, формовка и динамическая перемаркировка определяют поддержку услуг, обеспечивающих количественные метрики производительности.

Настройка взаимодействия между кондиционерами трафика и внутренними узлами должна осуществляться под административным контролем и может требовать операционного контроля с помощью протоколов и объектов управления. Существует множество возможных моделей такого контроля.

Точная природа и реализация взаимодействия находятся за пределами настоящей архитектуры. Однако масштабируемость требует, чтобы управление доменом могло осуществляться без управления сетевыми ресурсами на микроуровне. Наиболее масштабируемая модель управления будет работать с узлами в режиме «открытой петли» и будет требовать административного управления только в случаях изменения SLA. Эта простая модель может может оказаться в некоторых обстоятельствах неподходящей и могут оказаться желательными некоторые автоматизированные, но медленно изменяющиеся (минуты, а не секунды) средства рабочего контроля для балансирования загрузки сетевых ресурсов в соответствии с недавно загруженным профилем.

# *3. Рекомендации по заданию поведения на этапе*

Базовые требования по спецификации поведения на этапах приведены в [DSFIELD]. В этом разделе требования детализируются путём описания дополнительных рекомендаций относительно спецификаций для PHB и групп. Это сделано, прежде всего, для обеспечения согласованности реализаций. Прежде, чем группа PHB будет предложена для стандартизации, следует выполнить эти рекомендации подобающим образом, что обеспечит целостность архитектуры.

**G.1**: Стандарт PHB должен задавать рекомендуемое значение кода DS, выбранное из пространства, зарезервированного для стандартных отображений [DSFIELD]. Рекомендуемые коды распределяются IANA. Предложение PHB может рекомендовать временный код из пространства EXP/LU для использования в междоменных экспериментах. Определение PHB для пакетов не должно требовать использования полей заголовка, отличных от DS.

**G.2**: В спецификацию каждой предлагаемой новой группы PHB следует включать обзор поведения и цель предлагаемого поведения. В обзор следует включать описание проблем(ы), для решения которых предназначена данная группа PHB, а также базовые концепции, связанные с группой PHB. В эти концепции следует включать по крайней мере описание поведения очередей, отбрасывания пакетов, выбора выходного канала. В конце обзора следует включить спецификацию метода, посредством которого группа PHB решает проблемы, указанные в описании.

**G.3**: В спецификации группы PHB следует указывать число отдельных PHB в группе. При задании множества PHB следует чётко описать взаимодействие между всеми PHB в группе и ограничения при таком взаимодействии. В качестве примера укажем, что спецификация должна рассматривать возможность повышения вероятности изменения порядка следования пакетов в микропотоке в результате маркировки этих пакетов для разных PHB в группе.

**G.4**: Когда корректное функционирование группы PHB зависит от ограничений (таких, как ограничения в обеспечении), в определение PHB следует включать описание поведения в случае нарушения таких ограничений. Более того, если нужны такие действия, как отбрасывание пакетов или смена порядка, эти действия следует оговаривать явно.

**G.5**: Группа PHB может быть специфицирована для локального использования в домене с целью обеспечения специфической для данного домена функциональности или услуг. В этом случае спецификация PHB полезна для

предоставления производителям согласованного определения группы PHB. Однако любые группы PHB, определённые для локального применения, не следует стандартизовать, что не препятствует их публикации в виде информационных RFC. Напротив, группы PHB, предназначенные для общего пользования, должны строго следовать процессу стандартизации. Следовательно, все предложения PHB должны содержать сведения о локальном или глобальном использовании группы.

Понятно, что группы PHB могут разрабатываться для обеспечения услуг между хостами, краевыми узлами сетей WAN[1] и/или краевыми узлами доменов. Термин «сквозной» (end-to-end) в определении PHB следует трактовать как «между хостами» (host-to-host).

Группы PHB могут определяться и развёртываться локально внутри доменов для экспериментов или в рабочем режиме. К таким группам не предъявляется требование публикации документов, но для этих групп PHB следует использовать коды DS из пулов EXP/LU, определённых в [DSFIELD].

**G.6**: Возможны случаи, когда пакеты, помеченные для PHB в группе, могут быть перемаркированы для выбора другого PHB в той же группе, Такая перемаркировка может осуществляться внутри домена или для пакетов, проходящих через границу домена. Обычно замена PHB обусловлена одной из трёх причин:

a) Коды, связанные с группой PHB, предназначены для передачи информации о состоянии сети.

b) Существуют ситуации, когда от PHB требуется повышение или снижение уровня приоритета для пакета (это предполагает некое ранжирование PHB внутри группы).

c) Граница между доменами не покрывается SLA. К этом случае код/PHB для выбора при прохождении через пограничный канал определяется локальной политикой восходящего домена.

В спецификации PHB следует чётко указывать обстоятельства, при которых пакеты, помеченные для PHB в группе, можно или следует изменить (например, повысить или снизить приоритет) для отнесения к другому PHB в той же группе. Если изменение PHB для пакета нежелательно, в спецификации следует чётко описать риски, связанные со сменой PHB. Возможным риском, связанным с заменой для пакета PHB (внутри группы или на другую группу PHB), является изменение порядка следования пакетов в микропотоке. PHB в группе могут переносить некоторую семантику взаимодействия между хостами, краевыми узлами WAN или доменов и дублирование этой семантики при перемаркировке пакетов для выбора иного PHB может оказаться затруднительным.

Для некоторых групп PHB может оказаться желательным отражение смены состояния путём перемаркировки пакетов для задания другого PHB в той же группе. Если группа PHB разработана для отражения состояний сети, определение PHB должно подобающим образом описывать связи между PHB и состояниями, которые они отражают. Более того, если эти PHB ограничены операциями по пересылке, которые может выполнять узел, эти ограничения могут быть описаны, как действия, которые узлу следует или необходимо выполнять.

**G.7**: В спецификацию группы PHB следует добавить раздел, определяющий включение туннелирования в свойства группы PHB. В этом разделе следует описать использование группы PHB из внешнего заголовка, когда исходное поле DS из внутреннего заголовка инкапсулировано в туннель. Также в этом разделе следует рассмотреть возможные изменения, которые могут выполняться по отношению к внутреннему заголовку на выходе туннеля, когда становятся доступными коды как из внешнего, так и из внутреннего заголовка (см. параграф 6.2).

**G.8**: Очевидно, что процесс спецификации групп PHB является инкрементным по своей природе. Когда предлагается новая группа PHB, следует документировать её понятные взаимодействия с ранее определёнными группами PHB. Создаваемая группа PHB может быть совсем новой или являться расширением какой-либо из существующих групп PHB. Если группа PHB полностью независима от всех или некоторых существующих спецификаций PHB, в спецификацию такой группы PHB следует включать раздел, описывающий сосуществование новой группы PHB с ранее стандартизованными группами PHB. Например, такой раздел может указывать возможность изменения порядка следования пакетов в микропотоке для пакетов, промаркированных кодами, которые связаны с двумя разными группами PHB. Если одновременная работа двух (или более) разных групп PHB на одном узле невозможна или может причинять вред, этот факт следует указывать в спецификации. Если одновременная работа двух (или более) разных групп PHB требует некого специфического режима, когда пакеты, маркированные для PHB из этих разных групп будут одновременно обрабатываться узлом, такое поведение должно быть отражено в спецификации.

Следует принимать меры против возникновения цикличности в определениях групп PHB.

Если предлагаемая группа PHB является расширением существующей группы, в спецификацию группы PHB следует включать раздел, описывающей взаимодействие этого расширения с расширяемым режимом. Если расширение изменяет или сужает определение режима тем или иным способом, этот факт следует указывать явно.

**G.9**: В каждую спецификацию PHB следует включать раздел, описывающий минимальные требования к реализациям, соответствующим данной спецификации. Такой раздел предназначен для того, чтобы разработчики могли понимать, какие функции должны быть реализованы и какие могут служить расширением, дозволенным спецификацией. Этот раздел может быть представлен в виде правил, таблиц, псевдокода или тестов.

**G.10**: В спецификацию PHB следует включать раздел, рассматривающий вопросы защиты режима. В этот раздел следует включать обсуждение перемаркировки кодов во внутреннем заголовке на выходе туннеля и воздействие этой перемаркировки на желаемый режим пересылки.

В этот раздел также следует включать обсуждение вопроса возможности использования предлагаемой группы PHB для атак на службы, атак на нарушение сервисных контрактов и снижение уровня обслуживания. Наконец, в этом разделе следует обсудить возможные методы детектирования атак, имеющие отношение к предлагаемому режиму.

**G.11**: В спецификацию PHB следует включать раздел, описывающий вопросы настройки и управления, способные влиять на работу PHB и служб, которые могут использоваться PHB.

**G.12**: Настоятельно рекомендуется включать в спецификацию приложение, описывающее влияние предлагаемого поведения на существующий и возможный сервис, включая (но не ограничиваясь) услуги, связанные с

[1]Wide Area Network - распределенная (глобальная) сеть. *Прим. перев.*

пользователями, устройствами, доменами или сквозной сервис. Настоятельно рекомендуется также включать в приложение раздел, описывающий верификацию услуг со стороны пользователей, устройств и/или доменов.

**G.13**: В каждую спецификацию PHB рекомендуется включать приложение, содержащее руководство по выбору PHB для пакетов, которые пересылаются в домены, не поддерживающие данную группу PHB.

**G.14**: В каждую спецификацию PHB рекомендуется включать приложение, которое рассматривает влияние предлагаемой группы PHB на существующие протоколы вышележащих уровней. При некоторых обстоятельствах PHB может разрешать изменения в протоколах вышележащих уровней, способные повышать или снижать полезность предлагаемой группы PHB.

**G.15**: В каждую спецификацию PHB рекомендуется включать приложение, которое рекомендует отображения на механизмы QoS канального уровня для поддержки желаемого режима PHB в сетях с разделяемой и коммутируемой на канальном уровне средой. Выбор наиболее подходящего отображения между PHB и механизмами QoS канального уровня зависит от множества факторов, которые выходят за пределы данного документа, однако в спецификации следует привести некоторые рекомендации по выбору такого отображения.

## *4. Взаимодействие с узлами, не поддерживающими DS*

Определим узел, не поддерживающий дифференциацию услуг (non-DS-compliant), как любой узел, который не интерпретирует поле DS в соответствии с [DSFIELD] и/или не поддерживает некоторые или все стандартизованные PHB (или PHB используемые в определённом домене DS). Такое поведение может быть результатом настройки узла или отсутствия поддержки соответствующих функций. Определим унаследованный узел, как специальный случай не поддерживающего DS узла, который реализует функции классификации и пересылки IPv4 Precedence в соответствии с [RFC791, RFC1812], но не соответствует DS. Значения предпочтений в октете IPv4 TOS осознанно сделаны совместимыми со значениями кодов селекторов класса, определёнными в [DSFIELD], а режим рассылки с использованием предпочтений, определённый в [RFC791, RFC1812], совместим с требованиями Class Selector PHB, определёнными в [DSFIELD]. Ключевым различием между унаследованным узлом и узлом, поддерживающим DS, является то, что унаследованный узел может интерпретировать или не интерпретировать биты 3-6 октета TOS (биты DTRC), как определено в [RFC1349]; на практике такой узел не будет интерпретировать эти биты в соответствии с [DSFIELD]. Мы предполагаем, что использование маркировки TOS, определённой в [RFC1349], в настоящее время не принято. Узлы, не совместимые с DS и не относящиеся к категории унаследованных, могут вести себя непредсказуемо по отношению к пакетам с отличными от нуля кодами DS[1].

Дифференциация обслуживания зависит от механизмов распределения ресурсов, обеспечиваемых реализациями PHB на узлах. Параметры качества обслуживания или гарантии уровня сервиса могут нарушаться при прохождении трафика через узлы или домены, не поддерживающие DS.

Рассмотрим два случая. Первый относится к использованию не поддерживающих DS узлов в доменах DS. Отметим, что пересылка с использованием PHB полезна, прежде всего, для контролируемого распределения дефицитных ресурсов узла или каналов. На высокоскоростных, слабо загруженных каналах максимальная задержка, вариации задержки и уровень потери пакетов пренебрежимо малы и использование не поддерживающего DS узла на восходящем конце такого канала может и не приводить к деградации сервиса. В более реалистичных условиях отсутствие пересылки с использованием PHB на узле может приводить к невозможности обеспечения малых задержек, низкого уровня потерь или требуемой полосы для проходящих через узел путей. Однако использование унаследованного узла может быть приемлемым вариантом, если домен DS ограничивается использованием только кодов селекторов класса, определённых в [DSFIELD], и предполагается, что конкретная реализация пересылки по предпочтениям на унаследованном узле обеспечивает режим пересылки, совместимый с услугами, предлагаемыми для проходящего через этот узел пути. Отметим, что важно ограничить использование кодов исключительно значениями Class Selector, поскольку унаследованный узел не обязан интерпретировать биты 3-5 в соответствии с [RFC1349], а это может вести к непредсказуемому поведению.

Второй случай относится к поведению сервиса, проходящего через домен, который не поддерживает DS. Для простоты аргументации предполагается, что не поддерживающий DS домен, не реализует функций кондиционирования трафика на граничных узлах. Следовательно, даже при условии присутствия внутри домена унаследованных узлов или узлов, поддерживающих DS, отсутствие правил, применяемых к трафику на границе домена, будет ограничивать возможности предоставления некоторых типов обслуживания через такой домен. Домен DS и домен, не поддерживающий DS, могут согласовать условия маркировки трафика, выходящего из домена DS, до входа в домен, который не поддерживает DS. Мониторинг выполнения заключённого соглашения может осуществляться путём выборки трафика взамен жёсткого его кондиционирования. Когда известно о том, что не поддерживающий DS домен состоит из унаследованных узлов, восходящий домен DS может перемаркировать трафик дифференцированного обслуживания с использованием одного или нескольких кодов Class Selector. Когда информации о возможностях управления трафиком в нисходящем домене нет и отсутствует соглашение, выходной узел домена DS может принять решение о перемаркировке кодов DS с использованием нулевого значения в предположении, что не поддерживающий DS домен будет пытаться обеспечить однородное обслуживание пакетов с использованием доступных возможностей (best-effort).

Если не поддерживающий DS домен является партнёром домена DS, трафик из домена, не поддерживающего DS, следует кондиционировать на входном узле домена DS в соответствии с подходящим SLA или правилами.

## *5. Вопросы групповой адресации*

Использование дифференцированного обслуживания для трафика с групповой адресацией порождает множество вопросов. Во-первых, пакеты с групповой адресацией, которые входят в домен DS, могут на входном узле распределяться по множеству путей через некоторые сегменты домена в результате репликации пакетов с групповыми адресами. В таких случаях может потребляться больше сетевых ресурсов, нежели при обработке индивидуальных пакетов. Когда группы multicast являются динамическими, сложно заранее предсказать объем потребных сетевых ресурсов для обслуживания группового трафика из восходящей сети в ту или иную группу. В результате такой неопределённости может осложняться обеспечение гарантий обслуживания для отправителей группового трафика.

[1] См. RFC 3260. *Прим. перев.*

Более того, может потребоваться резервирование кодов и PHB для исключительного использования с групповым трафиком, чтобы обеспечить изоляцию ресурсов от группового трафика.

Второй проблемой является выбор кода DS для групповых пакетов, прибывающих на входной узел DS. Поскольку такие пакеты могут выходить из домена DS через множество выходных узлов DS, являющихся партнёрами разных нисходящих доменов, выбор используемого кода DS не следует осуществлять на основании запросов от нижестоящего домена DS, нарушающих партнерские SLA. При организации состояния классификации и кондиционирования на входном узле DS для принимаемого агрегата трафика с дифференцированным обслуживанием, который проходит через выходную границу домена, идентификация смежного нисходящего транзитного домена и специфика соответствующего партнерского SLA должны приниматься во внимание при выборе конфигурации (вопрос политики маршрутизации и стабильности инфраструктуры маршрутизации). В этом случае партнерские SLA с нисходящими доменами DS могут быть частично исполнены на входной границе восходящего домена, что позволит снизить нагрузку, связанную с классификацией и кондиционированием трафика на выходном узле восходящего домена. Это не совсем просто для группового трафика по причине динамической принадлежности к группам. Результатом может явиться воздействие на гарантии сервиса для индивидуального трафика. Одним из способов решения этой проблемы является заключение отдельного партнерского SLA для группового трафика и использование для групповых пакетов отдельного набора кодов или реализация требуемой классификации и кондиционирования трафика на выходных узлах DS для обеспечения предпочтительного выделения ресурсов индивидуальному трафику в соответствии с партнерским SLA для нисходящего домена.

# *6. Вопросы безопасности и туннелирования*

В этом разделе рассматриваются вопросы безопасности, возникающие в связи с введением дифференцированного обслуживания, прежде всего, потенциальные атаки на службы (DoS) и потенциальные возможности несанкционированного обслуживания трафика (параграф 6.1). В дополнение к этому рассматривается дифференцированное обслуживание в присутствии IPsec (параграф 6.2) и требования к аудиту (параграф 6.3). Рассматриваются вопросы, связанные с использованием туннелей (как IPsec, так и иных).

## 6.1 Несанкционированное обслуживание и атаки на службы

Основной задачей дифференциации обслуживания является обеспечение возможности предоставления потокам трафика разного уровня сервиса в одной сетевой инфраструктуре. Для достижения результата могут использоваться различные методы управления ресурсами, но конечный результат должен заключаться в том, что обслуживание одних пакетов будет отличаться от обслуживания других (например, будет лучше). Отображение сетевого трафика на соответствующий режим, приводящий к иному (лучше или хуже) обслуживанию, задаётся, прежде всего, значением поля DS и, следовательно, злоумышленник может добиться улучшения обслуживания путём изменения поля DS для отображения трафика на режим, используемый для обеспечения лучшего обслуживания, или вставки пакетов с соответствующим значением поля DS. С учётом ограниченности ресурсов такой обман может служить для организации атак на службы, когда изменение или вставка пакетов ведёт к исчерпанию ресурсов, доступных для пересылки других потоков трафика. Защита от таких «краж» и DoS-атак включает средства кондиционирования трафика на граничных узлах DS в сочетании с обеспечением защиты и целостности сетевой инфраструктуры домена DS.

Как описано в разделе 2, входные узлы DS должны кондиционировать весь трафик, входящий в домен DS, для обеспечения приемлемости имеющихся в пакетах кодов DS. Это означает, что коды должны соответствовать применимым TCA и принятой в домене политике обслуживания. Следовательно, входные узлы являются первой линией защиты от атак на службы, основанных на изменении кодов DS, поскольку успех такой атаки ведёт к нарушению соответствующих TCA и принятой в домене политики обслуживания. Важно понимать, что любой, генерирующий трафик, узел домена DS является входным узлом для этого трафика и узел должен гарантировать допустимость всех кодов DS в порождённом этим узлом трафике.

Как политика обслуживания в домене, так и TCA могут требовать смены на входных узлах кода DS для некоторых пакетов (например, входной маршрутизатор может устанавливать для пользовательского трафика код DS в соответствии с подходящим SLA). Входные узлы должны кондиционировать весь остальной входящий трафик для обеспечения приемлемости кодов DS; пакеты с неприемлемыми кодами должны отбрасываться или значение кода DS должно меняться на приемлемое до пересылки пакета. Например, входной узел, получая трафик от домена, для которого нет соглашения об улучшенном обслуживании, может сбрасывать код DS в значение, принятое для Default PHB [DSFIELD][1]. Для разрешения использования некоторых кодов DS (например, соответствующих улучшенному обслуживанию) может потребоваться идентификация трафика, которая может быть выполнена техническими (например, IPsec) и/или иными (например, подключение входного канала к единственному заказчику) средствами.

Междоменное соглашение может снижать или сводить к нулю требования по кондиционированию входящего трафика, частично или полностью перенеся ответственность за приемлемые значения кодов DS во входящем трафике на нисходящий домен. В таких случаях входной узел домена может сохранять (избыточную) проверку кондиционирования для снижения зависимости от восходящего домена (например, такая проверка может предотвращать распространение DoS-атак, связанных с дифференцированным обслуживанием через границу домена). Если при такой проверке наблюдается нарушение условий со стороны восходящего домена, результат проверки следует записывать в системный журнал с указанием даты и времени приёма пакета, IP-адресов отправителя и получателя, а также кода DS. На практике следует учитывать ограниченность такого контроля и затрату ресурсов на его реализацию.

Внутренние узлы домена DS могут опираться на значение поля DS для связывания трафика с дифференцированным обслуживанием и режимов, используемых для реализации улучшенного сервиса. Любой узел, выполняющий такое связывание, зависит от корректности работы домена DS по предотвращению приёма трафика с неприемлемыми кодами DS. В соответствии с требованиями жизнеспособности получение пакетов с недопустимыми кодами DS не должно вызывать отказов (сбоев в работе) на узлах сети. Внутренние узлы не несут ответственности за исполнение политики обслуживания (или отдельных SLA) и, следовательно, от них не требуется проверять коды DS перед их использованием. Внутренние узлы могут выполнять некоторые проверки кондиционирования трафика по кодам DS (например, проверку кодов DS, которые никогда не используются на заданном канале) для повышения уровня безопасности и отказоустойчивости (например, устойчивость к атакам, связанным с изменением кодов DS, - краже

[1]См. обсуждение в разделе 6 RFC 3260. *Прим. перев.*

услуг). При отрицательном результате проверки следует генерировать запись в системном журнале в указанием даты и времени, IP-адресов отправителя и получателя, а также кода DS. На практике следует учитывать ограниченность такого контроля и затрату ресурсов на его реализацию.

Любой канал, для которого невозможно обеспечить адекватную защиту от изменения кодов DS или внедрения несанкционированного трафика, следует трактовать, как граничный канал (и, следовательно, весь входящий через этот канал трафик должен обрабатываться как на входном узле домена). Определение «адекватной защиты» задаётся локальной политикой безопасности и может включать определение рисков и последствий, связанных с изменением кодов DS, которые не перекрываются дополнительными мерами по обеспечению безопасности для данного канала. Уровень защиты канала можно дополнительно повысить за счёт контроля доступа на физическом уровне и/или программными средствами (типа организации туннелей), обеспечивающими целостность пакетов.

## 6.2 Взаимодействие с IPsec и туннелями

Протокол IPsec, как указано в [ESP, AH], не включает операций с полем DS заголовков IP в криптографические преобразования (в туннельном режиме поле DS внешнего заголовка IP не шифруется). Следовательно, изменение поля DS в сети не оказывает влияния на сквозную защиту IPsec, поскольку такое изменение не способно дать отрицательный результат при проверке целостности IPsec. В результате IPsec не обеспечивает какой-либо защиты от злонамеренного изменения поля DS (т. е., перехвата и изменения с участием человека - MITM), поскольку такое изменение не оказывает влияния на уровень сквозной защиты IPsec. В некоторых средах возможность изменения поля DS без влияния на целостность данных IPsec может позволить создание скрытого канала; если требуется предотвратить создание таких каналов или снизить их полосу, домены DS следует настраивать так, чтобы требуемая обработка (например, установка одного значения для всех полей DS в чувствительном трафике) могла выполняться на выходных узлах DS, где трафик выходит из защищённых доменов.

Туннельный режим IPsec обеспечивает защиту для полей DS инкапсулированных заголовков IP. Пакет IPsec в туннельном режиме включает два заголовка IP - внешний заголовок, подставляемый на входе туннеля, и внутренний (инкапсулированный) заголовок от инициатора пакета. Когда туннель IPsec проходит (полностью или частично) через сети с дифференциацией услуг, промежуточные узлы оперируют с полем DS внешнего заголовка. На выходе из туннеля IPsec внешние заголовки удаляются и пакет пересылается (если нужно) с использованием внутреннего заголовка. Если внутренний заголовок IP не обрабатывается входным узлом DS при выходе из туннеля в домен DS, выходной узел туннеля является входным узлом DS для выходящего из туннеля трафика и, следовательно, для него должно осуществляться соответствующее кондиционирование (см. параграф 6.1). Если обработка IPsec включает достаточно строгую криптографическую проверку целостности инкапсулированного пакета (достаточность определяется локальной политикой безопасности), выходной узел туннеля может без опасений предполагать, что поле DS во внутреннем заголовке не изменилось по сравнению со значением на входе в туннель. Это позволяет выходному узлу туннеля, находящемуся в одном домене DS со входным узлом туннеля, без опасений относиться к пакетам, прошедшим такую проверку целостности, как к пакетам, полученным от узла в том же домене DS, избавляя входной узел домена DS от необходимости кондиционирования трафика, которое требуется в противном случае. Важным следствием этого является то, что незащищённые каналы, являющиеся внутренними по отношению к домену DS, могут быть защищены с помощью достаточно надёжного туннеля IPsec.

Этот анализ и сделанные выводы применимы ко всем протоколам туннелирования, которые обеспечивают контроль целостности, но уровень защищенности поля DS во внутреннем заголовке зависит от строгости контроля целостности, обеспечиваемого протоколом туннелирования. При отсутствии достаточных гарантий для туннеля, транзитные узлы которого могут находиться за пределами домена DS (иначе говоря, уязвимы), инкапсулированные пакеты должны трактоваться, как на входном узле DS при доставке пакетов из-за пределов домена.

Протокол IPsec в настоящее время требует, чтобы поле DS внутреннего заголовка не изменялось при декапсуляции IPsec на выходном узле туннеля. Это предотвращает возможность использования измененных значений поля DS внешнего заголовка для организации утечки или атаки на службы через оконечную точку туннеля IPsec, поскольку все изменения полей внешнего заголовка отбрасываются на выходе из туннеля. Данный документ не меняет указанного требования IPsec.

Если спецификация IPsec в будущем разрешит выходным узлам туннелей менять поле DS во внутреннем заголовке IP в соответствии со значением поля DS во внешнем заголовке (т. е., полностью или частично копировать внешнее поле DS в одноимённое поле внутреннего заголовка), потребуется дополнительное рассмотрение этого вопроса. Для туннелей, расположенных полностью в одном домене DS и обеспечивающих адекватную защиту от изменения внешнего поля DS, ограничивать изменение внутреннего поля DS будет только политика обслуживания в домене. В противном случае выходной узел домена, выполняющий такие изменения, будет действовать, как входной узел DS для выходящего из туннеля трафика и должен передавать входному узлу ответственность за кондиционирование трафика, включая защиту от утечки и атак на службы (см. параграф 6.1). Если туннель входит в домен DS на узле, который не совпадает с выходным узлом туннеля, тогда выходной узел туннеля может зависеть от того, насколько входной узел DS восходящего направления обеспечивает приемлемость внешнего поля DS. Даже в этом случае существуют способы проверки, которые доступны выходному узлу туннеля (например, проверка согласованности внутреннего и внешнего кода DS для шифрованного туннеля). Любые негативные результаты таких проверок должны фиксироваться системой аудита с генерацией записи в журнал аудита, включающей дату и время получения пакета, адреса отправителя и получателя, а также недопустимое значение кода DS.

Туннель IPsec с точки зрения архитектуры можно представлять по крайней мере в двух видах. Если туннель рассматривается как виртуальный провод с одним интервалом пересылки, действия промежуточных узлов по пересылке туннелируемого трафика не следует делать видимыми за пределами оконечных узлов туннеля и, следовательно, поле DS не следует менять в процессе декапсуляции. И напротив, при рассмотрении туннеля, как многоэтапной системы пересылки трафика, изменение поля DS при декапсуляции может оказаться желательным. Примером второго варианта является ситуация, когда туннель завершается на внутреннем узле домена DS, для которого администратор не хочет использовать логику кондиционирования трафика (т. е., желает упростить управление трафиком). Это может быть реализовано путём использования кода DS из внешнего заголовка IP, который устанавливается при кондиционировании трафика на входном узле туннеля, в качестве значения кода DS внутреннего заголовка IP, что позволяет перенести ответственность за кондиционирование трафика с выходного узла туннеля IPsec на соответствующий входной узел DS (который должен выполнять эту функцию для трафика до инкапсуляции).

## 6.3 Аудит

Не все системы, поддерживающие дифференцированное обслуживание, будут реализовать аудит. Однако, если поддержка дифференцированного обслуживания встроена в систему, поддерживающую аудит, реализации дифференцированного обслуживания также следует поддерживать аудит. Если такая поддержка присутствует, реализация должна позволять администратору системы включать или отключать аудит для дифференцированного обслуживания в целом и может поддерживать такое управление на уровне отдельных частей.

В большинстве случаев гранулярность аудита определяется локальными требованиями. Однако некоторые события упомянуты в этом документе, как протоколируемые и для каждого из таких событий определён минимальный набор информации, которую следует помещать в журнал аудита. Допускается также включение в журнал дополнительной информации (например, сведений о пакетах, связанных с теми, которые вызвали запись в журнал аудита). Кроме того, записи в журнал аудита могут быть связаны с другими событиями, не упомянутыми явно в этом документе. Не задаётся требований, заставляющих получателя передавать какие-либо сообщения соответствующему отправителю при обнаружении заносимого в журнал аудита события, поскольку введение такого требования создавало бы потенциальную угрозу атак на отказ служб (DoS).

# *7. Благодарности*

Этот документ основан на ранних работах Steven Blake, David Clark, Ed Ellesson, Paul Ferguson, Juha Heinanen, Van Jacobson, Kalevi Kilkki, Kathleen Nichols, Walter Weiss, John Wroclawski и Lixia Zhang.

Авторы выражают свою признательность за полезные комментарии и предложения Kathleen Nichols, Brian Carpenter, Konstantinos Dovrolis, Shivkumar Kalyana, Wu-chang Feng, Marty Borden, Yoram Bernet, Ronald Bonica, James Binder, Borje Ohlman, Alessio Casati, Scott Brim, Curtis Villamizar, Hamid Ould-Brahi, Andrew Smith, John Renwick, Werner Almesberger, Alan O'Neill, James Fu и Bob Braden.

# *8. Литература*

[802.1p] ISO/IEC Final CD 15802-3 Information technology - Tele-communications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) bridges, (current draft available as IEEE P802.1D/D15)[1].

[AH] Kent, S. and R. Atkinson, "IP Authentication Header", RFC 2402, November 1998.

[ATM] ATM Traffic Management Specification Version 4.0 <af-tm-0056.000>[2], ATM Forum, April 1996.

[Bernet] Y. Bernet, R. Yavatkar, P. Ford, F. Baker, L. Zhang, K. Nichols, and M. Speer, "A Framework for Use of RSVP with Diff-serv Networks", Work in Progress[3].

[DSFIELD] 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.

[EXPLICIT] D. Clark and W. Fang, "Explicit Allocation of Best Effort Packet Delivery Service", IEEE/ACM Trans. on Networking, vol. 6, no. 4, August 1998, pp. 362-373.

[ESP] Kent, S. and R. Atkinson, "IP Encapsulating Security Payload (ESP)", RFC 2406, November 1998.

[FRELAY] ANSI T1S1, "DSSI Core Aspects of Frame Rely", March 1990.

[RFC791] Postel, J., Editor, "Internet Protocol", STD 5, RFC 791, September 1981.

[RFC1349] Almquist, P., "Type of Service in the Internet Protocol Suite", RFC 1349, July 1992.

[RFC1633] Braden, R., Clark, D. and S. Shenker, "Integrated Services in the Internet Architecture: An Overview", RFC 1633, July 1994.

[RFC1812] Baker, F., Editor, "Requirements for IP Version 4 Routers", RFC 1812, June 1995.

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

[2BIT] K. Nichols, V. Jacobson, and L. Zhang, "A Two-bit Differentiated Services Architecture for the Internet", ftp://ftp.ee.lbl.gov/papers/dsarch.pdf, November 1997.

[TR] ISO/IEC 8802-5 Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 5: Token Ring Access Method and Physical Layer Specifications, (also ANSI/IEEE Std 802.5-1995), 1995.

# *Адреса авторов*

**Steven Blake**
Torrent Networking Technologies
3000 Aerial Center, Suite 140
Morrisville, NC 27560
Phone: +1-919-468-8466 x232
EMail: slblake@torrentnet.com

**David L. Black**
EMC Corporation
35 Parkwood Drive
Hopkinton, MA 01748

[1]В настоящее время это является частью стандарта IEEE 802.1D. *Прим. перев.*
[2]Документ доступен по ссылке http://www.ipmplsforum.org/ftp/pub/approved-specs/af-tm-0056.000.pdf. *Прим. перев.*
[3]Работа опубликована в RFC 2998. *Прим. перев.*

Phone: +1-508-435-1000 x76140
EMail: black_david@emc.com

**Mark A. Carlson**
Sun Microsystems, Inc.
2990 Center Green Court South
Boulder, CO 80301
Phone: +1-303-448-0048 x115
EMail: mark.carlson@sun.com

**Elwyn Davies**
Nortel UK
London Road
Harlow, Essex CM17 9NA, UK
Phone: +44-1279-405498
EMail: elwynd@nortel.co.uk

**Zheng Wang**
Bell Labs Lucent Technologies
101 Crawfords Corner Road
Holmdel, NJ 07733
EMail: zhwang@bell-labs.com

**Walter Weiss**
Lucent Technologies
300 Baker Avenue, Suite 100
Concord, MA 01742-2168
EMail: wweiss@lucent.com

**Перевод на русский язык**

Николай Малых

nmalykh@protokols.ru

## *Полное заявление авторских прав*