Network Working Group
Request for Comments: 2474
Obsoletes: 1455, 1349
Category: Standards Track

K. Nichols
Cisco Systems
S. Blake
Torrent Networking Technologies
F. Baker
Cisco Systems
D. Black
EMC Corporation

December 1998

Определение поля DS в заголовках IPv4 и IPv6

Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers

Статус документа

В этом документе содержится спецификация протокола, предложенного сообществу Internet. Документ служит приглашением к дискуссии в целях развития и совершенствования протокола. Текущее состояние стандартизации протокола вы можете узнать из документа "Internet Official Protocol Standards" (STD 1). Документ может распространяться без ограничений.

Авторские права

Copyright (C) The Internet Society (1998). All Rights Reserved.

Тезисы

Дифференцированное обслуживание добавляется в протокол IP для обеспечения масштабируемой дифференциации сервиса в Internet без необходимости поддержки для каждого потока состояния и сигнализации на каждом интервале маршрутизации. Различные уровни обслуживания могут создаваться с использованием небольшого и хорошо определенного набора элементов, которые разворачиваются на узлах сети. Уровни обслуживания могут быть сквозными или внутридоменными, они могут включать как гарантированную производительность (т. е., пиковую полосу), так и относительную производительность (т. е., дифференциацию по «классам»). Услуги могут строиться с использованием комбинаций следующих действий:

- установка битов в полях заголовка IP на границах сетей (границы автономных систем, внутренние административные границы, хосты);
- использование этих битов для управления пересылкой пакетов внутренними узлами сети;
- кондиционирование маркированных пакетов на границах сети в соответствии с требованиями правил для каждого уровня обслуживания.

Требования или правила для каждого типа сервиса должны устанавливаться с помощью механизмов административного управления, выходящих за пределы данного документа. Сетевой узел, поддерживающий дифференцированный сервис, включает средства классификации, которые выбирают пакеты на основе значения поля DS, вместе с механизмами буферизации и управления очередями, способными пересылать конкретные пакеты с учетом значения поля DS. Установка значений поля DS и управление порядком обработки маркированных пакетов требуются только на границе сети; эти операции могут быть достаточно сложными.

В этом документе определено поле заголовка IP, которое называется DS[1]. Для IPv4 документ определяет схему октета TOS, для IPv6 — октета Traffic Class. В дополнение к этому определен базовый набор вариантов обработки пакетов при пересылке.

Более подробное описание дифференцированных услуг содержится в документе [ARCH], посвященном архитектуре Diffserv.

Оглавление

1. Введение....2
2. Используемая терминология....3
3. Определение поля DS....4
4. Требования к старым кодам и PHB....5
4.1 PHB по умолчанию....5

[1]Differentiated services - дифференцированное обслуживание.

4.2 Настоящее и будущее поля IP Precedence....5
4.2.1 Краткая история развития IP Precedence....5
4.2.2 Отображение IP Precedence в коды селекторов класса....6
4.2.2.1 Коды селектора класса....6
4.2.2.2 Требования к PHB для селектора класса....6
4.2.2.3 Использование требований к PHB для селектора класса и совместимость с IP Precedence....6
4.2.2.4 Пример механизма реализации групп, совместимых с требованиями к PHB для селектора класса..7
4.3 Выводы....7
5. Руководство по стандартизации поэтапного поведения....7
6. Согласование с IANA....8
7. Вопросы безопасности....8
7.1 Несанкционированное использование услуг и DoS-атаки....8
7.2 Взаимодействие с IPsec и туннелями....8
8. Благодарности....9
9. Литература....9
Адреса авторов....9
Полное заявление авторских прав....10

1. Введение

Дифференциация услуг предназначена для создания основы построения масштабируемых систем дифференцированного обслуживания пакетов в сети Internet. Для ускоренного достижения цели предлагаемая архитектура разделена на две части, одна из которых достаточно хорошо разработана и понятна, а другая требует дополнительных исследований. В этом документе мы следуем исходной архитектуре Internet, где раздельно принимаются решения для компонент пересылки и маршрутизации. Пересылка пакетов является сравнительно простой задачей, которую требуется выполнять для каждого пакета независимо и максимально быстро. Для пересылки используется заголовок пакета, позволяющий найти запись в таблице маршрутизации, которая определяет выходной интерфейс. Маршрутизация создает записи в этой таблице, для чего может потребоваться учет множества транзитных и других правил, а также хранение сведений об отказах на маршрутах. Таблицы маршрутизации поддерживаются как фоновый процесс для пересылки пакетов. Маршрутизация является более сложной задачей и эта сложность продолжает расти в течение последних 20 лет.

Аналогично, архитектура дифференцированного обслуживания содержит две основных компоненты. Одна связана с хорошо понятным поведением на пути пересылки, а вторая более сложна и связана с политикой и распределением компонент, которые задают параметры, используемые при пересылке. Поведение пути пересылки включает дифференцированную трактовку получаемых пакетов и реализацию дисциплины очередей и управления ими. Такое поведение на каждом этапе весьма полезно и требуется от сетевых узлов для обеспечения дифференцированной обработки пакетов, не зависящего от построения сквозных и внутридоменных служб. Здесь рассматривается скорее общая семантика поведения, нежели специфические механизмы, используемые для реализации этого поведения, поскольку поведение меняется не так быстро, как механизмы.

Поведение на каждом этапе и механизмы выбора поведения для отдельных пакетов могут быть развернуты в узлах сети уже сегодня и этот аспект архитектуры дифференцированного обслуживания будет реализован в первую очередь. В дополнение к этому, путь пересылки может требовать некоторого мониторинга, выполнения правил и формовки трафика, для которого задана специальная обработка, чтобы выполнить требования, связанные с доставкой таких пакетов. Механизмы кондиционирования трафика также хорошо понятны. Повсеместное развертывание систем кондиционирования трафика важно для обеспечения возможности построения услуг, хотя их реальное использование может начаться не сразу.

Конфигурация элементов сети, за счет использования которых пакетам обеспечивается специальная обработка, и правила, применимые к использованию ресурсов, изучены гораздо слабее. Тем не менее, возможно развертывание систем дифференцированного обслуживания в сетях на основе простых правил и статических конфигураций. Как указано в работе [ARCH], существует множество способов организации поведения на каждом этапе и кондиционирования трафика для дифференциации услуг. В процессе эксплуатации таких систем накапливается дополнительный опыт, позволяющий создавать более сложные правила и распределение ресурсов. Базовое поведение на пути пересылки при этом может сохраняться, но компоненты архитектуры будут развиваться. Опытное использование такого сервиса должен быть достаточно продолжительным и мы не предлагаем сразу стандартизовать решение - это было бы преждевременным. Более того, многие детали дифференциации услуг связаны с контрактными соглашениями между различными предприятиями и организациями, а этот вопрос выходит за пределы компетенции IETF.

В этом документе рассматривается, прежде всего, компонента пути пересылки пакетов. На этом пути дифференцированное обслуживание реализуется за счет отображения кода из заголовка пакета IP на определенный режим пересылки или поэтапное поведение (PHB[1]) для каждого узла сети на пути пересылки. Коды могут выбираться из набора стандартных значений, определенных ниже в этом документе, набора рекомендуемых значений, которые могут быть определены в будущем, или из набора, имеющего исключительно локальное значение. Предполагается, что PHB будет реализоваться за счет создания службы очередей и/или порядка управления очередями на выходном интерфейсе сетевого устройства. Примером может служить взвешенное циклическое обслуживание очередей (WRR[2]) или управление очередями на основе приоритетов отбрасывания пакетов[3].

[1]Per-hop behavior.

[2]Weighted round-robin.

[3]В оригинале - «drop-preference queue management». *Прим. перев.*

Маркирование выполняется кондиционерами трафика на границах сетей, включая краевые узлы (первый маршрутизатор или хост-источник) и административные границы. Кондиционеры трафика могут включать примитивы маркировки, измерения, реализации политики и формовки (эти механизмы описаны в [ARCH]). Услуги реализуются путем использования классификации отдельных пакетов и механизмов кондиционирования трафика на границе, а также конкатенации поэтапного поведения на пути транзита трафика. Целью архитектуры дифференцированного обслуживания является определение «строительных блоков» с учетом расширения в будущем как числа, так и типов строительных блоков, а также расширения спектра услуг, обеспечиваемых на основе этих блоков.

Используемая в документе терминология определена в разделе 2. Поле дифференцированного обслуживания (DS) рассматривается в разделе 3. Раздел 4 посвящен описанию частичной совместимости со сложившейся практикой использования поля Precedence в IPv4. В качестве решения предлагаются коды селекторов класса и совместимые с селектором класса PHB. В разделе 5 приводятся рекомендации по стандартизации поэтапного поведения. Раздел 6 содержит рекомендации по распределению кодов. В разделе 7 рассматриваются вопросы безопасности. Документ в целом представляет собой описание поля DS и его использования. Документ следует читать вместе с описанием архитектуры дифференцированного обслуживания [ARCH].

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 [RFC2119].

2. Используемая терминология

Behavior Aggregate - агрегат поведения

Набор пакетов с одинаковым кодом DS, проходящих через канал в определенном направлении. Термины «агрегат» и «агрегат поведения» далее будут трактоваться, как синонимы.

Classifier - классификатор

Объект, который выбирает пакет на основе содержимого заголовков в соответствии с заданными правилами.

Class Selector Codepoint — код селектора класса

Любой из восьми кодов xxx000 (x может принимать значение 0 или 1). Коды селекторов класса рассматриваются в параграфе 4.2.2.

Class Selector Compliant PHB — совместимое с селектором класса поэтапное поведение

Поэтапное поведение, удовлетваляющее требованиям к Class Selector PHB, приведенным в параграфе 4.2.2.2.

Codepoint - код

Конкретное значение компоненты DSCP поля DS. Рекомендованные коды **следует** отображать на конкретные, стандартизованные PHB. Множество кодов **может** отображаться на один PHB.

Differentiated Services Boundary — граница дифференцированных услуг

Край домена DS, где развертываются классификаторы и кондиционеры трафика. Граница дифференцированных услуг может, в свою очередь, делиться на входные и выходные узлы (входные/выходные узлы являются нисходящими/восходящими узлами для заданного направления трафика). Граница дифференцированных услуг обычно находится у входа к первому маршрутизатору (или узлу сети), поддерживающему дифференцированное обслуживание, через который проходят пакеты хоста, выхода последнего маршрутизатора (или узла сети), поддерживающего дифференцированное обслуживание, или узла сети, через который проходят пакеты до прибытия на хост. Иногда границу дирренцированного обслуживания называют границей на маршрутизаторе ответвления. Граница дифференцированного обслуживания может быть совмещена с хостом (в зависимости от локальной политики). Используется также термин «граница DS».

Differentiated Services-Compliant — совместимость с дифференцированным обслуживанием

Соответствие требованиям данного документа. Используется также термин «совместимость с DS».

Differentiated Services Domain — домен дифференцированных услуг

Непрерывное подмножество Internet, в котором обеспечивается согласованный набор правил дифференцированного обслуживания. Домен дифференцированных услуг может быть представлен разными административными доменами, автономными системами, областями доверия, сетевыми технологиями (например, коммутация кадров/ячеек), хостами, маршрутизаторами и т. п. Используется тайже термин DS-домен.

Differentiated Services Field — поле DS

Октет TOS в заголовке IPv4 или октет Traffic Class в заголовке IPv6, который интерпретируется в соответствии с определениями данного документа.

Mechanism — механизм

Реализация одного или множества вариантов поэтапного поведения в соответствии с определенным алгоритмом.

Microflow — микропоток

Один экземпляр потока пакетов между приложениями, идентифицируемый адресами отправителя и получателя, номером протокола, а также (когда это применимо) номерами портов отправителя и получателя.

Per-hop Behavior (PHB) — поэтапное поведение

Описание наблюдаемого извне режима пересылки, применяемого на поддерживающих дифференциацию услуг узлах к агрегату поведения. **Следует** достаточно детально описывать PHB, чтобы можно было создавать прогнозируемые услуги, как описано в [ARCH].

Per-hop Behavior Group — группа PHB

Набор из одного или множества PHB, которые имеют значение и реализуются только совместно по причине общих ограничений, применимых ко всем PHB в группе (например, обслуживание очередей или правила управления очередями).

Traffic Conditioning — кондиционирование трафика

Функции управления, которые могут быть применены к агрегату поведения, потоку приложений или другому подмножеству трафика (например, маршрутным обновлениям). Кондиционирование **может** включать измерение, исполнение правил, формовку и маркировку пакетов. Кондиционирование трафика используется для реализации соглашения между доменами и приведения трафика в соответствие с политикой дифференцированного обслуживания в домене путем маркировки пакетов соответствующим кодом DS и мониторинга временных параметров агрегата с генерацией сигналов, когда это требуется. См. [ARCH].

Traffic Conditioner — кондиционер трафика

Элемент, выполняющий функции кондиционирования, которые **могут** включать измерение, исполнение правил, формовку и маркировку. Кондиционеры трафика обычно размещаются на узлах границы DS (т. е., не на внутренних узлах домена DS).

Service — сервис, услуга

Описание режима обслуживания (подмножества) трафика заказчика в отдельном домене, множестве соединенных между собой доменов DS или сквозного на всем пути пакетов. Описание услуг представляется административными правилами и услуги конструируются путем применения кондиционирования трафика для создания агрегатов поведения, для которых будет использоваться известный вариант PHB на каждом узле внутри домена DS. Одним вариантом поэтапного поведения может поддерживаться множество типов услуг, используемых в соответствии с кондиционированием трафика.

Классификаторы и кондиционеры трафика используются для выбора пакетов, относящихся к агрегатам поведения. Агрегатам предоставляется дифференцированное обслуживание в домене DS и кондиционеры трафика **могут** менять временные характеристики агрегатов в соответствии с некими требованиями. Поле DS в заголовках пакетов используется для обозначения агрегата поведения и впоследствии служит для определения используемого по отношению к пакетам режима пересылки. Классификатор агрегатов поведения, который может выбирать PHB (например, дисциплину обслуживания очереди) на основе кода в поле DS, **следует** включать во все сетевые узлы домена DS. Классификаторы и кондиционеры трафика на границах DS настраиваются в соответствии с некой спецификацией обслуживания, относящейся к административной политике, которая выходит за пределы настоящего документа.

Дополнительные определения дифференцированных услуг даны в документе [ARCH].

3. Определение поля DS

Здесь дано определение поля DS, которое предназначено для замены существующих определений октета TOS в заголовках IPv4 [RFC791] и октета Traffic Class в заголовках IPv6 [IPv6][1].

Шесть битов поля DS используются в качестве кода DSCP для выбора варианта PHB, с которым будет сталкиваться пакет на каждом узле. Двухбитовое поле CU является резервным, его определение и использование выходит за пределы данного документа. Значение битов CU игнорируется узлами, поддерживающими дифференцированное обслуживание, при определении поэтапного поведения для полученного пакета.

Структура поля DS показана на рисунке справа.

DSCP[2] - код дифференцированного обслуживания;

CU[3] - не используется в настоящее время.

```
  0   1   2   3   4   5   6   7
+---+---+---+---+---+---+---+---+
|          DSCP           |  CU   |
+---+---+---+---+---+---+---+---+
```

При указании значений DSCP используется нотация «xxxxxx» (x может принимать значения 0 и 1), где в крайней левой позиции указывается бит 0 поля DS (как показано на рисунке), а в крайней правой — бит 5.

Разработчикам следует отметить для себя, что размер поля DSCP составляет 6 битов. Поддерживающие DS узлы **должны** выбирать PHB, проверяя соответствие всему 6-битовому поля DSCP (например, трактуя значение поля, как индекс таблицы, используемой для выбора режима обслуживания пакетов, реализованного в данном устройстве). Значения поля CU **должны** игнорироваться при выборе PHB. Поле DSCP трактуется, как бесструктурное, чтобы обеспечить в будущем возможность определения разных вариантов поэтапного поведения.

С некоторыми исключениями, отмеченными ниже, отображение кодов на PHB **должно** быть настраиваемым. Соответствующие DS узлы **должны** поддерживать логический эквивалент настраиваемой таблицы отображения кодов в PHB. Спецификации PHB **должны** включать рекомендуемый по умолчанию код, который **должен** быть уникальным в стандартном пространстве кодов (см. раздел 6). Реализациям следует поддерживать рекомендуемые отображения кодов в PHB в принятой по умолчанию конфигурации. Операторы могут выбирать использование различных кодов для PHB в дополнение к рекомендуемым по умолчанию или вместо них. Отметим, что при таком выборе может потребоваться перемаркировка полей DS на административных границах даже в тех случаях, когда по обе стороны границы реализованы одинаковые PHB.

Вопросы перемаркировки более подробно рассмотрены в [ARCH].

Исключения из общих правил относятся к кодам xxx000 и рассмотрены в параграфах 4.2.2 и 4.3.

---

[1]См. также RFC 3260, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

[2]Differentiated services codepoint.

[3]Currently unused.

Пакеты, полученные с неизвестным кодом, **следует** пересылать, как будто они помечены для принятого по умолчанию поведения (см. раздел 4), не меняя в них кода[1]. Для узлов сети **недопустимо** нарушение функционирования при получении таких пакетов.

Показанная выше структура поля DS несовместима с существующим определением октета IPv4 TOS [RFC791]. Предполагается, что домены DS защищают себя, реализуя граничные узлы перемаркировки, как следует делать сетям, использующим поле RFC 791 Precedence. Процедуры перемаркировки **следует** выполнять в соответствии с [RFC791], где сказано: «Если та или иная сеть использует значение уровня предпочтения, она берет на себя ответственность за доступ к этому полю и его использование[2]». Проверка значения поля DS на границах DS осмысленна в любом случае, поскольку узел восходящего направления может легко установить для этого поля любое значение. Домены DS, не отделенные подобающим образом настроенными узлами, могут предлагать непредсказуемый сервис.

Узлы **могут** переписывать поле DS в соответсвии с потребностями обеспечения требуемого обслуживания (локального или сквозного). Спецификации преобразования поля DS на границах DS являются предметом соглашений об обслуживании (SLA[3]) между провайдерами и пользователями и выходят за пределы данного документа. Стандартизованные PHB позволяют провайдерам строить свои услуги на основе общепринятых режимов пересылки пакетов, поддержка которых предполагается в большинстве типов оборудования разных производителей.

4. Требования к старым кодам и PHB

Как отмечено в этом разделе, поле DS обеспечивает лишь ограниченную совместимость с современной практикой. Вопрос совместимости с более ранними версиями решается двумя способами. Во-первых существуют варианты поэтапного поведения, которые уже получили достаточно широкое распространение (например, варианты, удовлетворяющие требованиям к управлению очередями на основе поля IPv4 Precedence, заданным в [RFC1812]), и мы хотим сохранить возможность их использования на поддерживающих DS узлах. Кроме того, существуют некоторые коды, соответствующие сложившемуся использованию поля IP Precedence, и эти коды резервируются для отображения на PHB, которые соответствуют заданным в параграфе 4.2.2.2 общим требованиям, хотя конкретные PHB дифференцированного обслуживания, на которые отображаются эти коды, **могут** иметь дополнительные спецификации.

Здесь не предпринимается попыток обеспечить совместимость с DTR или битами TOS октета IPv4 TOS, определенного в [RFC791][4].

4.1 PHB по умолчанию

Используемый по умолчанию вариант PHB **должен** быть доступен на поддерживающих DS узлах. Это обычное поведение с обеспечением пересылки по возможности[5], доступное в существующих маршрутизаторах и стандартизованное в [RFC1812]. При отсутствии других соглашений предполагается, что пакеты относятся к этому агрегату. Такие пакеты **могут** передаваться в сеть без выполнения каких-либо специальных правил и сеть будет доставлять эти пакеты в таком количестве и с таким качеством, как это возможно в соответствии с существующими правилами ограничения ресурсов. Подходящей реализацией такого PHB будет дисциплина очередей, обеспечивающая передачу пакетов этого агрегата, когда выходной канал не требуется для других PHB. Подходящим правилом для конструирования сервиса будет «поддержание жизни» для этого агрегата. Это может быть реализовано на каждом узле с помощью механизма, резервирующего некие минимальные ресурсы (например, буферы, полосу) для принятого по умолчанию агрегата поведения. Такой подход позволит отправителям, не поддерживающим дифференцированное обслуживание, использовать сети так же, как это происходило без дифференциации. Влияние введения дифференцированных услуг в домене на обслуживание пользователей и партнеров является достаточно сложным вопросом, включающим политику домена, и выходит за пределы данного документа. **Рекомендуемым** кодом для принятого по умолчанию PHB является 000000; значение кода 000000 **должно** отображаться на PHB, который соответствует данной спецификации. Выбранный для принятого по умолчанию поведения код совместим с существующей практикой [RFC791]. Если код не отображается на стандартизованный или локальный PHB, его **следует** отображать на принятый по умолчанию PHB.

Пакет, изначально помеченный для принятого по умолчанию поведения, **может** быть перемаркирован с использованием другого кода при прохождении границы в домен DS и будет пересылаться в этом домене с использованием другого PHB (возможно, в результате партнерского соглашения между соседними доменами).

4.2 Настоящее и будущее поля IP Precedence

Мы хотим создать некую форму совместимости с современным использованием поля IP Precedence - битами 0 - 2 октета IPv4 TOS. Существуют маршрутизаторы, использующие поле IP Precedence для выбора различных режимов поэтапной пересылки, подобно использованию предложенного здесь поля DSCP. Таким образом, простой прототип архитектуры дифференцированного обслуживания может быть развернут достаточно быстро путем соответствующей настройки таких маршрутизаторов. Более того, IP-системы сегодня понимают местоположение поля IP Precedence и, таким образом, при использовании этих битов по аналогии с поддерживающим DS оборудованием, на начальном этапе не должно возникать серьезных проблем. Иными словами, строгое соответствие DS не является обязательным даже в сети одного сервис-провайдера, если биты 0 - 2 поля DSCP используются по аналогии с битами поля IP Precedence.

4.2.1 Краткая история развития IP Precedence

Поле IP Precedence в той или иной степени является предтечей поля DS. Определения предпочтений IP и поля IP Precedence изначально были даны в [RFC791]. Значения, которые может принимать трехбитовое поле IP Precedence, могут использоваться для разных типов трафика, включая трафик сетевого управления, маршрутную информацию и

[1]См. обсуждение в разделе 6 RFC 3260, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

[2]Здесь и далее RFC цитируются по переводам, опубликованным на сайте www.protocols.ru. *Прим. перев.*

[3]Service level agreement – соглашение об уровне обслуживания. *Прим. перев.*

[4]См. раздел 7 RFC 3260, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

[5]В оригинале: «best-effort forwarding» - пересылка с использованием разумно доступных возможностей.

различные уровни привилегий. В [RFC791] толкование поля Precedence было определено достаточно широко, как «степень важности дейтаграммы». Предполагалось, что не все значения поля IP Precedence сохраняют смысл при пересечении границ. Например, в [RFC791] было отмечено: «Уровень предпочтения Network Control (управление сетью) означает, что дейтаграмма предназначена для использования внутри сети. Реальная трактовка этого обозначения определяется местными условиями сети».

Хотя в ранних BBN IMP было реализовано использование поля Precedence, ранние коммерческие маршрутизаторы и код пересылки IP в UNIX не поддерживали эту функцию. По мере роста сложности сетей и требований пользователей производители коммерческих маршрутизаторов разработали различные варианты управления очередями пакетов, включая приоритетную обработку, которая в общем случае основывалась на правилах, закодированных в маршрутизаторах и использующих в качестве критериев адреса IP, номера протоколов IP, номера портов TCP или UDP и другие поля заголовков. Поле IP Precedence также включалось в число параметров.

Говоря коротко, можно отметить, что поле IP Precedence используется достаточно широко, но не в полном соответствии с [RFC791]. Это отмечено в [RFC1122], где указано, что использование поля IP Precedence осуществляется корректно, а конкретное распределение привилегий в [RFC791] стало достоянием истории.

4.2.2 Отображение IP Precedence в коды селекторов класса

Спецификация выбора режима пересылки пакетов на основе значений поля IP Precedence в настоящее время весьма неконкретна и ее может быть недостаточно для построения предсказуемого сервиса в схеме дифференцированного обслуживания. Для сохранения некоторой совместимости с прежними версиями при известной сложившейся практике использования поля IP Precedence без потери гибкости в будущем предлагается модель, описывающая минимальные треобвания к набору PHB, которые совместимы с развернутыми схемами выбора режима пересылки на основе поля IP Precedence. В дополнение к этому дается набор кодов, которые **должны** отображаться на PHB, соответствующие этим минимальным требованиям. PHB, отображающиеся на эти коды, **могут** иметь более детальные спецификации в дополнение к требованиям настоящего документа. Другие коды **могут** отображаться на те же PHB. Будем называть этот набор кодами селекторов класса[1], а минимальные требования к PHB, на которые эти коды могут отображаться, - требованиями к PHB для селектора класса[2].

4.2.2.1 Коды селектора класса

Спецификация режимов пересылки пакетов на основе значений поля DS xxx000|xx или DSCP = xxx000 без указания поля CU резервируется как набор кодов селекторов класса. PHB, которые отображаются на эти коды, **должны** соответствовать требованиям к PHB для селекторов класса, в дополнение к требованиям для Default PHB с кодом 000000 (параграф 4.1).

4.2.2.2 Требования к PHB для селектора класса

Мы будет говорить о коде селектора класса с большим числовым значением, нежели у другого кода селектора класса, как о коде с более высоким приоритетом. Набор PHB, отображаемых с помощью 8 битов кода селектора класса, **должен** обеспечивать по крайней мере два независимо пересылаемых класса трафика и PHB, выбираемым по коду селектора класса, **следует** давать возможногсть маркировать пакеты кодом селектора класса с более низким приоритетом при заданных условиях и уровне трафика. Отбрасываемые пакеты рассматриваются, как особый случай несвоевременной пересылки. В дополнение к этому, PHB, выбираемые по кодам 11x000, **должны** обеспечивать пакетам режим предпочтительной пересылки по отношению к пакетам PHB, выбранного по коду 000000 для сохранения сложившегося использования значений IP Precedence 110 и 111 для трафика протоколов маршрутизации.

Далее, для PHB, выбранных по разным кодам селектора класса, **следует** обеспечивать независимую пересылку, т. е., порядок пакетов с разными кодами селектра класса **может** изменяться. Узел сети **может** накладывать ограничения на ресурсы, отдаваемые каждому из таких PHB.

Группы PHB, чьи спецификации удовлетворяют приведенным здесь требованиям, называют совместимые с селектором класса PHB[3].

Требования к PHB для селектора класса с кодом 000000 сравнимы с требованиями к принятому по умолчанию PHB (см. параграф 4.1).

4.2.2.3 Использование требований к PHB для селектора класса и совместимость с IP Precedence

Поддерживающий DS сетевой узел может быть реализован с набором из одной или множества групп PHB, совместимых с селекторами класса. В этом документе устанавливается, что набор кодов xxx000 **должен** отображаться на такой набор PHB. Допускается отображение множества кодов на один PHB – производитель оборудования или администратор сети **может** настроить сетевой узел для отображения на PHB без учета битов 3 - 5 поля DSCP для организации работы сети в соответствии со сложившейся практикой использования поля IP Precedence. Например, код 011010 может отображаться на один вариант PHB с кодом 011000.

---

[1]Class Selector Codepoint.

[2]Class Selector PHB Requirement.

[3]Class Selector Compliant PHB.

4.2.2.4 Пример механизма реализации групп, совместимых с требованиями к PHB для селектора класса

Поддерживающие селекторы класса PHB можно реализовать множеством способов, включая строгую приоритизацию очередей[1], беспристрастные взвешенные очереди (WFQ[2]), WRR[3] или варианты [RPS, HPFQA, DRR], очереди по классам [CBQ[4]]. Более предметное рассмотрение PHB и механизмов их реализации приведено в разделе 5.

Важно отметить, что эти механизмы могут быть доступны через другие PHB (возможно, стандартизованные), которые поддерживаются оборудованием конкретного производителя. Например, в будущих документах может быть стандартизована группа Strict Priority Queueing PHB для набора рекомендованных кодов. Администратор сети может настроить маршрутизаторы на выбор такой группы для кодов xxx000 в соответствии с данным документом.

В качестве другого примера может служить производитель, реализовавший в своих маршрутизаторах механизм CBQ. Этот механизм может использоваться для реализации PHB со строгой приоритизацией очередей, а также поддерживающих селекторы класса PHB с широким спектром возможностей, которые будут доступны набору PHB, удовлетворяющих лишь требованиям к PHB для селекторов класса.

4.3 Выводы

Этот документ определяет группу кодов xxx000, как коды селекторов класса и соответствующие этим кодам PHB **должны** удовлетворять требованиям к PHB для селектора класса, описанным в параграфе 4.2.2.2. Это сделано для сохранения совместимости со сложившейся практикой использования поля IP Precedence в сети Internet без снижения уровня гибкости в будущем. Код 000000 используется для принятого по умолчанию PHB в сети и, следовательно, не настраивается. Оставшиеся семь ненулевых кодов селектора класса настраиваются для расширений PHB, соответствующих требованиям параграфа 4.2.2.2.

5. Руководство по стандартизации поэтапного поведения

Стандартизуются характеристики поведения PHB, а не отдельные механизмы или алгоритмы, используемые для реализации PHB. Узел может иметь (возможно большой) набор параметров, которые могут использоваться для управления пакетами на выходном интерфейсе (например, N отдельных очередей с управляемым уровнем приоритета, размер очередей, веса при циклическом переборе, алгоритм отбрасывания, веса и пороги при отбрасывании и т. п.). Чтобы показать различия между PHB и механизмами, отметим, что PHB, совместимые с селекторами класса, могут быть реализованы с использованием нескольких механизмов, включая очереди со строгой приоритизацией, WFQ, WRR или варианты [HPFQA, RPS, DRR], CBQ [CBQ], а также комбинации этих механизмов.

Спецификации PHB могут задаваться индивидуально или группами (отдельный PHB является вырожденной группой). Группа PHB обычно представляет собой набор из двух или более PHB, которые могут быть специфицированы и реализованы только совместно по причине наличия общих ограничений, применимых к каждому PHB в группе (таких, как обслуживание очередей или политика управления очередями). В спецификации группы PHB **следует** описывать условия, при которых маркировка пакетов может изменяться для выбора другого PHB из группы. PHB **рекомендуется** не менять порядка пакетов в микропотоке. Спецификация группы PHB **должна** идентифицировать все возможные случаи изменения порядка пакетов, которые могут происходить для каждого отдельного PHB или в ситуациях, когда пакеты одного микропотока маркируются для разных PHB в группе.

Стандартизовать **следует** только те PHB, которые не включены в существующие стандарты, но реализованы, развернуты и показали свою применимость. Поскольку опыта использования дифференцированных услуг пока не достаточно, преждевременно строить какие-либо гипотезы в части спецификации PHB.

Для каждого стандартизованного PHB **должен** предлагаться **рекомендуемый** код, выделенный из пространства 32 кодов (см. раздел 6). В данной спецификации пространство кодов не распределяется, чтобы обеспечить возможность развития. Таким образом, коды xxx000 оставлены свободными преднамеренно.

Производители сетевого оборудования вольны предлагать те параметры и возможности, которые представляются полезными и востребованными. При реализации на узле конкретного стандартизованного PHB производитель **может** использовать любые алгоритмы, которые соответствуют определению PHB в стандарте. Режим обработки пакетов определяется возможностями узла и его конфигурационными параметрами.

Сервис-провайдеры не обязаны использовать одинаковые механизмы и конфигурации в своей сети для предоставления дифференцированных услуг и могут выбирать конфигурационные параметры узлов в соответствии с предлагаемыми услугами и задачами управления трафиком. Ясно, что с течением времени могут появиться и широко распространиться PHB, которые особенно полезны для организации сквозной дифференциации услуг. Такие варианты поведения могут быть связаны с конкретными кодами EXP/LU[5] PHB в поле DS, что позволит дифференцировать услуги через границы доменов (см. раздел 6). Такие PHB являются кандидатами для стандартизации.

Спецификации стандартизованных PHB **рекомендуется** задавать в соответствии с рекомендациями [ARCH].

---

[1]Strict Priority Queueing.

[2]Weighted fair queueing.

[3]Weighted Round Robin - взвешенный циклический перебор. *Прим. перев.*

[4]Class-Based Queuing.

[5]Для экспериментального и локального использования.

6. Согласование с IANA

Набор	Пространство кодов	Правила распределения
1	xxxxx0	Стандартизация
2	xxxx11	EXP/LU
3	xxxx01	EXP/LU

Поле DSCP внутри поля DS позволяет задать 64 различных кода. Пространство кодов делится на три части: 32 **рекомендуемых** кода (набор 1) выделяются путем стандартизации (Standards Action) в соответствии с [CONS], группа из 16 кодов (набор 2) резервируется для экспериментов и локального использования[1] (EXP/LU) в соответствии с [CONS], а последняя группа из 16 кодов (набор 3) изначально была выделена для экспериментов и локального применения, но ее следует сохранять, прежде всего, для стандартного распределения по мере исчерпания набора 1. Наборы кодов показаны в таблице (x может принимать значения 0 или 1):

В этом документе выделено восемь **рекомендуемых** кодов (xxx000) из набора 1. Эти коды **должны** отображаться не на конкретные PHB, а на PHB, которые удовлетворяют, по меньшей мере, требованиям параграфа 4.2.2.2 для обеспечения минимального уровня совместимости с полем IP Precedence, определенным в [RFC791], и сложившейся практикой использования этого поля.

7. Вопросы безопасности

В этом разделе рассматриваются вопросы безопасности, связанные с введением дифференциации услуг, к которым, прежде всего, отноятся возможные атаки на отказ служб (DoS) и возможности несанкционированного использования ресурсов (параграф 7.1). В параграфе 7.2 рассматривается диффренциация услуг при использовании IPsec, включая интеграцию с туннельным режимом IPsec и другими протоколами туннелирования. Более детальное рассмотрение вопросов безопасности, связанных с дифференциацией услуг, содержится в документе [ARCH].

7.1 Несанкционированное использование услуг и DoS-атаки

Основной целью дифференцированных услуг является предоставление различных уровней сервиса для потоков трафика в сетевой инфраструктуре общего пользования. Для достижения этой цели может использоваться множество методов, но в конечном итоге одни пакеты будут обслуживаться лучше, нежели другие. Отображение сетевого трафика на специфические варианты поведения будет приводить к другому (лучше или хуже) обслуживанию в зависимости, прежде всего, от значения кода DS и, следовательно, возникает соблазн получить более высокий уровень обслуживания за счет изменения кодов, определяющих поведение, или загрузить сеть пакетами с кодом для высокого уровня обслуживания. Возможности такие действий ограничены, но они, по сути, позволяют организовать атаки на отказ служб, поскольку изменение кода или вставка пакетов с кодом высокоприоритетного обслуживания может приводить к исчерпанию ресурсов, потребных для обработки других потоков трафика.

Защита от этого класса атак обеспечивается комбинацией кондиционирования трафика на границах домена DS со средствами контроля целостности и защиты сетевой инфраструктуры внутри домена DS. Граничные узлы домена DS **должны** гарантировать для всего входящего в домен трафика маркировку кодами, подходящими для трафика и защиты от несанкционированного использования ресурсов и атак на службы, основанных на несанкционированном изменении кодов. Важно отметить, что любой генерирующий трафик узел внутри домена DS является граничным узлом для этого трафика. Внутренние узлы домена DS на основе кодов DS связывают трафик с PHB и от этих узлов **не требуется** проверять коды перед его использованием. В результате внутренние узлы зависят от работы граничных узлов домена DS в части предотвращения трафика с неприемлемыми кодами и перерасхода ресурсов, которые могут нарушить работу домена.

7.2 Взаимодействие с IPsec и туннелями

Протокол IPsec в соответствии с определением [ESP, AH] не включает поля DS заголовков IP в криптографические преобразования (в туннельном режиме не включается в преобразования поле внешнего заголовка IP). Следовательно, изменение поля DS в сети не оказывает влияния на сквозную защиту IPsec, поскольку такое изменение не влияет на контроль целостности IPsec. В результате IPsec не обеспечивает никакой защиты против изменения значений поля DS (т. е., MITM[2]-атак), поскольку такое изменение не оказывает влияния на сквозную защиту IPsec.

Туннельный режим IPsec обеспечивает защиту для поля DS инкапсулированных заголовков IP. Пакет в туннельном режиме IPsec включает два заголовка IP - внешний, который создается на входе в туннель, и инкапсулированный внутренний заголовок созданный отправителем пакета. При организации туннеля IPsec (полностью или частично) в сети с дифференцированным обслуживанием промежуточные узлы сети имеют дело с полем DS внешнего заголовка. На выходе из туннеля IPsec удаляет внешний заголовок и пересылает (при необходимости) пакет в соответствии с внутренним заголовком. Протокол IPsec **требует**, чтобы поле DS внутреннего заголовка не менялось при декапсуляции дабы предотвратить возможность организации атак путем изменения DS через конечную точку туннеля IPsec. Этот документ не меняет данного требования. Если внутренний заголовок IP не обрабатывается граничным узлом DS для выходного узла туннеля в домене DS, выход туннеля является граничным узлом для выходящего из туннеля трафика и, следовательно, **должен** гарантировать, что вышедший из туннеля трафик имеет приемлемые коды DS.

Когда процесс декапсуляции на выходе туннеля IPsec включает достаточно строгую криптографическую проверку целостности инкапсулированного пакета (строгость определяется локальной политикой), этот узел может безопасно предполагать, что значение поля DS во внутреннем заголовке совпадает со значением этого поля на входе в туннель. Важным следствием этого является возможность защиты каналов внутри домена DS с помощью туннелей IPsec с достаточно строгой криптозащитой. Этот вывод применим ко всем протоколам туннелирования, обеспечивающим контроль целостности, но уровень защиты поля DS во внутреннем заголовке зависит от строгости проверки целостности, выполняемой туннельным протоколом. При отсутствии достаточных гарантий для туннелей, промежуточные узлы которых находятся за пределами домена DS (уязвимы), инкапсулированные пакеты **должны** трактоваться на границе домена DS, как пакеты, полученные извне.

[1]См. одноименный раздел в RFC 3260, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

[2]Man-in-the-middle attack — атака с перехватом и изменением данных на пути доставки при участии человека.

8. Благодарности

Авторы признательны членам рабочей группы Differentiated Services за полезные дискуссии в процессе подготовки документа.

9. Литература

[AH] Kent, S. and R. Atkinson, "IP Authentication Header", RFC 2402[1], November 1998.

[ARCH] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", RFC 2475[2], December 1998.

[CBQ] S. Floyd and V. Jacobson, "Link-sharing and Resource Management Models for Packet Networks", IEEE/ACM Transactions on Networking, Vol. 3 no. 4, pp. 365-386, August 1995.

[CONS] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 2434, October 1998.

[DRR] M. Shreedhar and G. Varghese, Efficient Fair Queueing using Deficit Round Robin", Proc. ACM SIGCOMM 95, 1995.

[ESP] Kent, S. and R. Atkinson, "IP Encapsulating Security Payload (ESP)", RFC 2406[3], November 1998.

[HPFQA] J. Bennett and Hui Zhang, "Hierarchical Packet Fair Queueing Algorithms", Proc. ACM SIGCOMM 96, August 1996.

[IPv6] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC791] Postel, J., Editor, "Internet Protocol", STD 5, RFC 791[2], September 1981.

[RFC1122] Braden, R., "Requirements for Internet hosts - communication layers", STD 3, RFC 1122[2], October 1989.

[RFC1812] Baker, F., Editor, "Requirements for IP Version 4 Routers", RFC 1812[2], June 1995.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119[2], March 1997.

[RPS] D. Stiliadis and A. Varma, "Rate-Proportional Servers: A Design Methodology for Fair Queueing Algorithms", IEEE/ACM Trans. on Networking, April 1998.

Адреса авторов

Kathleen Nichols

Cisco Systems

170 West Tasman Drive

San Jose, CA 95134-1706

Phone: +1-408-525-4857

EMail: kmn@cisco.com

Steven Blake

Torrent Networking Technologies

3000 Aerial Center, Suite 140

Morrisville, NC 27560

Phone: +1-919-468-8466 x232

EMail: slblake@torrentnet.com

Fred Baker

Cisco Systems

519 Lado Drive

Santa Barbara, CA 93111

Phone: +1-408-526-4257

EMail: fred@cisco.com

David L. Black

EMC Corporation

[1]Этот документ устарел и заменен RFC 4302, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

[2]Перевод этого документа имеется на сайте www.protocols.ru. *Прим. перев.*

[3]Этот документ устарел и заменен RFC 4306, перевод которого имеется на сайте www.protocols.ru. *Прим. перев.*

35 Parkwood Drive

Hopkinton, MA 01748

Phone: +1-508-435-1000 x76140

EMail: black_david@emc.com

Перевод на русский язык

Николай Малых

nmalykh@bilim.com

Полное заявление авторских прав

Copyright (C) The Internet Society (1998). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.