Network Working Group
Request for Comments: 3168
Updates: 2474, 2401, 793
Obsoletes: 2481
Category: Standards Track

K. Ramakrishnan
TeraOptic Networks
S. Floyd
ACIRI
D. Black
EMC

September 2001

Добавление явных уведомлений о перегрузке (ECN) в IP

The Addition of Explicit Congestion Notification (ECN) to IP

Статус документа

Документ содержит спецификацию стандартного протокола Internet, предложенного сообществу Internet, и служит приглашением к дискуссии в целях развития протокола. Информацию о текущем состоянии стандартизации протокола можно найти в документе Internet Official Protocol Standards (STD 1). Документ можно распространять без ограничений.

Аннотация

В документе приведена спецификация включения ECN[1] в протоколы TCP и IP, включая использование для ECN двух битов заголовка IP.

Оглавление

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

1. Введение

Сначала рассматривается использование протоколом TCP процедуры отбрасывания пакетов для индикации перегрузки. Далее даются пояснения по поводу новых возможностей, обусловленных добавлением активного управления очередями (например, RED[1]) в инфраструктуру Internet, когда маршрутизаторы детектируют перегрузку до того, как будет переполнена очередь, и уже не обязаны отбрасывать пакеты для индикации перегрузки. Маршрутизаторы могут устанавливать маркер CE[2] в заголовках пакетов IP от поддерживающих ECN транспортных протоколов. Описано, когда маршрутизаторы устанавливают маркер CE, а также рассмотрены изменения, которые нужно внести в протокол TCP для поддержки ECN. Изменения других протоколов транспортного уровня (например, негарантированная доставка пакетов unicast или multicast, гарантированная доставка multicast-пакетов, другие протоколы гарантированной доставки пакетов unicast) могут рассматриваться при разработке или стандартизации таких протоколов. В данном документе описываются также вопросы использования ECN в туннелях IP и, в частности, в IPsec-туннелях.

Одним из основных принципов предлагаемого расширения протокола является возможность постепенного развёртывания. Наличие некоторых туннелей IP, не совместимых с ECN, является одним из препятствующих постепенному развёртыванию обстоятельств. По мере развёртывания. ECN несовместимые туннели IP должны быть изменены для того, чтобы соответствовать данной спецификации.

Данный документе отменяет RFC 2481 «A Proposal to add Explicit Congestion Notification (ECN) to IP», в котором ECN определяется как экспериментальный протокол для сообщества Internet. Кроме того, данный документ обновляет RFC 2474 «Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers» в части определения поля ECN в заголовке IP, RFC 2401 «Security Architecture for the Internet Protocol» (изменение обработки байтов IPv4 TOS и IPv6 Traffic Class Octet при создании в туннельном режиме заголовков, совместимых с ECN) и RFC 793 «Transmission Control Protocol» в части определения новых флагов заголовка TCP.

Алгоритмы контроля и предотвращения перегрузки протокола TCP основаны на представлении сети, как «чёрного ящика» [Jacobson88, Jacobson90]. Перегрузка или её отсутствие определяется конечными системами путём проверки состояния сети за счёт увеличения нагрузки (расширения окна перегрузки - числа пакетов, остающихся в сети) до тех пор, пока не возникнет перегрузка и связанная с ней потеря пакетов. Трактовка сети, как чёрного ящика, и потери пакетов, как индикатора перегрузки, приемлема для случаев передачи по протоколу TCP данных, которые не чувствительны или не критичны к задержкам или потере отдельных пакетов. Алгоритмы контроля перегрузки в TCP используют встроенные методы (такие, как Fast Retransmit и Fast Recovery[3]) для минимизации влияния потерь с точки зрения пропускной способности. Однако эти алгоритмы не подходят для приложений, которые чувствительны к задержкам или потере одного или нескольких пакетов. Интерактивные системы (например, telnet, просмотр web-страниц, передача аудио и видео-потоков) могут быть чувствительны к потере пакетов (при использовании транспорта без гарантии доставки, такого как UDP) или увеличению задержки, вызванному повтором передачи в результате потери пакета (при использовании гарантированного транспорта, такого как TCP).

Поскольку TCP определяет приемлемый размер окна перегрузки путём увеличения размера этого окна до тех пор, пока не начнётся потеря (отбрасывание) пакетов, это может приводить к переполнению очередей в маршрутизаторах, являющихся узким местом в сети. Большинство алгоритмов отбрасывания пакетов в маршрутизаторах не чувствительно к нагрузке, создаваемой отдельным потоком, что означает возможность отбрасывания пакетов из некоторых потоков, чувствительных к задержкам.

Активные механизмы управления очередями детектируют перегрузку до того, как переполнится очередь, и обеспечивают индикацию перегрузки для конечных узлов. Преимущества активного управления очередями обсуждаются в RFC 2309 [RFC2309]. Такое управление позволяет избавиться от некоторых негативных свойств систем управления очередями, основанных на отбрасывании пакетов при переполнении (в частности, от нежелательной синхронизации потери пакетов во множестве потоков данных). Более важно, что в системах активного управления

[1]Random Early Detection - упреждающее детектирование.
[2]Congestion Experienced - наблюдается перегрузка.
[3]Ускоренный повтор передачи и быстрое восстановление

очередями транспортные протоколы с контролем перегрузки (например, TCP) не используют переполнение буферов в качестве индикатора перегрузки.

Механизмы активного управления очередями могут применять один из нескольких методов индикации перегрузки для конечных узлов. Одним из методов является отбрасывание пакетов, как это делается сейчас. Однако активное управление очередями позволяет маршрутизатору отделить правила буферизации (включения в очередь) и отбрасывания пакетов от политики индикации перегрузки. Таким образом, активное управление очередями позволяет маршрутизаторам использовать маркер CE в заголовке пакета для индикации перегрузки вместо того, чтобы полагаться исключительно на отбрасывание пакетов. Это может снизить избыточные задержки в очереди для всех типов трафика, использующих эту очередь.

В сети Internet существуют промежуточные устройства (межсетевые экраны, системы распределения нагрузки, системы детектирования попыток вторжения), которые отбрасывают пакеты TCP SYN, предназначенные для согласования ECN, или отвечают на них пакетами RST. В данном документе описывается процедура, которая может использоваться в реализациях TCP для обеспечения надёжной работы даже при наличии таких устройств.

2. Соглашение об использовании терминов

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].

3. Исходные допущения и основные принципы

В этом разделе рассматриваются некоторые важные принципы и допущения на которых основана работа предлагаемого расширения.

- В силу очевидности постепенной адаптации ECN существенное значение имеет стратегия перехода. Некоторые маршрутизаторы могут по-прежнему использовать лишь отбрасывание пакетов для индикации перегрузки, а часть конечных систем может не поддерживать ECN. Наиболее подходящей стратегией является постепенное развёртывание без возникновения «островков», которые поддерживают (или, наоборот, не поддерживают) ECN.
- Новые механизмы контроля и предотвращения перегрузки должны сосуществовать и кооперироваться с имеющимися механизмами контроля перегрузки. В частности, новые механизмы должны сосуществовать с методами, используемыми в TCP, и принятой в современных маршрутизаторах практикой отбрасывания пакетов в периоды перегрузки.
- Длительность перегрузки может меняться в широких пределах и превышать время кругового обхода (RTT).
- Число пакетов в отдельном потоке (например, в соединении TCP или сеансе обмена данными по протоколу UDP) также может меняться в широких пределах. Мы заинтересованы в контроле перегрузки, создаваемой потоком, который передаёт достаточно большое количество данных, чтобы такой поток оставался активным до того, как будет получен сигнал обратной связи из сети.
- Очевидно, что асимметрия маршрутов является нормальным явлением в Internet. Путь (последовательность каналов и маршрутизаторов), по которому данные следуют из одной точки в другую в одном направлении, может отличаться от пути между той же парой точек в обратном направлении (например, для передачи подтверждений).
- Многие маршрутизаторы более эффективно обрабатывают заголовки пакетов IP без опций, нежели опции IP. Это служит предпосылкой включения индикации перегрузки в обычный заголовок пакета IP, а не в поле опций.
- Следует признать, что не все конечные системы могут кооперироваться в плане контроля перегрузки. Однако новым механизмам не следует упрощать для приложений TCP запрет контроля перегрузки. Преимущества от использования новых механизмов типа ECN не должны быть значительными.

4. Активное управление очередями AQM

Упреждающее детектирование RED представляет собой один из механизмов активного управления очередями (Active Queue Management или AQM), предложенный для детектирования начинающейся перегрузки [FJ93] и развёрнутый уже в Internet [RFC2309]. AQM отбрасывает пакеты в тех случаях, когда средний размер очереди превышает пороговое значение, не дожидаясь переполнения очереди. Однако, когда AQM отбрасывает пакеты до переполнения очереди, это отбрасывание не всегда вызвано нехваткой памяти.

AQM может устанавливать маркер CE в заголовке пакета вместо того, чтобы отбросить пакет, если такое поле присутствует в заголовке IP и понятно транспортному протоколу. Использование маркера CE с ECN будет позволять получателю избавиться от избыточной задержки, связанной с повтором передачи после отбрасывания пакета. Для обозначения пакетов с установленным маркером CE далее будет использоваться термин CE-пакет.

5. Явное уведомление о перегрузке в IP

В этом документе указывается, что Internet обеспечивает индикацию наступающей перегрузки (как в RED и более ранней работе [RJ90]) путём маркировки пакетов, а не их отбрасывания. Этот механизм использует двухбитовое поле ECN в заголовке IP, которое может содержать значения от 00 до 11. В поддерживающем ECN транспорте (ECN-Capable Transport или ECT) отправитель устанавливает маркер 10 или 01 для индикации поддержки конечной точкой возможностей ECN, эти маркеры обозначаются ECT(0) и ECT(1), соответственно. Термин «маркер ECT» в данном документе будет использоваться применительно к обоим случаям. Маршрутизаторы трактуют коды ECT(0) и ECT(1) одинаково[1]. Отправитель может выбрать любой из маркеров ECT(0) или ECT(1) для индикации ECT и использовать этот маркер для последующих пакетов.

Использование двух маркеров ECT обусловлено, прежде всего, желанием предоставить отправителю механизм проверки того, что маркер CE не теряется в сети и получатель ответит отправителю пакетов с маркером CE в

[1] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*

соответствии с требованиями транспортного протокола. Рекомендации для отправителей и получателей по дифференциации маркеров ECT(0) и ECT(1) следует указать в отдельных документах для каждого из транспортных протоколов. В частности, данный документ не описывает различий между маркерами ECT(0) и ECT(1) для протокола TCP. Протоколам и отправителям, которым требуется единственный маркер ECT, **следует** использовать ECT(0)[1].

```
+-----+-----+
| Поле  ECN |
+-----+-----+
  ECT   CE     [устаревшие] обозначения битов ECN в RFC 2481.
   0     0     Not-ECT
   0     1     ECT(1)
   1     0     ECT(0)
   1     1     CE
```

Рисунок 1. Поле ECN в заголовке IP.

Маркер not-ECT (00) указывает пакет, в котором не используется ECN. Маркер CE (11) устанавливают маршрутизаторы для индикации перегрузки конечным точкам. Маршрутизаторы, получающие пакеты при отсутствии места в очереди, просто отбрасывают такие пакеты, как это делается без ECN.

Применение двухбитовых маркеров ECT вступает в противоречие с использованием временных полей ECN в заголовках пакетов и маршрутизаторы должны удалять такие флаги при установке маркера CE [SCWA99]. Например, маршрутизатор, удаляющий маркер CE, будет сталкиваться с дополнительными сложностями при попытках восстановления исходного поля и, таким образом, повторное удаление маркера CE будет с большей очевидностью обнаружено конечными точками. Временные поля ECN могут также приводить к проблемам, связанным с тем, что некоторые получатели не будут информировать отправителя о наличии в пакете маркера CE. Мотивация выбора двухбитовых маркеров ECT более детально рассматривается в разделе 20 вместе с обсуждением варианта использования четвёртого маркера ECT (01). Вопросы совместимости с ранними реализациями ECN, которые не понимают маркеров ECT(1), рассматриваются в разделе 11.[2]

В RFC 2481 [RFC2481] поле ECN было разделено на две части - бит ECT и бит CE. Поле ECN, в котором установлен только бит ECT (в соответствии с RFC 2481), эквивалентно маркеру ECT(0), а поле ECN, в котором установлены оба флага ECT и CE, - маркеру CE. Маркер 01 не определён в RFC 2481 и по этой причине следует использовать маркер ECT(0) в тех случаях, когда требуется единственный маркер ECT.

```
    0     1     2     3     4     5     6     7
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Поле DS, DSCP           | Поле  ECN |
+-----+-----+-----+-----+-----+-----+-----+-----+

    DSCP: код дифференцированного обслуживания
    ECN:  явное уведомление о перегрузке
```

Рисунок 2. Поля Differentiated Services и ECN в заголовке IP.

Биты 6 и 7 октета IPv4 TOS обозначены как поле ECN. Октет IPv4 TOS соответствует октету Traffic Class в IPv6, а поле ECN определяется одинаково для обоих случаев. Определения для октета IPv4 TOS [RFC791] и октета IPv6 Traffic Class были отменены определением шестибитового поля DS (Differentiated Services) [RFC2474, RFC2780]. Биты 6 и 7 указаны в [RFC2474], как неиспользуемые (Currently Unused), а в RFC 2780 - как предложенные для экспериментального использования в ECN. В разделе 22 приведена краткая ретроспектива применения октета TOS.

По причине изменений в характере использования TOS описанное здесь применение поля ECN не может гарантировать совместимости со всеми прежними вариантами использования двух битов, отведённых для поля ECN. Потенциальные проблемы, связанные с отсутствием совместимости рассматриваются в разделе 22.

При получении поддерживающим ECN транспортным протоколом одного CE-пакета алгоритм контроля перегрузки в конечной системе **должен** работать в точности так же, как при индикации перегрузки путём отбрасывании одного пакета[3]. Например, для поддерживающей ECN реализации протокола TCP требуется, чтобы отправитель уменьшил вдвое размер окна перегрузки для любого окна данных, в котором произошло отбрасывание пакета или был получен индикатор ECN.

Одной из причин требования идентичной реакции на индикацию перегрузки при получении CE-пакета и отбрасывании пакета является необходимость обеспечения постепенного развёртывания. ECN как в конечных системах, так и в маршрутизаторах. Некоторые маршрутизаторы могут отбрасывать ECN-пакеты[4] (например, при использовании некоторых правил обнаружения перегрузки в AQM), тогда как другие маршрутизаторы будут устанавливать маркер CE при таких же условиях перегрузки. Подобно этому, маршрутизатор может отбрасывать не поддерживающие ECN пакеты или устанавливать бит CE в ECN-пакетах при одинаковых условиях перегрузки. Разная реакция на индикацию перегрузки путём установки бита CE и отбрасывания пакетов может приводить к различным (неправильным) трактовкам для разных потоков.

Дополнительное требование заключается в том, что конечным системам следует реагировать на перегрузку не более одного раза в расчёте на окно данных (т. е., не более одного раза за период кругового обхода), чтобы избежать множественной реакции на несколько фактов индикации перегрузки в течение одного периода кругового обхода.

Маршрутизаторам[5] **следует** устанавливать маркер CE в ECN-пакетах только в тех случаях, когда маршрутизатор должен был бы отбросить пакет для индикации перегрузки конечным узлам. Когда буфер ещё не заполнен, но маршрутизатор уже приготовился к отбрасыванию пакетов для уведомления конечных узлов о перегрузке, этому

[1] RFC 8311 заменяет этот абзац и последнее предложение предыдущего текстом: «Протоколы и отправители **должны** применять код ECT(0) для индикации ECT, если не указано иное в Experimental RFC потока документов IETF. Протоколам и отправителям **недопустимо** использовать код ECT(1) для индикации ECT, если не указано иное в Experimental RFC потока документов IETF. Рекомендации для отправителей и получателей по различению кодов ECT(0) и ECT(1) будут приведены в отдельных документах для каждого транспортного протокола. Этот документ не рассматривает механизмы различения кодов ECT(0) и ECT(1) для конечных узлов TCP.».

[2] В соответствии с RFC 8311 этот абзац исключён. *Прим. перев.*

[3] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*

[4] В оригинале ECN-Capable - пакеты, для которых может поддерживаться ECN. *Прим. перев.*

[5] RFC 8311 заменяет это слово фразой: «Если иное не указано в Experimental RFC потока документов IETF, ». *Прим. перев.*

маршрутизатору следует сначала проверить наличие маркера ECT в заголовке IP. Если маркер установлен, вместо отбрасывания пакетов маршрутизатор **может** помещать маркер CE в заголовок IP.

Среда, где все конечные узлы поддерживают ECN, позволяет разработать новые критерии установки маркера CE и механизмы контроля перегрузки для реакции конечных узлов на CE-пакеты. Однако рассмотрение этого вопроса выходит за рамки данного документа.

Когда маршрутизатор получает CE-пакет, маркер CE остаётся без изменения и пакет передаётся как обычно. При возникновении некоторой перегрузки и заполнении очереди маршрутизатора, последний может принять решение об отбрасывании некоторых пакетов при поступлении новых. Предполагается, что такое отбрасывание пакетов станет сравнительно редким явлением, когда большая часть конечных систем будет поддерживать ECN и механизмы контроля перегрузки TCP или аналогичные механизмы. В корректно организованной сети, работающей в среде с поддержкой ECN, потери пакетов могут происходить в периоды неустойчивости или при наличии не поддерживающих контроль перегрузки отправителей.

Рассмотренная выше ситуация, когда установка CE используется взамен отбрасывания пакетов, применяется по умолчанию ко всем Differentiated Service PHB[1] [RFC 2475]. Спецификации для PHB **могут** более детально указывать, как совместимые реализации должны выбирать между отбрасыванием пакетов и установкой CE, но такая детализация **не требуется**. Для маршрутизатора **недопустима** установка CE вместо отбрасывания пакета, когда это отбрасывание обусловлено причинами, отличными от подступающей перегрузки или желания указать конечным узлам на начинающуюся перегрузку (например, граничный узел diffserv может быть настроен на безусловное отбрасывание некоторых классов трафика на входе в домен).

Предполагается, что маршрутизаторы будут устанавливать маркер CE при начинающейся перегрузке, определяемой по среднему размеру очереди с использованием алгоритма RED, предложенного в [FJ93, RFC2309]. По имеющимся у авторов сведениям этот вариант является единственным предложением, обсуждаемым IETF, для упреждающего отбрасывания пакетов маршрутизаторами до переполнения буферов. Однако данный документ не пытается задавать тот или иной механизм активного управления очередями, оставляя решение этого вопроса (если он возникнет) за IETF. Хотя использование ECN связано с вопросом о необходимости иметь подходящий механизм активного управления очередями в маршрутизаторах, авторы не настаивают на использовании именно этого механизма. Методы активного управления очередями были разработаны и развёрнуты независимо от ECN с отбрасыванием пакетов для индикации перегрузки ещё до начала использования ECN в архитектуре IP.

5.1. ECN как индикация установившейся перегрузки

Подчеркнём, что **единичный** пакет с установленным маркером CE в заголовке IP приводит к тому, что система контроля перегрузки транспортного протокола реагирует на, это как на отбрасывание пакета. Очевидно, что размер очереди постоянно меняется, если маршрутизатор не находится в состоянии установившейся перегрузки. Важно, чтобы кратковременная перегрузка маршрутизатора, отражающаяся в том, что мгновенный размер очереди превышает пороговое значение, которое существенно меньше ёмкости очереди, не вызывала реакции транспортного уровня, как при перегрузке. Следовательно, маркер CE не следует устанавливать на основе мгновенных значений размера очереди в маршрутизаторе.

Например, механизмы индикации перегрузки ATM и Frame Relay обычно определяются без связи со средним размером очереди, как индикатором перегрузки промежуточного узла. Разумно предположить, что такая индикация будет создавать избыточный шум. Реакция отправителя TCP в соответствии с данной спецификацией для ECN не **является** подходящим вариантом для такого шумного сигнала о перегрузке. Однако, если маршрутизаторы, имеющие интерфейсы в сети ATM, получат способ определения среднего размера очереди для интерфейса и станут использовать этот размер для надёжного обнаружения перегрузки в подсети ATM, они могут использовать уведомления ECN, описанные в данной спецификации.

Мы продолжаем поощрять эксперименты с методами реализации преимуществ ECN на уровне 2 (например, в коммутаторах ATM или Frame Relay). К примеру, используя схемы типа RED (когда пакеты маркируются на основе превышения средним размером очереди заданного порога), устройства канального уровня могут обеспечивать достаточно надёжную индикацию перегрузки. Когда все устройства уровня 2 на пути установят принятый на этом уровне маркер возможной перегрузки (например, бит EFCI для ATM или FECN для Frame Relay) с использованием надёжного детектирования перегрузки, интерфейс маршрутизатора в сеть уровня 2 сможет транслировать такую индикацию в маркеры CE заголовков IP. Мы признаем, что в сегодняшней практике и стандартах такого не наблюдается. Однако продолжение экспериментов в этом направлении может дать информацию, которая позволит найти способ перехода от имеющихся механизмов канального уровня к более надёжной индикации перегрузки с использованием одного бита.

5.2. Отброшенные и повреждённые пакеты

Для предлагаемого в этом документе использования ECN (т. е., для транспортных протоколов, таких как TCP, где отбрасывание данных является индикацией перегрузки) конечные узлы видят отбрасывание пакетов данных и отклик (о перегрузке) от обнаруживших такое отбрасывание конечных узлов имеет по крайней мере такую же силу, как отклик на получение пакета CE. Для гарантированной доставки индикации перегрузки с помощью кода CE **недопустимо** передавать код ECT в пакете, пока потеря пакета в сети не будет обнаружена конечными узлами и интерпретирована как индикация перегрузки[2].

Транспортные протоколы, такие как TCP, не обязательно детектируют отбрасывание любых пакетов (в частности, пакетов, содержащих лишь подтверждение ACK), например, TCP не снижает скорость доставки последующих пакетов ACK в ответ на ранее отброшенные пакеты ACK. Любые предложения по расширению ECN на такие пакеты будут приводить к возникновению проблем, таких, как маркировка пакета ACK кодом CE и последующее отбрасывание такого пакета в сети. Мы надеемся, что этот аспект будет исследован специально, поэтому в настоящем документе указывается, что «чистые» пакеты ACK **недопустимо** использовать для индикации поддержки ECN.

[1] Per-Hop Behavior - поведение для интервала пути.

[2] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*

Аналогично, если пакет с маркировкой CE отбрасывается в сети по причине повреждения (битовые ошибки), конечным узлам следует по-прежнему вводить контроль перегрузки так же, как TCP реагирует в настоящее время на отбрасывание пакета данных. Вопрос повреждения пакетов CE будет рассматриваться в любых предложениях по способам определения был пакет отброшен в результате повреждения или по причине перегрузки/переполнения буфера. В частности, повсеместное развёртывание ECN не будет мерой, достаточной для того, чтобы позволить конечным узлам интерпретировать отбрасывание пакетов, как результат повреждения, а не перегрузки.

5.3. Фрагментация

Пакеты с поддержкой ECN **могут** иметь установленный флаг DF (не фрагментировать). При сборке фрагментов **недопустимо** терять индикацию перегрузки. Иными словами, если любой из фрагментов собираемого пакета IP имеет код CE, **должны** выполняться одно из двух действий, указанных ниже.

- Установка кода CE для собранного пакета. **Недопустимо** устанавливать этот код, если хотя бы один из собираемых фрагментов имеет код Not-ECT.
- Отбрасывание пакета вместо сборки фрагментов (по любой причине).

Если применимы оба варианта, **можно** выбирать любой. При сборке фрагментов **недопустимо** менять код ECN, если значения кода совпадают во всех фрагментах.

Отметим, что в результате того, что в RFC 2481 не было задано поведение при сборке фрагментов, старые реализации ECN, соответствующие экспериментальному RFC, не обязательно будут выполнять сборку фрагментов корректно в плане сохранения кода CE во фрагментах. Отправитель может предотвратить последствия такой некорректной обработки, устанавливая бит DF в пакетах, поддерживающих ECN.

Могут возникать ситуации, когда приведённая выше спецификация сборки фрагментов будет недостаточно точна. Например, при наличии вредоносных или сбойных элементов пути в точке фрагментации или после неё, фрагменты могут содержать смесь кодов ECT(0), ECT(1), Not-ECT. Спецификация сборки фрагментов, приведённая выше, не включает требований для такого случая. В ситуациях, когда требуется более чёткое поведение при сборке фрагментов, спецификации протокола **следует** вместо этого указывать, что во всех передаваемых протоколом пакетах, способных поддерживать ECN, **должен** устанавливаться флаг DF.

6. Поддержка в транспортном протоколе

ECN требует поддержки со стороны транспортного протокола в дополнение к функциональности, обеспечиваемой полем ECN в заголовке пакета IP. Транспортный протокол может требовать согласования между конечными точками при организации соединения для проверки поддержки ими ECN, чтобы отправитель мог устанавливать код ECT в передаваемых пакетах. Во-вторых, транспортный протокол должен быть способен соответствующим образом реагировать на получение пакетов CE. Эта реакция может выражаться в форме информирования отправителя данных о полученном пакете CE (например, TCP), отказа получателя от участия в многоуровневой multicast-группе (например, RLM [MJV96]) или в ином виде, обеспечивающем, в конечном итоге, снижение скорости поступления потока данных через перегруженное соединение. Пакеты CE показывают скорее долговременную, чем краткосрочную перегрузку (см. параграф 5.1) и поэтому реакция на получение пакетов CE должна соответствовать продолжающейся перегрузке.

В этом документе рассматривается добавление поддержки ECN только для протокола TCP, а рассмотрение вопросов использования ECN в других транспортных протоколах оставлено для будущих исследований. Для TCP добавление ECN требует поддержки трёх новых функций - согласования между конечными точками в процессе организации соединения для проверки поддержки ECN на обеих сторонах, флага ECN-Echo (ECE) в заголовке TCP для информирования отправителя о получении пакета CE, флага CWR[1] в заголовке TCP для информирования получателя о снижении размера окна перегрузки. Очевидно, что функциональность, требуемая от других протоколов (в частности, протоколов групповой адресации с гарантиями доставки и без таковых), будет отличаться и определится при стандартизации этих транспортных протоколов в IETF.

В этом документе используется термин «пакеты TCP» вместо «сегментов TCP», хоть это и не вполне корректно.

6.1. TCP

В последующих параграфах подробно рассматривается предложенное использование ECN в TCP. Эти предложения представлены в той же форме, что и в работе [Floyd94]. Предполагается, что TCP на стороне отправителя использует стандартные алгоритмы контроля перегрузки Slow-start, Fast Retransmit и Fast Recovery [RFC2581].

Это предложение задаёт два новых флага в резервном поле заголовка TCP. Механизм TCP для согласования поддержки ECN использует флаг ECE (ECN-Echo) в заголовке TCP (бит 9 в поле Reserved заголовка TCP). Местоположение 6-битового резервного поля заголовка TCP показано на рисунке 3[2] в RFC 793 [RFC793], приведённом здесь для полноты. Данная спецификация поля ECN оставляет 4-битовое резервное поле (биты 4 - 7).

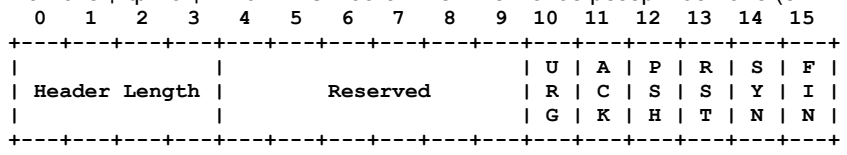


Рисунок 3. Старое определение байтов 13 и 14 в заголовке TCP.

Чтобы обеспечить получателю TCP возможность определения момента прекращения установки флага ECN-Echo, в заголовок TCP добавлен ещё один флаг - CWR. Для флага CWR выделен бит 8 резервного поля в заголовке TCP.

Таким образом, ECN использует флаги ECT и CE в заголовке IP (Рисунок 1) для сигнализации между маршрутизаторами и конечными точками соединений, а также флаги ECN-Echo и CWR в заголовке TCP (Рисунок 4) для

[1]Congestion Window Reduced - окно перегрузки уменьшено.
[2]В оригинале ошибочно дана ссылка на рисунок 4. *Прим. перев.*

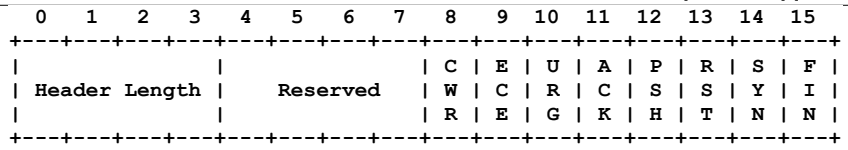


Рисунок 4. Новое определение байтов 13 и 14 в заголовке TCP.

сигнализации между конечными точками TCP. Для соединения TCP типичная последовательность событий в основанной на ECN реакции на перегрузку представлена ниже.

- В передаваемых пакетах устанавливается код ECT для индикации поддержки ECN транспортными элементами.
- Поддерживающий ECN маршрутизатор детектирует приближающуюся перегрузку и видит код ECT в пакете, намеченном для отбрасывания. Вместо отбрасывания пакета маршрутизатор устанавливает код CE в заголовке IP и пересылает пакет.
- Получатель принимает пакет с кодом CE и устанавливает флаг ECN-Echo в следующем пакете TCP ACK, передаваемом отправителю.
- Отправитель получает пакет TCP ACK с флагом ECN-Echo и реагирует на перегрузку как в случае отбрасывания пакета.
- Отправитель устанавливает флаг CWR в заголовке TCP следующего пакета, передаваемого получателю, для подтверждения приёма и реакции на флаг ECN-Echo.

Согласование использования ECN транспортными элементами TCP, а также использование флагов ECN-Echo и CWR более подробно описаны в последующих параграфах.

6.1.1 Инициализация TCP

На этапе организации соединения TCP модули TCP на стороне отправителя и получателя обмениваются информацией о своём намерении использовать ECN. После завершения этого согласования отправитель TCP устанавливает код ECT в заголовке IP пакета данных для указания сетевым устройствам возможности и желания использовать ECN для этого пакета. Этот код показывает маршрутизаторам, что они могут маркировать данный пакет кодом CE, если они хотят использовать такой метод индикации перегрузки. Если соединение TCP не хочет использовать ECN-уведомление для отдельного пакета, передающая сторона TCP устанавливает в качестве кода ECN значение not-ECT, а получатель TCP игнорирует код CE в полученном пакете.

В дальнейшем обсуждении будем обозначать инициирующий соединение хост, как A, а отвечающий - B. Будем называть пакет SYN с флагами ECE и CWR «SYN-пакетом организации ECN[1]», а пакет SYN, в котором флаг ECE или CWR сброшен - «пакетом SYN без организации ECN». Аналогично будем называть пакет SYN-ACK с флагом ECE, но без флага CWR «SYN-ACK-пакетом организации ECN[2]», а пакет SYN-ACK с другой комбинацией флагов ECE и CWR - пакетом «SYN-ACK без организации ECN».

Прежде, чем соединение TCP сможет использовать ECN, хост A передаёт SYN-пакет организации ECN, а хост B передаёт пакет SYN-ACK с организацией ECN. Для пакета SYN установка обоих флагов ECE и CWR в SYN-пакете с организацией ECN определяется, как индикация поддержки ECN на передающей стороне TCP, а не индикация перегрузки или отклика на неё. Говоря точнее, пакет SYN с организацией ECN показывает, что реализация TCP, передающая пакет SYN, будет принимать участие в ECN, как отправитель и получатель. В качестве получателя она будет отвечать на пакеты данных с кодом CE в заголовке IP установкой флага ECE в исходящих подтверждениях TCP ACK. В качестве отправителя она будет отвечать на входящие пакеты с флагом ECE снижением размера окна перегрузки и установкой флага CWR, когда это следует делать. Пакет SYN с организацией ECN не обязывает отправителя TCP устанавливать код ECT в каком-либо или всех пакетах, которые он может передать. Однако обязанность подобающим образом отвечать на входящие пакеты с кодом CE сохраняется даже в том случае, когда отправитель TCP позднее в сеансе TCP передаёт пакет SYN без установленных флагов ECE и CWR.

Когда хост B передаёт пакет SYN-ACK с организацией ECN, он устанавливает флаг ECE, но не устанавливает флаг CWR. Пакет SYN-ACK с организацией ECN определяется, как индикация того, что узел TCP, передавший пакет SYN-ACK, поддерживает ECN. Как и SYN, пакет SYN-ACK с организацией ECN не обязывает хост TCP устанавливать код ECT в передаваемых пакетах.

Приведённые ниже правила применяются для пакетов организации ECN в соединении TCP, определяемом стандартными правилами организации и разрыва соединений TCP.

- Если хост получил SYN-пакет с организацией ECN, он **может** передать пакет SYN-ACK с организацией ECN. В остальных случаях передача пакетов SYN-ACK в организацией ECN **недопустима**.
- Для хоста **недопустима** установка ECT в пакетах данных, пока этот хост не передал хотя бы один пакет SYN или SYN-ACK с установкой ECN и не получил хотя бы один пакет SYN или SYN-ACK с организацией ECN, не передавая пакетов SYN или SYN-ACK без организации ECN. Если хост получил хотя бы один пакет SYN или SYN-ACK без организации ECN, ему **не следует** устанавливать ECT в пакетах данных.
- Если хост установил код ECT в пакете данных, он **должен** корректно устанавливать/сбрасывать флаг CWR в заголовках TCP всех последующих пакетов данного соединения.
- Если хост передал хотя бы один пакет SYN или SYN-ACK с организацией ECN и не получал пакетов SYN или SYN-ACK без организации ECN, тогда этот хост при получении пакетов данных TCP с кодами ECT и CE в заголовке IP **должен** обрабатывать их, как задано для соединений, поддерживающих ECN.
- Хосту, не желающему использовать ECN в соединении TCP, **следует** сбрасывать оба флага ECE и CWR во всех пакетах SYN и SYN-ACK без организации ECN, которые он передаёт для индикации своего нежелания. Получатели

[1]ECN-setup SYN packet.
[2]ECN-setup SYN-ACK packet.

должны корректно обрабатывать все формы пакетов SYN и SYN-ACK без организации ECN.

- Для хоста **недопустимо** устанавливать ECT в пакетах SYN или SYN-ACK[1].

Клиент TCP переходит в состояние TIME-WAIT после получения пакета FIN-ACK и потом в состояние CLOSED по истечении тайм-аута. Многие реализации TCP создают в состоянии TIME-WAIT новое соединение, если получают в окне пакет SYN. Когда хост TCP переходит в состояние TIME-WAIT или CLOSED, ему следует игнорировать все прежние данные о согласовании ECN для этого соединения.

6.1.1.1. Промежуточные устройства

ECN добавляет 2 флага (ECN-Echo и CWR) в заголовок TCP (Рисунок 3) для инициализации. В Internet имеются некорректно работающие межсетевые экраны, устройства балансировки и системы детектирования вторжений, которые отбрасывают SYN-пакеты с организацией ECN или передают в ответ пакет RST, ошибочно воспринимая установленные в заголовке флаги, как сигнатуры сканирования портов для организации атак на службы (DoS). Часть таких устройств идентифицирована и на сайте [FIXES] приведён их список с указанием рекомендуемых производителями исправлений, если они имеются. На сайте [TBIT] перечислены некоторые web-серверы, на которые оказывает влияние такое оборудование. Эти списки могут служить предостережением о наличии описанной проблемы.

Для обеспечения отказоустойчивых соединений даже при наличии некорректно работающих устройств хост, получающий пакет RST в ответ на передачу пакета SYN с организацией ECN, **может** повторить передачу пакета SYN со сброшенными флагами CWR и ECE. Это может позволить организацию соединения TCP без использования ECN.

Хост, не получивший отклика на пакет SYN с организацией ECN в течение обычного времени ожидания для повтора SYN, **может** повторить передачу пакета SYN со сброшенными флагами CWR и ECE. Для предотвращения влияния обычной потери пакета SYN исходный хост может передать один или несколько повторных пакетов SYN с организацией ECN до начала передачи пакетов SYN со сброшенными флагами CWR и ECE.

Отметим, что в таких случаях возможен следующий сценарий:

(1) хост A передаёт пакет SYN с организацией ECN;

(2) хост B передаёт пакет SYN/ACK с организацией ECN, который отбрасывается или задерживается;

(3) хост A передаёт пакет SYN без организации ECN;

(4) хост B передаёт пакет SYN/ACK без организации ECN.

Отметим, что в этом случае, следуя описанной выше процедуре, хосты A и B не могут устанавливать бит ECT в пакетах данных. Более того, важным следствием правил организации и использования ECN, приведённых в параграфе 6.1.1, является то, что хосту в таких случаях запрещено принимать пакеты данных с ECT, поскольку это неявно говорит об отсутствии поддержки ECN на другом хосте.

Если не указано иное в Experimental RFC потока документов IETF, промежуточным устройствам **не следует** отбрасывать пакеты управления TCP и повторно передаваемые пакеты TCP лишь по потому, что поле ECN в заголовке IP не содержит Not-ECT. Исключением из этого правила может быть реакция на атаку, использующую коды ECN, отличные от Not-ECT. Например, в качестве части отклика на атаку может быть приемлемо отбрасывание пакетов TCP SYN с маркировкой ECT с большей вероятностью, нежели пакетов TCP SYN с маркером Not-ECT. Такие исключения с отбрасыванием пакетов управления TCP и повторно передаваемых пакетов TCP в ответ на атаку **недопустимо** применять при отсутствии атак и **следует** разрешать лишь в том случае, когда ясно, что использование ECN способствует атаке.[2]

6.1.1.2. Отказоустойчивая инициализация TCP с возвратом битов резервного поля

Возникает вопрос, почему для пакетов SYN используется два связанных с ECN флага в резервном поле заголовка TCP, тогда как в ответном пакете SYN-ACK устанавливается только один связанный с ECN флаг. Такая асимметрия нужна для отказоустойчивого согласования поддержки ECN с некоторыми имеющимися реализациями TCP. Существует по меньшей мере одна некорректно работающая реализация TCP, в которой получатели устанавливают в поле Reserved заголовка TCP пакетов ACK (и, следовательно, SYN-ACK) просто отражение поля Reserved из заголовка TCP принятого пакета данных. Поскольку в пакете TCP SYN для индикации поддержки ECN устанавливаются флаги ECN-Echo и CWR, а в пакетах SYN-ACK - только флаг ECN-Echo, передающая сторона TCP корректно интерпретирует отражение получателем своих флагов, как индикацию отсутствия поддержки ECN на приёмной стороне. Передающая сторона TCP в результате не вводится в заблуждение некорректными реализациями TCP, передающими пакеты SYN-ACK с отражением поля Reserved из принятого пакета SYN.

6.1.2. Отправитель TCP

Для соединений TCP, использующих ECN, новые пакеты данных передаются с кодом ECT в заголовке IP. Когда отправителю нужен только один код ECT для всех пакетов, передаваемых в соединении TCP, **следует** использовать ECT(0). Если отправитель получает пакет ECE ACK (т. е., пакет ACK с флагом ECN-Echo в заголовке TCP), отправитель узнает о том, что в сети на пути к получателю имеется перегрузка. Индикацию перегрузки следует трактовать просто как потерю в результате перегрузки для TCP без поддержки ECN[1]. Т. е. отправитель TCP снижает вдвое размер окна перегрузки cwnd и уменьшает порог замедленного старта ssthresh. Передающему модулю TCP **не следуе**т увеличивать окно перегрузки в ответ на получение пакета ECN-Echo ACK.

TCP не следует реагировать на индикацию перегрузки более одного раза в каждом окне данных (или более одного раза за период кругового обхода). Т. е. окно перегрузки у отправителя TCP следует уменьшать однократно в ответ на серию отброшенных или помеченных CE пакетов из одного окна данных. Кроме того, отправителю TCP не следует уменьшать значение порога ssthresh, если оно уже было снижено в последний период кругового обхода. Однако отбрасывание повторно переданных пакетов интерпретируется отправителем TCP, как новый факт перегрузки.

[1] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*
[2] Этот текст отсутствует в оригинальном документе и добавлен в перевод в соответствии с RFC 8311. *Прим. перев.*

После того, как отправитель TCP уменьшает окно перегрузки в ответ на пакет CE, входящие подтверждения, которые продолжают приходить, могут влиять на передачу пакетов, дозволенных уменьшенным окном перегрузки. Если окно перегрузки содержит только 1 MSS (максимальный размер сегмента) и передающий модуль TCP получает пакет ECE ACK, передающему TCP следует, в принципе, продолжать уменьшение окна перегрузки вдвое. Однако размер этого окна ограничен снизу значением в 1 MSS. Если передающий модуль TCP будет продолжать передачу, используя окно перегрузки размером 1 MSS, это приведёт к передаче одного пакета за период кругового обхода. Требуется дальнейшее снижение скорости передачи TCP в ответ на получение пакета ECN-Echo при окне перегрузки размером 1 MSS. Мы используем таймер повтора передачи в качестве меры снижения скорости в таких ситуациях, поэтому передающий модуль TCP **должен** сбрасывать таймер повтора при получении пакета ECN-Echo в случае единичного размера окна перегрузки. Передающий модуль TCP в результате сможет передать новый пакет только после завершения отсчёта таймера повтора.

Когда поддерживающий ECN отправитель TCP снижает по какой-либо причине (тайм-аут повтора, ускоренный повтор - Fast Retransmit или отклик на ECN Notification) размер окна перегрузки, он устанавливает флаг CWR в заголовке TCP первого пакета, передаваемого после сокращения окна. Если пакет данных отбрасывается в сети, передающий модуль TCP будет снова уменьшать окно перегрузки и повторять передачу отброшенного пакета.

Мы гарантируем, что информация о сокращении окна перегрузки (флаг CWR) надёжно доставляется получателю TCP. Это основано на том, что при отбрасывании нового пакета данных с флагом CWR отправитель TCP будет повторно сокращать окно перегрузки и передавать новый пакет данных с установленным флагом CWR. Таким образом, бит CWR в заголовке TCP **не следует** устанавливать для повторно передаваемых пакетов.

Когда отправитель TCP готов установить флаг CWR после снижения размера окна перегрузки, ему **следует** устанавливать этот флаг только в первом передаваемом после сокращения окна пакете данных.

В работе [Floyd94] обсуждается реакция TCP на ECN более подробно. В работе [Floyd98] рассматривается тест с использованием эмулятора ns, который иллюстрирует множество сценариев ECN, включая ECN, за которым следует другой ECN, Fast Retransmit или Retransmit Timeout, Retransmit Timeout или Fast Retransmit, за которым следует ECN, окно перегрузки в один пакет, за которым следует ECN.

TCP следует существующим алгоритмам передачи пакетов данных в ответ на приём пакета ACK, дубликаты подтверждений или тайм-аут повтора [RFC2581]. TCP также следует обычным процедурам увеличения размера окна перегрузки при получении пакетов ACK без флага ECN-Echo [RFC2581].

6.1.3. Получатель TCP

Когда TCP принимает пакет данных CE, приёмный модуль TCP устанавливает флаг ECN-Echo в заголовке TCP следующего пакета ACK. Если на приёмной стороне уже есть ожидающий пакет ACK (как в современных реализациях TCP с задержкой подтверждений, передающих пакет ACK по прибытии каждого второго пакета данных), тогда флаг ECN-Echo устанавливается в пакете ACK, если код CE был установлен для любого из подтверждаемых пакетов данных. Т. е. при наличии в любом из подтверждаемых пакетов маркера CE, возвращаемый пакет ACK будет иметь флаг ECN-Echo.

Для обеспечения устойчивости к отбрасыванию пакетов ACK с флагом ECN-Echo, получатель TCP устанавливает этот флаг в передаваемых позднее пакетах ACK. Прекращение передачи флага ECN-Echo получатель TCP инициирует при получении флага CWR в пакете данных от передающей стороны TCP.

После того, как получатель TCP передаёт пакет ACK с установленным флагом ECN-Echo, он продолжает устанавливать этот флаг во всех передаваемых пакетах ACK (подтверждающих как пакеты данных с маркером CE, так и пакеты без маркера), пока не получит пакет с флагом CWR. После получения пакета CWR подтверждения для последующих пакетов без маркера CE передаются без флага ECN-Echo. Если получатель данных принимает другой пакет CE, он снова начинает передавать пакеты ACK с флагом ECN-Echo. Хотя приём пакета CWR не гарантирует получение отправителем пакета с флагом ECN-Echo, это событие говорит о том, что отправитель уменьшил размер окна перегрузки в какой-то момент «после» передачи пакета данных, для которого был установлен маркер CE.

Выше уже было отмечено, что отправитель TCP не должен снижать размер окна перегрузки более одного раза в окне данных. Требуются некоторые меры по предотвращению многократного снижения размера окна, когда окно данных включает как отброшенные пакеты, так и пакеты с маркером CE. Этот вопрос рассматривается в работе [Floyd98].

6.1.4. Перегрузка на пути пакета ACK

В имеющихся реализациях механизмов контроля перегрузки TCP чистые пакеты подтверждения (т. е. пакеты, содержащие только подтверждение без дополнительных данных) **должны** передаваться с кодом not-ECT[1]. Современные получатели TCP не имеют механизмов снижения трафика на пути пакетов ACK в ответ на индикацию перегрузки. Механизмы отклика на перегрузку в пути доставки пакетов ACK являются предметом современных и будущих исследований (одним из возможных вариантов может служить снижение отправителем размера окна перегрузки при получении чистого пакета ACK с кодом CE). Для современных реализаций TCP отбрасывание одного пакета ACK в общем случае оказывает пренебрежимо малое влияние на скорость передачи TCP.

6.1.5. Повторно переданные пакеты TCP

Этот документ указывает, что поддерживающим ECN реализациям TCP **недопустимо** устанавливать код ECT(0) или ECT(1) в заголовке IP для повторно передаваемых пакетов данных, а получателю данных TCP **следует** игнорировать поле ECN в прибывающих пакетах данных, которые находятся за пределами текущего окна получателя[1]. Это сделано для обеспечения более эффективной защиты от атак на службы, а также устойчивости к индикации перегрузки ECN в пакетах, которые позднее отбрасываются в сети.

Во-первых, отметим, что если отправитель TCP будет устанавливать код ECT в повторно передаваемых пакетах, тогда в случаях отбрасывания без необходимости переданных повторно пакетов в сети конечные точки никогда не получат индикации перегрузки от маршрутизатора, устанавливающего код CE. Таким образом, установка кода ECT в повторно

[1] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*

передаваемых пакетах несовместима с отказоустойчивой доставкой индикации перегрузки даже при последующем отбрасывании пакетов в сети.

Кроме того, атакующий может подменить IP-адрес отправителя TCP и передать пакеты данных с произвольными порядковыми номерами и установленным кодом CE в заголовке IP. При получении такого обманного пакета данных приёмный модуль TCP будет считать, что данные не относятся к текущему окну приёма и возвращать дубликаты подтверждений. Мы определяем пакеты, находящиеся за пределами окна на стороне получателя TCP, как пакеты, лежащие за пределами текущего окна получателя. При получении такого пакета принимающий модуль TCP решает, следует ли считать код CE в заголовке пакета корректной индикацией перегрузки и, следовательно, нужно ли возвращать индикацию ECN-Echo отправителю данных TCP. Если получатель TCP игнорирует код CE в пакете данных за пределами окна, отправитель TCP не получит (возможно легитимной) индикации перегрузки в сети, что приведёт к нарушению сквозного контроля перегрузки. С другой стороны, если получатель данных TCP воспримет индикацию CE из пакетов за пределами окна и покажет перегрузку отправителю данных TCP, вредоносный узел, создавший обманные пакеты, лежащие за пределами окна, сможет успешно атаковать соединение TCP, заставляя отправителя без необходимости снижать (вдвое) размер окна перегрузки. Для предотвращения таких атак на службы мы указываем, что для легитимного отправителя TCP **недопустима** установка кода ECT в передаваемых повторно пакетах данных, а получателю TCP **следует** игнорировать код CE в пакетах, лежащих за пределами окна.

Негативным эффектом отказа от установки ECT(0) или ECT(1) в повторно передаваемых пакетах является то, что в результате этого отключается ECN-защита для передаваемых повторно пакетов. Однако для поддерживающих ECN соединений TCP в полностью поддерживающей ECN среде со средним уровнем перегрузки пакеты будут редко отбрасываться в результате перегрузки, поэтому повторно передаваемые пакеты в большинстве случаев не будут теряться. Повторная передача пакета говорит о том, что потери уже наблюдались (в результате повреждения или перегрузки) и ECN все равно не поможет предотвратить возможную потерю.

Отметим, что если маршрутизатор устанавливает код CE для совместимого с ECN пакета данных в соединении TCP, для этого соединения гарантируется получение индикации перегрузки в том же окне данных даже при отбрасывании или нарушении порядка доставки пакетов в сети. Мы рассматриваем здесь два случая - когда пакеты позднее передаются повторно и когда они повторно не передаются. В первом случае при отбрасывании или задержке пакета и повторе передачи этот повтор является результатом использования алгоритма Fast Retransmit или Retransmit Timeout для данного пакета или какого-то из его предшественников в том же окне данных. Поскольку отправитель уже передал пакет повторно, мы знаем, что он уже откликнулся на индикацию перегрузки для какого-то из пакетов в том же окне данных, что и исходный пакет. Таким образом, даже при отбрасывании или задержке первого пакета в сети, если у него был установлен код CE и пакет был потом проигнорирован получателем, как выходящий за пределы окна, это не создаёт проблем, поскольку отправитель уже отреагировал на индикацию перегрузки в этом окне данных. Во втором случае, если пакет никогда не передаётся повторно, такой пакет будет единственной копией соответствующих данных на стороне получателя и, следовательно, попадёт к получателю в пределах окна данных, независимо от задержки или нарушения порядка в сети. В этом случае код CE устанавливается для пакета в сети и будет трактоваться получателем, как корректная индикация перегрузки.

6.1.6. Пробы окна TCP

Когда получатель TCP анонсирует окно нулевого размера, отправитель TCP передаёт зонды для определения возможности увеличения окна на приёмной стороне. Пакеты проб не содержат пользовательских данных за исключением однобайтового порядкового номера. Если зонд отбрасывается в сети, получатель не видит такой потери. Следовательно, отправителю TCP **недопустимо** устанавливать код ECT или флаг CWR в пробных пакетах[1]. Однако, благодаря порядковым номерам в пробных пакетах, эти пакеты невозможно просто подменить в DoS-атаке. Следовательно, если пробный пакет приходит с кодом CE, получателю **следует** реагировать на индикацию ECN.

7. Неподатливость конечных узлов

В этом разделе рассматривается опасность ECN для неподатливых[2] конечных узлов (т. е., узлов, которые устанавливают код ECT в передаваемых пакетах, но не реагируют на получение пакетов с кодом CE). Мы понимаем, что добавление ECN в архитектуру IP не повышает сколь-нибудь существенно общий уровень уязвимости архитектуры со стороны невосприимчивых потоков.

Даже для сред, не поддерживающих ECN, следует серьёзно рассматривать возможность нарушений, которые могут быть вызваны неподатливыми или невосприимчивыми потоками (т. е. потоками, которые не отвечают на индикацию перегрузки снижением скорости доставки через загруженный канал). Например, конечная точка может «отключить контроль перегрузки», не снижая размер окна перегрузки в ответ на отбрасывание пакетов. Эта проблема важна для современного состояния Internet. Ясно, что в маршрутизаторах нужно реализовать механизмы обнаружения и дифференцированной трактовки пакетов из неподатливых потоков [RFC2309, FF99]. Предполагается также, что такие методы, как сквозное планирование на уровне потока и изоляция потоков друг от друга, дифференцированные услуги или сквозное резервирование, могут устранить некоторые из наиболее разрушительных эффектов от невосприимчивых потоков.

Может показаться, что отбрасывание пакетов само по себе является средством сдерживания неподатливости, а использование ECN блокирует эту возможность. Мы утверждаем в ответ на это, что (1) поддерживающие ECN маршрутизаторы сохраняют возможность отбрасывания пакетов при сильной перегрузке и (2) даже в случаях сильной перегрузки отбрасывание пакетов не является сдерживающим неподатливость фактором. Во-первых, поддерживающие ECN маршрутизаторы будут только маркировать пакеты (вместо их отбрасывания), пока частота маркирования достаточно мала. В периоды, когда средний размер очереди превышает верхний порог и, следовательно, потенциальная скорость маркировки пакетов будет велика, мы рекомендуем маршрутизатору отбрасывать пакеты вместо установки кода CE в заголовках.

В периоды с низкой и средней скоростью маркировки пакетов, когда поддержка ECN реализована, будет возникать некий негативный эффект для невосприимчивых потоков в виде отбрасывания пакетов вместо их маркировки. Например, для нечувствительных к задержкам потоков, использующих гарантированную доставку, при отбрасывании

[1] RFC 8311 добавляет в конце предложения слова: «если иное не указано в Experimental RFC потока документов IETF». *Прим. перев.*
[2] Non-compliant.

пакетов может наблюдаться увеличение скорости вместо её снижения. Аналогично для чувствительных к задержкам потоков без гарантированной доставки может возрастать использование FEC в ответ на рост частоты отбрасывания пакетов, приводящее скорее к росту, чем к снижению скорости передачи. По тем же причинам мы не верим, что отбрасывание пакетов, само по себе, является эффективным средством сдерживания неподатливости даже в средах с высокой частотой отбрасывания пакетов, когда вероятность отбрасывания делится между всеми потоками.

Было предложено несколько методов идентификации и ограничения неподатливых и невосприимчивых потоков. Добавление ECN в сетевую среду никак не усложнит разработку и развёртывание таких механизмов. Во всяком случае, добавление ECN в архитектуру будет существенно упрощать работу по идентификации невосприимчивых потоков. Например, в среде с поддержкой ECN маршрутизаторы не ограничиваются информацией о том, что пакет был отброшен или получил код на данном маршрутизаторе - в таких средах маршрутизаторы могут также отмечать прибытие пакетов с кодом CE, показывающих перегрузку, встреченную пакетом на своём пути раньше.

8. Неподатливость в сети

В этом разделе рассматриваются ситуации, когда маршрутизатор (возможно с враждебными целями) меняет какие-либо биты поля ECN. Отметим, что в IPv4 заголовок IP защищён от битовых ошибок контрольной суммой, но такая защита отсутствует в IPv6. Таким образом, для IPv6 поле ECN может быть непреднамеренно изменено в результате битовых ошибок в каналах или маршрутизаторах и такое изменение не будет обнаружено по причине отсутствия контрольной суммы.

Манипулируя битами поля ECN, враждебный (или некорректно работающий) маршрутизатор может приводить к возникновению ряда негативных эффектов, включая ложную информацию о перегрузках, запрет поддержки ECN для отдельных пакетов, удаление ECN-индикации перегрузки или ложная индикация поддержки ECN. В разделе 18 результаты изменения полей ECN рассмотрены подробно и систематизированы. Важным критерием учёта последствий такого изменения является то, что оно может вести к ухудшению поведения по целому ряду параметров (пропускная способность, задержка, беспристрастность, функциональность) даже по сравнению с отбрасыванием пакетов в маршрутизаторе.

В двух первых случаях (ложная информация о перегрузке и запрет ECN для отдельного пакета) ситуация не хуже, чем при простом отбрасывании пакета маршрутизатором. С точки зрения системы контроля перегрузки установка кода CE в отсутствии перегрузки неподатливым маршрутизатором будет не хуже, чем неоправданное отбрасывание им пакета. За счёт «удаления» кода ECT в пакете, который позднее будет отброшен в сети, действия маршрутизатора могут привести впоследствии к неоправданному отбрасыванию пакета в сети.

Однако, как отмечено в разделе 18, маршрутизатор, способный удалять ECN-индикацию перегрузки или ложно указывать поддержку ECN, может наносить больший ущерб, нежели простое отбрасывание пакета. Враждебный или некорректно работающий маршрутизатор, который «удаляет» код CE в приходящих к нему пакетах CE, будет удалять индикацию перегрузки, приходящую от получателей (нисходящих, по отношению к маршрутизатору). Это может приводить к отказам системы контроля перегрузки для потоков и дальнейшему росту перегрузки в сети, ведущему к нарастающему отбрасыванию последующих пакетов потока по мере роста среднего размера очереди на перегруженном шлюзе.

В разделе 19 рассмотрены возможные результаты нарушения сквозного контроля перегрузки за счёт ложной индикации поддержки ECN или удаления индикации перегрузки (кода CE). Показано, что в результате нарушения основанного на ECN контроля перегрузки может теряться беспристрастность промежуточных устройств, но это влияние явно не будет хуже, чем при потере конечными узлами контроля перегрузки на основе отбрасывания пакетов или ECN.

8.1. Осложнения, связанные с расщеплением пути

Если маршрутизатор или другой элемент сети имеет доступ ко всем пакетам потока, это устройство не может нанести путём изменения поля ECN большего вреда, чем простое отбрасывание пакетов из потока. Однако в некоторых случаях враждебный или некорректно настроенный маршрутизатор может получить доступ лишь к части пакетов потока. Возникает вопрос - сможет ли такой маршрутизатор путём изменения поля ECN в данном подмножестве пакетов нанести больший вред, нежели путём простого отбрасывания этого набора пакетов?

Этот вопрос также подробно рассматривается в разделе 18 и делается следующее заключение - верно, что злоумышленник, имеющий доступ лишь к части пакетов, может путём нарушения контроля перегрузки на основе ECN оказаться способным свести на нет преимущества ECN для других пакетов потока. Такое поведение нежелательно, но не может служить достаточным основанием для отказа от ECN.

9. Инкапсулированные пакеты

9.1. Пакеты IP, инкапсулированные в IP

Инкапсуляция заголовков пакетов IP в туннели используется во многих случаях, включая IPsec и IP-in-IP [RFC2003]. В этом разделе рассматриваются вопросы, связанные со взаимодействием между ECN и IP, а также описаны два дополнительных решения. Это обсуждение дополняется рассмотрением в RFC 2983 взаимодействия между дифференцированным обслуживанием (DifServ) и различными формами туннелей IP [RFC 2983], а также использования в DifServ оставшихся 6 битов заголовка IP, не занятых ECN (Рисунок 2).

Некоторые режимы туннелей IP основаны на добавлении нового «внешнего» заголовка IP, который инкапсулирует исходный или «внутренний» заголовок IP и связанный с ним пакет. Во многих случаях внешний заголовок может добавляться и удаляться на промежуточных точках соединения, что позволяет создавать туннели без участия конечных точек в их организации. Будем называть туннели, которые задают отбрасывание внешнего заголовка на выходе «простыми туннелями».

ECN использует поле в заголовке IP для сигнализации между маршрутизаторами и конечными точками соединений, взаимодействуя с туннелями IP на базе трактовки поля ECN в заголовке IP. В простых туннелях IP октет, содержащий поле ECN, копируется или отображается из внутреннего заголовка IP во внешний на входе туннеля IP, а на выходе из туннеля внешняя копия поля просто отбрасывается. Если внешний заголовок отбрасывается без обработки поля ECN,

а поддерживающий ECN маршрутизатор установил код CE (перегрузка) в заголовке простого туннеля IP, эта индикация будет отброшена на выходе из туннеля и конечный узел не узнает о перегрузке.

Таким образом, использование ECN с простыми туннелями IP приведёт к тому, что маршрутизаторы будут пытаться сигнализировать о перегрузке с помощью внешнего заголовка, но эта индикация не будет получена в результате отбрасывания поля при декапсуляции на выходе из туннеля. Эта проблема возникает при использовании ECN с IPsec в туннельном режиме и RFC 2481 рекомендует отказаться от использования ECN со старыми простыми туннелями IPsec во избежание упомянутого негативного эффекта и его последствий. По мере распространения ECN простые туннели должны будут измениться для обеспечения передачи поддерживающего ECN трафика. Если трафик ECN передаётся через простой туннель и сталкивается с перегрузкой на поддерживающем ECN маршрутизаторе, это может привести к отбрасыванию последующих пакетов в зависимости от роста среднего размера очередей на перегруженном маршрутизаторе, как было отмечено в разделе 8.

С точки зрения безопасности использование ECN во внешнем заголовке туннеля IP может вызывать проблемы, поскольку злоумышленник может изменять информацию ECN, передаваемую между конечными точками туннеля. На основе результатов анализа в разделах 18 и 19 с учётом отмеченной проблемы и связанного с ней риска предлагается включать поддержку ECN, как опцию для туннелей IP, чтобы при настойке туннеля можно было задать использование ECN во внешнем заголовке туннеля или отказ от этого. Таким образом, в средах и протоколах с туннелированием, где риск в результате использования ECN больше, чем обеспечиваемые преимущества, туннель может просто не использовать ECN в своём внешнем заголовке. В этом случае единственным способом индикации перегрузки маршрутизатора становится отбрасывание пакетов.

В результате возникают два жизнеспособных варианта поведения поддерживающих ECN соединений через туннели IP, включая IPsec:

- ограниченная функциональность, когда ECN сохраняется во внутреннем заголовке, но отбрасывается из внешнего и единственным механизмом сигнализации о перегрузке является отбрасывание пакетов;
- полная функциональность с поддержкой ECN во внутренних и внешних заголовках и передачей сигналов о перегрузке от узлов внутри туннеля конечным точкам.

Поддержка этих опций требует различного объёма изменений, вносимых в обработку заголовков IP на входах и выходах туннелей. Для поддержки ограниченной функциональности будет достаточно небольшого набора изменений, обеспечивающего устранение всех несовместимостей между ECN и туннелями IP.

Одной из задач данного документа является определение компромисса между полнофункциональным и ограниченным вариантами. Обсуждение возможных эффектов враждебного изменения поля ECN приведено в разделах 18 и 19.

9.1.1. Опции ограниченной и полной функциональности

Опция ограниченной функциональности для инкапсуляции ECN в туннели IP обеспечивается путём установки кода not-ECT во внешнем заголовке, независимо от значения поля ECN во внутреннем заголовке. В этом случае поле ECN внутреннего заголовка не меняется при декапсуляции. Недостатком этого варианта является отсутствие поддержки ECN для части пути, использующей туннелирование IP, даже в тех случаях, когда инкапсулированный пакет (от исходного отправителя TCP) поддерживает ECN. Если инкапсулированный пакет приходит на перегруженный маршрутизатор, который поддерживает ECN, и маршрутизатор принимает решение об отбрасывании или маркировке пакета для индикации перегрузки конечному узлу, маршрутизатору не разрешается устанавливать код CE в заголовке пакета и вместо этого приходится отбрасывать пакет.

Полнофункциональная инкапсуляция ECN использует копирование кода ECN из внутреннего заголовка во внешний, если внутренний заголовок not-ECT или ECT, и установки для внешнего заголовка ECT(0), если код ECN во внутреннем заголовке имеет значение CE. При декапсуляции код CE, если он установлен, копируется из внешнего заголовка во внутренний. В остальных случаях код ECN во внутреннем заголовке остаётся неизменным. Т. е. при полной поддержке ECN процессы инкапсуляции и декапсуляции включают ряд операций. На входе туннеля полнофункциональный вариант устанавливает код ECN во внешнем заголовке. Если код ECN во внутреннем заголовке имеет значение not-ECT или ECT, этот код копируется во внешний заголовок. Если код ECN во внутреннем заголовке имеет значение CE, для кода ECN во внешнем заголовке устанавливается значение ECT(0). При декапсуляции на выходе туннеля полнофункциональный вариант устанавливает код CE во внутреннем заголовке, если этот код был установлен во внешнем. В остальных случаях это поле внутреннего заголовка не меняется.

С полнофункциональной поддержкой поток может использовать преимущества ECN на тех участках пути, где может встречаться туннелирование IP. Недостаток полнофункционального варианта с точки зрения безопасности заключается в том, что туннель IP не может защитить поток от некоторых изменений поля ECN в заголовке IP. Потенциальные опасности такой замены битов ECN подробно рассмотрены в разделах 18 и 19.

(1) Туннель IP **должен** изменять обработку октета поля DS в конечных точках туннеля IP путём реализации ограниченной или полной функциональности.

(2) В дополнение к этому туннель IP **может** разрешить своим конечным точкам согласование выбора между полной и ограниченной функциональностью для ECN в туннеле.

Минимальным требованием для использования ECN с туннелями IP является поддержка ограниченной функциональности, при которой ECN не применяется для внешних заголовков туннеля. Для полной поддержки ECN требуется полнофункциональная опция. Если нет дополнительных механизмов согласования между концами туннеля выбора полнофункциональной или ограниченной опции, может использоваться предварительное соглашение о выборе той или иной опции поддержки ECN для туннеля.

Все туннели IP **должны** поддерживать опцию ограниченной функциональности и **следует** поддерживать также полнофункциональную опцию.

В дополнение к этому **рекомендуется** отбрасывать пакеты с кодом CE во внешнем заголовке, если они приходят на выход туннеля с ограниченной функциональностью или туннель поддерживает полнофункциональную опцию, но во внутреннем заголовке пакета установлен код not-ECT. Это нужно для совместимости с более ранними версиями и

предотвращения несанкционированного изменения поля ECN. Более подробно этот вопрос рассматривается в следующем параграфе.

9.1.2. Изменения для поля ECN внутри туннелей IP

Наличие копии поля ECN во внутреннем заголовке туннелируемого пакета IP обеспечивает возможность обнаружения несанкционированного изменения поля ECN во внешнем заголовке. Для реализаций, соответствующих данному документу, при сравнении полей ECT во внутреннем и внешнем заголовке, нужно принимать во внимание два случая:

- если туннель IP использует полнофункциональную опцию, то код not-ECT следует устанавливать во внешнем заголовке тогда и только тогда, когда он установлен во внутреннем заголовке;
- если туннель использует опцию ограниченной функциональности, во внешнем заголовке следует устанавливать код not-ECT.

Получение пакета, не удовлетворяющего соответствующему условию, может быть поводом для беспокойства.

Рассмотрим случай, когда входная точка туннеля IP не была обновлена в соответствии с требованиями этого документа, а выходная точка обновлена для поддержки ECN. В этом случае туннель IP не настроен явно на полнофункциональную поддержку ECN. Однако поведение входной точки идентично поведению входной точки туннеля с полнофункциональной поддержкой. Если пакет из поддерживающего ECN соединения использует такой туннель, на входе туннеля может быть установлен код ECT во внешнем заголовке. Перегрузка в туннеле может привести к установке поддерживающим ECN маршрутизатором кода CE во внешнем заголовке. Поскольку туннель явно не настроен на поддержку полнофункциональной опции, выходная точка ожидает присутствия во внешнем заголовке кода not-ECT. Когда поддерживающая ECN выходная точка туннеля получает пакет с кодом ECT или CE во внешнем заголовке туннеля, который не настроен на поддержку полнофункциональной опции, этот пакет следует обрабатывать с учётом наличия кода CE. **Рекомендуется** для туннелей, не настроенных на поддержку полнофункциональной опции, отбрасывать пакет на выходе, если код CE установлен во внешнем заголовке, но отсутствует во внутреннем. Остальные пакеты следует пересылать.

Туннель IP не может обеспечить защиту от удаления индикации перегрузки путём замены кода ECN с CE на ECT. Удаление индикации перегрузки может влиять на сеть и другие потоки так, как невозможно было бы повлиять в отсутствие ECN. Важно отметить, что удалить можно лишь те индикаторы перегрузки, которые были установлены узлами внутри туннеля. Копия поля ECN во внутреннем заголовке сохраняет индикацию перегрузки от узлов, расположенных до входа в туннель (если внутренний заголовок также не был изменён). Если удаление индикации перегрузки связано с риском, превышающим преимущества от контроля перегрузки с помощью ECN, туннель следует настроить на поддержку ограниченной функциональности.

9.2. Туннели IPsec

IPsec поддерживает защищённую связь через потенциально небезопасные компоненты сети, такие как промежуточные маршрутизаторы. Протоколы IPsec поддерживают два режима работы (туннельный и транспортный), которые обеспечивают выполнение широкого спектра требований защиты и работу в различных средах. Заголовок протокола защиты в транспортном режиме помещается между заголовком IP (IPv4 или IPv6) и заголовком протокола вышележащего уровня (например, TCP), следовательно в транспортном режиме обеспечивается сквозная защита (между конечными точками). Туннельный режим IPsec основан на добавлении нового внешнего заголовка IP, который инкапсулирует исходный заголовок и связанный с ним пакет. Заголовки защиты в туннельном режиме вставляются между внешним и внутренним заголовками IP. В отличие от транспортного режима внешний заголовок IP и заголовки защиты туннельного режима могут удаляться и добавляться на промежуточных узлах пути, позволяя шлюзам безопасности защищать уязвимые части соединения без необходимости включения конечных точек в обеспечение защиты. Важным свойством туннельного режима в соответствии с исходной спецификацией является отбрасывание внешнего заголовка на выходе туннеля, в результате чего угрозы, связанные с изменением заголовков IP не распространяются дальше конечной точки туннеля. Дополнительную информацию о IPsec можно найти в [RFC2401].

Протокол IPsec, изначально определённый в [ESP, AH], требует, чтобы поле ECN внутреннего заголовка не менялось при декапсуляции IPsec в выходном узле туннеля - это требование противоречит возможностям полнофункциональной поддержки ECN. В то же время это обеспечивает защиту от враждебного изменения поля ECN с целью организации атак через конечные точки туннелей IPsec, поскольку в конечной точке все изменения теряются.

В принципе, при разрешении использовать функциональность ECN во внешнем заголовке туннеля IPsec возникают проблемы безопасности, связанные с тем, что враждебные стороны могут искажать информацию, распространяющуюся за пределы конечной точки туннеля. На основе анализа (разделы 18 и 19) этих опасностей мы рекомендуем обеспечивать конфигурационную поддержку для изменений IPsec, позволяющих разрешить конфликт с ECN.

В частности, в туннельном режиме туннель IPsec **должен** поддерживать опцию ограниченной функциональности, кратко рассмотренную в параграфе 9.1.1, и **следует** также поддерживать полнофункциональную опцию, описанную в параграфе 9.1.1.

Это делает разрешение на использование функциональности ECN во внешнем заголовке туннеля IPsec настраиваемой частью соответствующей защищённой связи IPsec (SA[1]), которая может быть отключена в тех случаях, когда риск превышает достигаемые преимущества. В результате администратору безопасности IPsec предоставляется два варианта поведения поддерживающих ECN соединений в туннелях IPsec - с полной и ограниченной функциональностью, как описано выше.

В дополнение к сказанному в этом документе также даётся спецификация согласования используемой функциональности ECN между конечными точками туннеля IPsec с учётом политики безопасности. Возможность согласования ECN между конечными точками туннеля будет позволять администратору безопасности отключать поддержку ECN в ситуациях, когда возможный риск превышает преимущества от использования ECN (например, при потере уведомлений о перегрузках).

[1]Security Association.

Протокол IPsec, определённый в [ESP, AH], не включает поля ECN заголовка IP в какие-либо криптографические преобразования (в туннельном режиме внешний заголовок IP не включает поля ECN). Следовательно, изменение поля ECN любым узлом в сети не оказывает влияния на сквозную защиту IPsec, поскольку не позволяет нарушить целостность данных IPsec. В результате этого IPsec не обеспечивает никакой защиты от враждебного изменения поля ECN (например, от атак MITM[1]), поскольку такие изменения не оказывают влияния на сквозную защиту IPsec. В некоторых средах возможность изменения поля ECN без воздействия на проверку целостности IPsec позволяет создавать скрытые каналы, для предотвращения такой возможности или снижения полосы скрытого канала для туннеля IPsec следует выбирать режим ограниченной функциональности.

9.2.1. Согласование между конечными точками туннеля

В этом параграфе описаны изменения, которые требуется внести для использования ECN с туннелями IPsec, включая согласование поддержки ECN между конечными точками туннеля. Для такой поддержки в IPsec нужно внести три изменения.

- Необязательное поле SAD[2], показывающее способность процессов инкапсуляции и декапсуляции туннеля разрешать и запрещать использование ECN во внешнем заголовке IP.
- Необязательный атрибут защитной связи, разрешающий для данного поля SAD согласование между двумя конечными точками SA, поддерживающей туннельный режим.
- Изменения процессов инкапсуляции и декапсуляции, позволяющие разрешать и запрещать использование ECN во внешнем заголовке IP на основе значения поля SAD. Когда использование ECN во внешнем заголовке IP разрешено, в этом заголовке устанавливается код ECT для поддерживающих ECN соединений и уведомления о перегрузках (код CE) из таких соединений передаются во внутренний заголовок на выходе туннеля.

Если реализовано согласование опций применения ECN, **следует** также реализовать поле SAD. С другой стороны, согласование использования ECN в любом случае является **опциональным**, даже если реализация поддерживает поле SAD. Изменение обработки при инкапсуляции и декапсуляции **требуется**, но **может** быть реализовано без реализации двух других изменений в предположении, что использование ECN всегда запрещено. Полнофункциональный вариант использования ECN с туннелями IPsec включает поле SAD и полный вариант изменения процессов инкапсуляции и декапсуляции с **опциональной** поддержкой согласования. Вариант с ограниченной функциональностью включает часть изменений процессов инкапсуляции и декапсуляции, которая всегда запрещает использование ECN.

Эти изменения рассмотрены более подробно в трёх следующих параграфах.

9.2.1.1. Поле ECN Tunnel в SAD

Полнофункциональная поддержка ECN добавляет новое поле в SAD (см. [RFC2401]):

> ECN Tunnel: разрешён (allowed) или запрещён (forbidden).
>
> Показывает, разрешено ли совместимому с ECN соединению, использующему данную SA в туннельном режиме, получать индикацию ECN для перегрузок, возникающих в туннеле. Значение allowed разрешает уведомления ECN. Значение forbidden запрещает такие уведомления, вынуждая использовать для индикации перегрузки отбрасывание пакетов.
>
> [**Опционально**. Для реализаций, не поддерживающих этот атрибут в данном поле, **следует** устанавливать значение forbidden.]

Если этот атрибут реализован, спецификация SA в базе данных SPD[3] **должна** поддерживать соответствующий атрибут и этот атрибут SPD **должен** быть включён в административный интерфейс SPD (в настоящее время описан в параграфе 4.4.1 [RFC2401]).

9.2.1.2. Атрибут ECN Tunnel в SA

Определён новый атрибут IPsec SA, позволяющий поддерживать согласование использования индикации перегрузки ECN во внешнем заголовке IP для туннелей IPsec (см. [RFC2407]). Этот атрибут является **опциональным**, поддерживающим его реализациям **следует** поддерживать также поле SAD, определённое в параграфе 9.2.1.1.

Тип атрибута

Класс	*Значение*	*Тип*
ECN Tunnel	10	Basic

Значение атрибута IPsec SA, равное 10, выделено IANA для индикации согласования атрибута ECN Tunnel SA, тип этого атрибута - Basic (см. параграф 4.5 [RFC2407]). Значения класса атрибутов используются при согласовании. Дополнительная информация, включая форматы кодирования и требования по согласованию этого атрибута SA, приведена в [RFC2407, RFC2408, RFC2409].

Значения класса

ECN Tunnel

Определяет возможность использования функциональности ECN с туннельной инкапсуляцией. Значение определяет процессы инкапсуляции и декапсуляции (см. параграф 9.2.1.3).

Резерв 0

Разрешено 1

Запрещено 2

[1]A man-in-the-middle attack - атака с участием человека на пути передачи данных.
[2]Security Association Database - база данных ассоциаций защиты.
[3]Security Policy Database - база данных политики.

Значения 3 - 61439 зарезервированы IANA, значения 61440 - 65535 выделены для приватного использования.

По умолчанию следует предполагать запрет (2).

ECN Tunnel является новым атрибутом SA и, следовательно, при его использовании могут возникать проблемы непонимания этого атрибута и отказа от предложения его использовать. Для совместимости со старыми версиями новым реализациям **следует** во всех случаях также предлагать работу без атрибута ECN. RFC 2407 в настоящее время требует от отвечающей стороны отвергать все предложения с неизвестными атрибутами, предполагается, что это требование будет заменено требованием для отвечающего не выбирать предложения или преобразования с неизвестными атрибутами.

9.2.1.3. Изменения в обработке заголовков туннелей IPsec

Для полной поддержки ECN процессы инкапсуляции и декапсуляции для полей IPv4 TOS и IPv6 Traffic Class меняются по сравнению с указанными в [RFC2401], как показано ниже.

Связь внешнего заголовка с внутренним		
Поле	***Внешний заголовок на инкапсуляторе***	***Внутренний заголовок на декапсуляторе***
IPv4		
DS	Копируется из внутреннего заголовка (5)	Не меняется
ECN	Создаётся (7)	Создаётся (8)
IPv6		
DS	Копируется из внутреннего заголовка (6)	Не меняется
ECN	Создаётся (7)	Создаётся (8)

(5)(6) Если пакет незамедлительно будет вводиться в домен, для которого значение DSCP во внешнем заголовке неприемлемо, это значение **должно** быть отображено на приемлемое для домена значение [RFC 2474]. Дополнительная информация по этому вопросу содержится в [RFC 2475].

(7) Если поле ECN Tunnel в записи SAD для данной SA имеет значение allowed (разрешено) и поле ECN во внутреннем заголовке имеет значение, отличное от CE, поле ECN копируется во внешний заголовок. Если поле ECN во внутреннем заголовке имеет значение CE, в поле ECN внешнего заголовка устанавливается значение ECT(0).

(8) Если поле ECN Tunnel в записи SAD для данной SA имеет значение allowed и поле ECN во внутреннем заголовке имеет значение ECT(0) или ECT(1), а поле ECN во внешнем заголовке имеет значение CE, поле ECN из внешнего заголовка копируется во внутренний. В остальных случаях значение поля ECN во внутреннем заголовке не меняется.

(5) и (6) идентичны в соответствии с использованием в [RFC2401], хотя в [RFC2401] они различаются.

Приведённое выше описание применимо к реализациям, поддерживающим поле ECN Tunnel в SAD, такие реализации **должны** обеспечивать описанную здесь обработку вместо использования обработки октетов IPv4 TOS и IPv6 Traffic Class, описанной в [RFC2401]. Это обеспечивает полнофункциональную поддержку ECN с туннелями IPsec.

Реализации, не поддерживающие поле ECN Tunnel в SAD, **должны** обеспечивать обработку в предположении, что поле ECN Tunnel в SAD имеет значение forbidden (запрещено) для каждой SA. В этом случае обработка поля ECN сводится к двум операциям:

(7) установить для поля ECN во внешнем заголовке значение not-ECT;

(8) не менять поле ECN во внутреннем заголовке.

Такое поведение обеспечивает ограниченную поддержку ECN с туннелями IPsec.

Для совместимости с более ранними версиями пакеты с кодом CE во внешнем заголовке **следует** отбрасывать, если они приходят в SA, использующую опцию ограниченной функциональности, или при полнофункциональной поддержке, когда во внутреннем заголовке установлен код not-ECN.

9.2.2. Изменения поля ECN в туннелях IPsec

Если поле ECN неподобающим образом меняется в туннеле IPsec и эти изменения детектируются на выходе туннеля, получение пакета, не удовлетворяющего условиям для данной SA, протоколируется системой аудита. Реализация **может** создавать записи аудита с учётом некорректных пакетов для каждой SA в течение определённого периода вместо создания отдельной записи для каждого некорректного пакета. Во все записи аудита **следует** включать заголовки по крайней мере из одного некорректного пакета, но не требуется включать заголовки всех пакетов, представленных в записи.

9.2.3. Комментарии к поддержке IPsec

Приведённые здесь комментарии к двум частям этого документа были получены при просмотре документа членами рабочей группы IPsec. В этом параграфе рассматриваются полученные комментарии и даются объяснения причин того, что предложенные изменения не были включены в документ.

В первом комментарии отмечалось, что настройка конфигурации на уровне узла проще в реализации, нежели настройка на уровне SA. После серьёзного обдумывания и обсуждения исходное предложение о настройке конфигурации на уровне узла перестало казаться хорошей идеей. Причина заключается в том, что осведомлённость о ECN становится все шире в IPsec - многие знающие о ECN реализации IPsec осознают, что они взаимодействуют как с поддерживающими ECN конечными точками туннелей IPsec, так и с точками, не понимающими ESN. В такой среде при настройке конфигурации на уровне узла остаётся лишь запретить использование ECN для всех туннелей IPsec, которые не являются нужным выходом.

Во втором комментарии ряд рецензентов отметил, что согласование SA является достаточно сложным и его добавление - нетривиальная задача. Другие предлагали в качестве альтернативы использование ICMP после организации туннеля. Поддержка согласования SA в этом документе указана, как **опциональная** и останется таковой -

разработчики сами принимают решение о реализации этой опции. Авторы надеются, что приведённые здесь аргументы будут полезны в разных ситуациях. Если эти рекомендации не будут использованы на практике, они могут быть удалены на последующих этапах процесса стандартизации. Использование ICMP для согласования ECN после организации туннеля более сложно, нежели расширение согласования атрибутов SA. Некоторые туннели не принимают трафик, адресованный узлу выходной точки туннеля, поэтому пакеты ICMP нужно будет направлять в какой-то иной адрес - они будут «сканироваться» на выходе туннеля и отбрасываться здесь или у конечного получателя. Кроме того, ICMP не обеспечивает гарантированной доставки и поэтому существует возможность отбрасывания пакетов ICMP, для учёта которой потребуется создать дополнительный механизм подтверждений и повторов передачи. **Опциональное** расширение существующего механизма согласования SA представляется более эффективным решением.

9.3. Пакеты IP, инкапсулированные в пакеты других протоколов

При инкапсуляции пакетов IP в пакеты других протоколов возникают иные вопросы, связанные с ECN. Такие ситуации возникают для MPLS [MPLS], GRE [GRE], L2TP [L2TP] и PPTP [PPTP]. Для этих протоколов не возникает конфликта с ECN - дело в том, что ECN просто не может использоваться внутри туннеля, пока код ECN не может быть включён в заголовок инкапсулирующего протокола. Выполнены начальные разработки по встраиванию ECN в MPLS, а предложения по встраиванию ECN в GRE, L2TP или PPTP будут рассматриваться по мере их появления.

10. Проблемы, создаваемые «карательными» устройствами

Возможно, что в сети будут устройства мониторинга и реализации политики (точнее было бы назвать их «карательными» устройствами), которые будут просматривать потоки на предмет адекватной реакции на перегрузки и отбрасывать в первую очередь пакеты из тех потоков, которые не используют подобающих процедур сквозного контроля перегрузки.

Для «карательных» устройств, которые обнаруживают, что отдельный поток или группа потоков не обеспечивает сквозного контроля перегрузки, следует сначала менять для таких потоков маркировку пакетов на отбрасывание и только после этого принимать дополнительные меры по ограничению полосы, доступной для потока. Таким образом, сначала маршрутизатор будет отбрасывать пакеты, которые при иных условиях он бы пометил кодом CE. Сюда включается отбрасывание пакетов, поступающих из совместимых с ECN потоков и уже имеющих код CE. В этом случае любая перегрузка, которую маршрутизатор видит для потока, будет видимой и для конечных узлов даже при наличии враждебных или некорректно работающих маршрутизаторов на пути передачи. Если предположить, что первым действием для любого «карательного» устройства по отношению к поддерживающему ECN потоку будет отбрасывание пакетов вместо их маркировки, тогда у злоумышленника уже не останется возможности нарушить сквозной контроль перегрузки на базе ECN путём ложного представления этого потока, как несовместимого с контролем перегрузки, для принятия к нему более жёстких мер в «карательном» устройстве.

Устройства мониторинга и реализации политики, разворачиваемые на практике, могут отличаться от описанного выше «идеального» устройства мониторинга в том, что отслеживание происходит применительно не к отдельному потоку, а к агрегату потоков (например, к потокам одного туннеля IPsec). В этом случае переход от маркировки к отбрасыванию будет осуществляться для всех потоков агрегата, сводя на нет преимущества использования ECN для этих потоков. При высоком уровне агрегирования наблюдается другой вариант запрета ECN даже при отсутствии устройств мониторинга и реализации политики, когда поддерживающие ECN очереди RED переключаются с маркировки пакетов на их отбрасывание в качестве индикации перегрузки в тех случаях, когда средний размер очереди превышает заданный порог.

11. Оценка ECN

11.1. Работы по оценке использования ECN

В этом разделе рассмотрены некоторые работы в которых проводится оценка использования ECN. На Web-странице ECN [ECN] приведены ссылки на работы по ECN и реализации ECN.

В работе [Floyd94] рассмотрены преимущества и недостатки, связанные с добавлением ECN в архитектуру TCP/IP. Как показано в основанном на модели сравнении, одним из преимуществ ECN является избавление от ненужного отбрасывания пакетов для краткосрочных и чувствительных к задержкам соединений TCP. Вторым преимуществом является предотвращение некоторых ненужных повторов передачи по тайм-аутам TCP. В этой статье подробно обсуждается интеграция ECN с механизмами контроля перегрузки TCP. К отмеченным в статье возможным недостаткам ECN относится то, что неподатливые соединения TCP могут ложно анонсировать себя, как поддерживающие ECN, а также возможность отбрасывания в сети пакетов TCP ACK, содержащих сообщения ECN-Echo. Первый из этих недостатков обсуждается в приложении к данному документу, а второй снимается путём добавления флага CWR в заголовок TCP.

Экспериментальные оценки ECN проведены в [RFC2884, K98] и сделано заключение о том, что ECN TCP обеспечивает умеренное повышение производительности по сравнению с TCP без использования ECN, потоки ECN TCP не нарушают работу потоков без поддержки ECN и ECN TCP обеспечивает отказоустойчивость для случаев перегрузки в обоих направлениях и наличия на пути множества перегруженных маршрутизаторов. Эксперименты со множеством коротких соединений для передачи web-трафика показали, что для большинства коротких соединений в результате использования ECN время передачи значительно сокращалось.

11.2. Обсуждение ECN nonce[1]

Использование двух кодов ECT - ECT(0) и ECT(1) - может обеспечивать однобитовый маркер ECN nonce в заголовках пакетов [SCWA99]. Основной целью введения такой маркировки является предоставление отправителю возможности обнаружения фактов удаления кода CE элементами сети на пути и проверки корректности информирования со стороны получателя о приёме пакетов с кодом CE, требуемого транспортным протоколом. В этом параграфе обсуждаются вопросы совместимости с реализациями IP ECN в маршрутизаторах, соответствующих RFC 2481, которые поддерживают только один код ECT. Мы полагаем, что расширяющееся развёртывание реализаций ECN, понимающих

[1] В соответствии с RFC 8311 этот и следующий (11.2.1) параграфы исключены. *Прим. перев.*

код ECT(1), не будет вызывать каких-либо проблем. Отсутствие таких проблем очевидно для случаев, когда поддержка ECT(1) в маршрутизаторах реализуется до того, как этот код начинают использовать конечные узлы.

11.2.1. Поэтапное развёртывание ECT(1) в маршрутизаторах

ECN имеет статус экспериментального стандарта с января 1999 и уже имеются реализации ECN в маршрутизаторах, которые не понимают кода ECT(1). При стандартизации использования ECT(1) для TCP или других транспортных протоколов это может привести к тому, что отправители данных будут использовать код ECT(1), а некоторые перегруженные маршрутизаторы на пути доставки пакетов не будут понимать этот код.

Если транспортный протокол позволяет, отправитель данных может совсем не использовать код ECT(1) и передавать все поддерживающие ECN пакеты с кодом ECT(0). Если поддерживающий ECN отправитель использует ECT(1), а перегруженный маршрутизатор на пути передачи не понимает код ECT(1), этот маршрутизатор будет маркировать некоторые пакеты с кодом ECT(0) и отбрасывать некоторые пакеты с кодом ECT(1) для индикации перегрузки. Поскольку протокол TCP должен реагировать как на маркировку, так и на отбрасывание пакетов, при отбрасывании пакетов, которые могли бы быть промаркированы, не возникает каких-либо существенных проблем в сети и такое отбрасывание совместимо в общей моделью работы ECN, позволяющей маршрутизаторам самостоятельно принимать решение об отбрасывании или маркировке пакетов (см. раздел 5).

12. Список требуемых изменений для IP и TCP

В этом документе определено использование для ECN двух битов заголовка IP. Код not-ECT показывает, что транспортный протокол будет игнорировать код CE. Этот код является принятым по умолчанию значением ECN. Коды ECT показывают, что транспортный протокол хочет и может участвовать в работе ECN.

Маршрутизатор устанавливает код CE для индикации перегрузки конечным узлам. Маршрутизаторам **недопустимо** сбрасывать код CE в заголовках пакетов.

Для поддержки ECN в протокол TCP требуется внести три изменения - это фаза организации соединения и два новых флага в заголовке TCP. Флаг ECN-Echo используется получателем данных для информирования отправителя о получении пакета CE. Флаг сокращения окна перегрузки (CWR) используется отправителем для информирования получателя о снижении размера окна.

При использовании ECN (явное уведомление о перегрузке) требуется, чтобы индикаторы перегрузки, генерируемые в туннеле IP, не терялись на выходе из туннеля. Мы вносим незначительные изменения в протокол IP для обработки поля ECN при инкапсуляции и декапсуляции, позволяющие потокам, проходящим через туннели IP, использовать ECN.

В туннелях для ECN определяются две опции.

1) Опция ограниченной функциональности без использования ECN внутри туннеля IP путём установки в поле ECN значения not-ECT и сохранения внутреннего заголовка при декапсуляции.

2) Полнофункциональная опция, когда в поле ECN внешнего заголовка может использоваться код not-ECT или один из кодов ECT в зависимости от значения поля ECN во внутреннем заголовке. При декапсуляции код CE копируется во внутренний заголовок, если этот код присутствует во внешнем заголовке, а во внутреннем установлен один из кодов ECT.

Для туннелей IPsec в этом документе также определён необязательный атрибут защищённой связи (SA), управляющий согласованием использования ECN в туннеле IPsec и необязательное поле SAD для индикации возможности использования ECN в туннельном режиме для SA. Требуемые для использования ECN изменения туннелей IPsec вносят изменения в документ RFC 2401 [RFC2401], определяющий архитектуру IPsec и задающий некоторые аспекты реализации. Новый атрибут IPsec SA дополняет атрибуты, определённые в параграфе 4.5 [RFC2407].

Этот документ отменяет действие RFC 2481 «A Proposal to add Explicit Congestion Notification (ECN) to IP[1]», который определял ECN в качестве экспериментального протокола для сообщества Internet. В оставшейся части этого параграфа описаны связи настоящего документа с его предшественником.

RFC 2481 включает краткое обсуждение использования ECN с инкапсулированными пакетами, где отмечено, что спецификация IPsec на тот момент (январь 1999) говорит о невозможности безопасного использования ECN через туннели IPsec. В RFC 2481 также описаны изменения, которые нужно было внести в спецификации туннелей IPsec для обеспечения совместимости с ECN.

В этом документе учтены работы, выполненные с момента публикации RFC 2481. Во-первых, подробно описаны изменения для туннелей IPsec и влияние ECN на безопасность (разделы 18 и 19). Во-вторых, рассмотрение туннелей IPsec расширено для всех туннелей IP. Поскольку старые туннели IP не совместимы с потоками, использующими ECN, развёртывание ECN в сети Internet оказывало сильное давление на процессы обновления старых реализаций туннелей до совместимых с ECN версий на основе полной или ограниченной функциональности.

В этом документе не рассматриваются вопросы использования ECN в туннелях других протоколов (не IP), таких как MPLS, GRE, L2TP, PPTP. В настоящее время предварительные документы по включению поддержки ECN в MPLS находятся в стадии разработки.

В-третьих, после публикации RFC2481 были описаны процедуры ECN для повторно передаваемых пакетов данных, когда код ECT при повторной передаче устанавливать не следует. Мотивом отказа от использования кодов при повторе передачи послужило желание предотвратить возможные атаки на службы (DoS) для существующих соединений TCP. Некоторые ранние реализации TCP с поддержкой ECN могут не соответствовать (новым) требованиям по отказу от установки кода ECT в повторно передаваемых пакетах. Мы полагаем, что на практике это не вызовет существенных проблем.

Этот документ также несколько расширяет спецификацию использования пакетов SYN для согласования ECN. Некоторые ранние реализации TCP с поддержкой ECN могут не соответствовать заданным здесь требованиям, но мы

[1]Предложение по добавлению явных уведомлений о перегрузке (ECN) в протокол IP.

полагаем, что это не вызовет проблем в организации и производительности соединений TCP в которых участвуют реализации с различными возможностями.

Этот документ также включает спецификацию кода ECT(1), который может использоваться протоколом TCP, как часть реализации ECN nonce.[1]

13. Заключение

На основе опыта развёртывания. AQM мы полагаем, что настало время развёртывания. механизмов предотвращения перегрузок, не зависящих от отбрасывания пакетов. По мере развёртывания. приложений и транспортного сервиса, чувствительных к задержкам и потере данных (например, трафик в реальном масштабе времени, короткие web-транзакции), использование потери пакетов в качестве механизма индикации перегрузки становится недостаточно эффективным (по крайней мере, неоптимальным).

Мы оценили последствия изменения поля ECN в сети, проанализировав все возможные варианты враждебного изменения поля ECN. Во многих случаях изменение поля ECN не представляет большей опасности, чем отбрасывание пакета. Однако мы отмечаем, что некоторые изменения могут приводить к более серьёзным последствиям с нарушением сквозного контроля перегрузки. Но даже в таких случаях потенциальные нарушения ограничены и подобны угрозам, создаваемым время от времени возникающими отказами при взаимодействии систем для организации сквозного контроля перегрузки.

14. Благодарности

В подготовку этого документа внесло свой вклад много людей, включая тех, кто не участвовал в создании документа непосредственно. Мы хотим явно поблагодарить Kenjiro Cho за предложения по механизмам TCP для согласования поддержки ECN, Kevin Fall за предложения по флагу CWR, Steve Blake за материалы по пересчету контрольной суммы заголовков IPv4, Jamal Hadi-Salim за обсуждение проблем ECN, а также Steve Bellovin, Jim Bound, Brian Carpenter, Paul Ferguson, Stephen Kent, Greg Minshall и Vern Paxson за обсуждение вопросов безопасности. Мы также благодарим участников исследовательской группы Internet End-to-End за обсуждение многих вопросов.

Обмен информацией по электронной почте со многими людьми, включая Dax Kelson, Alexey Kuznetsov, Jamal Hadi-Salim и Venkat Venkatsubra, помог в решении вопросов, связанных с несовместимым оборудованием в сети Internet, которое не реагирует на пакеты TCP SYN с установленными флагами ECE и CWR. Мы благодарим Mark Handley, Jitentra Padhye и других за обсуждение вопросов инициализации TCP.

Обсуждение вопросов взаимодействия ECN и туннелей IP было прочно связано и дискуссиями и документами рабочей группы Differentiated Services. Мы благодарим Tabassum Bint Haque из Dhaka, Bangladesh за отклики по туннелям IP. Мы также благодарны Derrell Piper и Kero Tivinen за предложенные изменения RFC 2407, позволившие повысить уровень применимости согласования атрибута SA при использовании ECN с туннелями.

Мы благодарим David Wetherall, David Ely и Neil Spring за их предложения по ECN. Благодарим также Stefan Savage за обсуждение этой тематики. Мы признательны Bob Briscoe и Jon Crowcroft за поднятый вопрос о фрагментации IP, дополнительную семантику четвёртого кода ECN и обсуждение других тем. Мы благодарны Richard Wendland за отклики по нескольким затронутым в документе вопросам.

Мы также благодарим IESG и, в частности, руководителей направления Transport за их отклики и усилия по стандартизации ECN.

15. Литература

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

[ECN] "The ECN Web Page", URL "http://www.aciri.org/floyd/ecn.html".[2]

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

[FIXES] ECN-under-Linux Unofficial Vendor Support Page, URL "http://gtf.org/garzik/ecn/".[2]

[FJ93] Floyd, S., and Jacobson, V., "Random Early Detection gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, V.1 N.4, August 1993, p. 397-413.

[Floyd94] Floyd, S., "TCP and Explicit Congestion Notification", ACM Computer Communication Review, V. 24 N. 5, October 1994, p. 10-23.

[Floyd98] Floyd, S., "The ECN Validation Test in the NS Simulator", URL "http://www-mash.cs.berkeley.edu/ns/", test tcl/test/test-all-ecn.[2]

[FF99] Floyd, S., and Fall, K., "Promoting the Use of End-to-End Congestion Control in the Internet", IEEE/ACM Transactions on Networking, August 1999.

[FRED] Lin, D., and Morris, R., "Dynamics of Random Early Detection", SIGCOMM '97, September 1997.

[GRE] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994.

[Jacobson88] V. Jacobson, "Congestion Avoidance and Control", Proc. ACM SIGCOMM '88, pp. 314-329.

[Jacobson90] V. Jacobson, "Modified TCP Congestion Avoidance Algorithm", Message to end2end-interest mailing list, April 1990. URL "ftp://ftp.ee.lbl.gov/email/vanj.90apr30.txt".

[K98] Krishnan, H., "Analyzing Explicit Congestion Notification (ECN) benefits for TCP", Master's thesis, UCLA, 1998. Цитируется лишь в качестве благодарности.

[1] В соответствии с RFC 8311 этот абзац исключён. *Прим. перев.*
[2] Приведено только для справки.

[L2TP] Townsley, W., Valencia, A., Rubens, A., Pall, G., Zorn, G. and B. Palter, "Layer Two Tunneling Protocol "L2TP"", RFC 2661, August 1999.

[MJV96] S. McCanne, V. Jacobson, and M. Vetterli, "Receiver-driven Layered Multicast", SIGCOMM '96, August 1996, pp. 117-130.

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

[PPTP] Hamzeh, K., Pall, G., Verthein, W., Taarud, J., Little, W. and G. Zorn, "Point-to-Point Tunneling Protocol (PPTP)", RFC 2637, July 1999.

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

[RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.

[RFC1141] Mallory, T. and A. Kullberg, "Incremental Updating of the Internet Checksum", RFC 1141, January 1990.

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

[RFC1455] Eastlake, D., "Physical Link Security Type of Service", RFC 1455, May 1993.

[RFC1701] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation (GRE)", RFC 1701, October 1994.

[RFC1702] Hanks, S., Li, T., Farinacci, D. and P. Traina, "Generic Routing Encapsulation over IPv4 networks", RFC 1702, October 1994.

[RFC2003] Perkins, C., "IP Encapsulation within IP", RFC 2003, October 1996.

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

[RFC2309] Braden, B., et al., "Recommendations on Queue Management and Congestion Avoidance in the Internet", RFC 2309, April 1998.

[RFC2401] Kent, S. and R. Atkinson, Security Architecture for the Internet Protocol, RFC 2401, November 1998.

[RFC2407] Piper, D., "The Internet IP Security Domain of Interpretation for ISAKMP", RFC 2407, November 1998.

[RFC2408] Maughan, D., Schertler, M., Schneider, M. and J. Turner, "Internet Security Association and Key Management Protocol (ISAKMP)", RFC 2408, November 1998.

[RFC2409] Harkins D. and D. Carrel, "The Internet Key Exchange (IKE)", RFC 2409, November 1998.

[RFC2474] Nichols, K., Blake, S., Baker, F. and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", RFC 2474, December 1998.

[RFC2475] Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z. and W. Weiss, "An Architecture for Differentiated Services", RFC 2475, December 1998.

[RFC2481] Ramakrishnan K. and S. Floyd, "A Proposal to add Explicit Congestion Notification (ECN) to IP", RFC 2481, January 1999.

[RFC2581] Alman, M., Paxson, V. and W. Stevens, "TCP Congestion Control", RFC 2581, April 1999.

[RFC2884] Hadi Salim, J. and U. Ahmed, "Performance Evaluation of Explicit Congestion Notification (ECN) in IP Networks", RFC 2884, July 2000.

[RFC2983] Black, D., "Differentiated Services and Tunnels", RFC 2983, October 2000.

[RFC2780] Bradner S. and V. Paxson, "IANA Allocation Guidelines For Values In the Internet Protocol and Related Headers", BCP 37, RFC 2780, March 2000.

[RJ90] K. K. Ramakrishnan and Raj Jain, "A Binary Feedback Scheme for Congestion Avoidance in Computer Networks", ACM Transactions on Computer Systems, Vol.8, No.2, pp. 158-181, May 1990.

[SCWA99] Stefan Savage, Neal Cardwell, David Wetherall, and Tom Anderson, TCP Congestion Control with a Misbehaving Receiver, ACM Computer Communications Review, October 1999.

[TBIT] Jitendra Padhye and Sally Floyd, "Identifying the TCP Behavior of Web Servers", ICSI TR-01-002, February 2001. URL "http://www.aciri.org/tbit/".

16. Вопросы безопасности

Вопросы безопасности рассматриваются в разделах 7, 8, 18 и 19.

17. Пересчёт контрольной суммы в заголовке IPv4

При пересчёте контрольной суммы заголовка IPv4 возникает проблема с некоторыми маршрутизаторами, использующими буферизацию на выходе, поскольку большинство (если не все) операций с заголовком выполняются на входе, а решение для ECN нужно принимать локально по состоянию выходного буфера. Этой проблемы не возникает для IPv6, поскольку этот протокол не использует контрольных сумм для заголовков. Октет IPv4 TOS является последним байтом 16-битового полуслова.

В RFC 1141 [RFC1141] обсуждается обновление контрольной суммы IPv4 после уменьшения значения поля TTL. Обновление контрольной суммы IPv4 после установки кода CE описано ниже. Обозначим HC исходную контрольную сумму заголовка для пакета ECT(0), а HC' будет новой контрольной суммой после установки бита CE (т. е., поле ECN изменит значение с 10 на 11). Тогда контрольная сумма заголовка вычисляется путём вычитания дополнения до 1

```
HC' = { HC - 1     HC > 1
```

```
              { 0x0000      HC = 1
```

Для расчёта контрольной суммы на машинах с дополнением до двух HC' после установки флага CE будет

```
        HC' = { HC - 1      HC > 0
              { 0xFFFE      HC = 0
```

Похожее изменение контрольной суммы IPv4 может выполняться при изменении поля ECN с ECT(1) на CE (с 01 на 11).

18. Возможные изменения поля ECN в сети

В этом разделе подробно рассматривается вопрос возможного изменения в сети поля ECN и связанные с этим последствия - ложные уведомления о перегрузках, запрет поддержки ECN для отдельных пакетов, удаление индикации ECN о перегрузке или ложная индикация поддержки ECN.

18.1. Возможные изменения заголовка IP

18.1.1. Удаление индикатора перегрузки

Сначала рассмотрим изменения, которые может внести маршрутизатор и которые будут приводить к эффективному удалению индикации перегрузки. При нормальной работе код ECN принятого пакета определяет код ECN передаваемого пакета.

Замена кода CE на ECT(0) или ECT(1) фактически удаляет индикацию перегрузки. Однако при использовании двух кодов ECT маршрутизатор, удаляющий код CE, не знает, какой код ECT был в пакете изначально - ECT(0) или ECT(1). Таким образом, транспортный протокол может реализовать механизмы детектирования фактов удаления кода CE.

Результат удаления кода CE для маршрутизатора восходящего потока состоит в потенциальной возможности организации перегрузки, поскольку индикация не достигает отправителя. Однако пакет будет получен и подтверждён.

Потенциальный эффект от удаления индикации перегрузки является комплексным и подробно рассматривается в разделе 19. Отметим, что эффект удаления индикации перегрузки может отличаться от случая отбрасывания пакета в сети. При отбрасывании пакета этот факт обнаруживается отправителем TCP и трактуется, как индикация перегрузки. Аналогично при достаточном числе отброшенных пакетов подтверждения, вызывающем прекращение роста значения кумулятивного поля подтверждений на стороне отправителя последний ограничивает дальнейшую передачу пакетов размером окна перегрузки и, в конечном итоге, повторяет передачу по таймеру.

В противоположность этому систематическое отбрасывание бита CE маршрутизатором нисходящего потока будет приводить к созданию очереди на маршрутизаторе восходящего потока с возможностью потери пакетов в результате переполнения буфера. При этом может теряться беспристрастность, поскольку другие потоки через перегруженный маршрутизатор могут реагировать на установленный бит CE, тогда как поток с удаленным битом CE будет по-прежнему отнимать ресурсы буфера. Подавление этого негативного эффекта подробно рассматривается в разделе 19.

Последним из трёх возможных изменений является замена кода CE кодом not-ECT, которая одновременно удаляет индикацию перегрузки и запрещает использование ECN.

Удаление индикации перегрузки даёт эффект только в том случае, когда пакет впоследствии не будет промаркирован заново или отброшен маршрутизатором нисходящего потока. Если код CE заменяется кодом ECT, пакет сохраняет совместимость с ECN и может быть снова промаркирован или отброшен маршрутизатором нисходящего потока для индикации перегрузки. Если код CE меняется на not-ECT, пакет утрачивает совместимость с ECN и может быть отброшен, но не может быть промаркирован для индикации перегрузки.

18.1.2. Ложная информация о перегрузке

Это изменение заключается в установке кода CE при уже установленном коде ECT вне зависимости от наличия реальной перегрузки. Данное изменение не влияет на трактовку пакета по пути его передачи. В частности, маршрутизаторы не проверяют наличие кода CE при решении вопроса о маркировке или отбрасывании пакета. Однако это изменение может приводить к необоснованному включению механизма сквозного контроля перегрузки и снижению скорости доставки пакетов. Само по себе это не создаёт дополнительных проблем (для приложений или сети) по сравнению с отбрасыванием пакетов маршрутизатором.

18.1.3. Запрет поддержки ECN

Это изменение заключается в отключении кода ECT для пакета. Если такой пакет в дальнейшем столкнётся с перегрузкой (например, при попадании в RED-очередь с умеренным размером буфера), он будет отброшен вместо маркировки. Само по себе это не создаёт больших проблем (для приложений), чем обычное отбрасывание пакета маршрутизатором. Сдерживающим фактором в данном случае является отсутствие маршрутизатора восходящего потока, ожидающего реакции на установку бита CE.

18.1.4. Ложная индикация поддержки ECN

Это изменение состоит в некорректной индикации для пакета поддержки ECN. На практике пакет может быть передан как поддерживающим, так и не поддерживающим ECN транспортом.

Если пакет впоследствии столкнётся с умеренной перегрузкой на поддерживающем ECN маршрутизаторе, этот маршрутизатор может установить код CE вместо отбрасывания пакета. Если транспортный протокол реально не поддерживает ECN, транспорт не будет получать такую индикацию перегрузки и не будет снижать скорость передачи. Возможные последствия ложной индикации поддержки ECN рассматриваются в разделе 19.

Если пакет на остальном пути не сталкивается с перегрузкой на поддерживающих ECN маршрутизаторах, первое изменение не будет давать никакого эффекта, кроме возможного создания помех использованию ECN транспортным протоколом. Второе изменение, однако, будет давать эффект за счёт ложной индикации перегрузки устройствам мониторинга на пути доставки. Если транспортный протокол поддерживает ECN, это изменение может также оказывать влияние на транспортный уровень за счёт комбинирования ложной индикации поддержки ECN с ложными отчётами о перегрузке. Для поддерживающего ECN транспорта это будет вызывать необоснованную реакцию на перегрузку со

стороны транспортного протокола. В данном частном случае маршрутизатор, некорректно изменивший поле ECN, может отбросить пакет. Таким образом, для этого случая при использовании совместимого с ECN транспорта последствия такого изменения поля ECN не будут хуже, чем при обычном отбрасывании пакета.

18.2. Информация, передаваемая в транспортном заголовке

Для протокола TCP поддерживающий ECN получатель может информировать партнёра TCP о своей поддержке ECN на уровне протокола TCP, помещая эту информацию в заголовок TCP на этапе организации соединения. В этом документе не рассматриваются угрозы, связанные с возможностью изменения заголовка транспортного уровня в сети. Отметим, что при использовании IPsec заголовок транспортного уровня защищён как в туннельном, так и в транспортном режиме [ESP, AH].

Другой вопрос связан с пакетами TCP, содержащими обманный адрес IP и некорректную информацию ECN в заголовке транспортного уровня. Для полноты мы рассмотрели возможные варианты, когда узел, подставляющий обманный IP-адрес отправителя, может использовать два флага ECN в заголовке TCP для организации DoS-атаки. Для таких атак злоумышленнику нужны корректные порядковые номера TCP и атакующий, способный создавать корректные номера и использовать обманный IP-адрес отправителя, может повредить соединение TCP даже без использования флагов ECN. Следовательно, ECN в данном случае не добавляет новых уязвимостей.

Пакет подтверждения с обманным IP-адресом отправителя, указывающим на получателя TCP, может иметь флаг ECE. Если этот пакет будет принят отправителем данных TCP как корректный, это может привести к ненужному снижению вдвое размера окна перегрузки на передающей стороне TCP. Однако для того, чтобы отправитель принял обманный пакет подтверждения, в этом пакете должен быть указан корректный 32-битовый порядковый номер, а также корректный номер подтверждения. Атакующий, который способен создавать обманные пакеты с корректными порядковыми номерами, может просто передать обманный пакет RST или иной пакет, способный нарушить работу соединения TCP.

Пакеты с обманным IP-адресом отправителя, указывающим на передающую сторону TCP, могут иметь флаг CWR. Для такого пакета также требуется указать корректный порядковый номер. Кроме того, пакеты такого типа могут оказать лишь незначительное влияние на работу соединения. Обманный пакет данных с флагом CWR может привести к тому, что получатель TCP станет передавать меньше пакетов ECE, чем обычно, если он передавал такие пакеты в момент получения пакета с флагом CWR.

18.3. Расщеплённые пути

В некоторых случаях враждебные или некорректно настроенные маршрутизаторы могут иметь доступ лишь к части пакетов потока. Возникает вопрос, может ли такой маршрутизатор, меняя значения поля ECN в доступном ему подмножестве пакетов потока, повреждать этот поток сильнее, чем при простом отбрасывании пакетов?

Разделим пакеты потока на две группы - A и B, предположив, что злоумышленнику доступны только пакеты группы A. Предположим, что враждебное воздействие выражается в нарушении сквозного контроля перегрузки на пути, по которому проходят пакеты группы A, путём ложной индикации поддержки ECN для восходящего направления, на котором наблюдается перегрузка, или путём удаления индикации перегрузки в нисходящем направлении. Предположим также, что имеется устройство мониторинга, которое видит пакеты обеих групп и будет «карать» пакеты групп, если суммарный поток, доступный устройству, не реагирует должным образом на индикацию перегрузки. Другим важным моментом (полагаем, что это достаточно очевидно) является то, что устройство мониторинга до использования «карательных» мер по отношению к потоку A и B будет сначала отбрасывать пакеты вместо установки в них кода CE, а также будет отбрасывать прибывающие пакеты, в которых уже установлен код CE. Если конечные узлы на практике используют сквозной контроль перегрузки, они будут видеть все индикаторы перегрузки, которые видны устройству мониторинга, и будут должным образом реагировать на перегрузку. Таким образом, устройство мониторинга успешно обеспечивает индикацию для потока на начальном этапе.

Атакующий, который имеет доступ только к пакетам A, нарушая работу системы контроля перегрузки на основе ECN, способен свести на нет преимущества ECN для других пакетов суммарного потока A и B. Это печальный факт, но он не является достаточным основанием для отказа от ECN.

Вариант ложной индикации перегрузки возникает в том случае, когда на пути имеется два злоумышленника, из которых первый создаёт ложную индикацию, а второй «удаляет» её (в отличии от случая отбрасывания пакетов индикация перегрузки в ECN может быть «реверсирована» на пути враждебным или некорректно настроенным маршрутизатором, однако использование ECN nonce может помочь в обнаружении такого поведения). Несмотря на то, что описанные выше искажения не заметны для конечных узлов, между первым и вторым злоумышленником может оказаться устройство мониторинга, которое увидит ложную индикацию перегрузки. Как было отмечено выше, в этом случае до начала «карательных» действий по отношению к потоку, не реагирующему должным образом на перегрузку, маршрутизатор сначала перейдёт на режим отбрасывания пакетов потока взамен их маркировки. Когда этот процесс включает отбрасывание прибывающих пакетов с кодом CE, конечные узлы получат индикацию перегрузки на пути. Таким образом, не возникает дополнительных угроз в результате действия множества конфликтующих злоумышленников.

19. Влияние нарушения сквозного контроля перегрузки

В этом разделе рассматриваются возможные последствия нарушения сквозного контроля перегрузки путём ложной индикации поддержки ECN или удаления индикации перегрузки в ECN (код CE). Нарушение сквозного контроля перегрузки любым из этих способов может оказывать влияние как на приложения, так и на сеть. Ниже эти варианты подробно рассматриваются по отдельности.

Первый метод нарушения сквозного контроля перегрузки заключается в ложной индикации поддержки ECN. Эффективно нарушить сквозной контроль перегрузки при этом удаётся лишь в тех случаях, когда на пути пакета действительно наблюдается перегрузка, приводящая к установке кода CE. В этом случае транспортный протокол (который может не поддерживать ECN) не будет получать индикацию насыщения от перегруженных маршрутизаторов нисходящего направления.

Вторым методом нарушения сквозного контроля перегрузки является «удаление» кода CE из пакетов. Эффективное нарушение контроля перегрузки при этом получается лишь в тех случаях, когда в пакете уже был код CE установленный перегруженным маршрутизатором. В этом случае транспортный протокол не получает индикации насыщения от перегруженных маршрутизаторов восходящего направления.

Любой из этих методов нарушения сквозного контроля перегрузки потенциально способен наносить больший вред сети (и, возможно, самому потоку), нежели простое отбрасывание враждебным устройством пакетов из потока. Однако, как сказано в разделе 7 и будет сказано ниже, спектр возможных нарушений весьма ограничен.

19.1. Влияние на сеть и конкурирующие пути

Код CE в поле ECN используется лишь маршрутизаторами для индикации перегрузки в период «умеренной» перегрузки. Поддерживающим ECN маршрутизаторам в периоды существенных перегрузок следует отбрасывать пакеты, а не маркировать их даже в тех случаях, когда в очередях маршрутизатора ещё имеется место. Например, маршрутизаторам, использующим активное управление очередями на основе RED, следует отбрасывать пакеты вместо их маркировки, если средний размер очередей превышает верхний порог для очередей RED.

Одним из следствий нарушения сквозного контроля перегрузки для сетей является то, что потоки, не получающие индикации перегрузки от сети, могут увеличивать свою скорость передачи, пока не будет достигнута существенная перегрузка в сети. Тогда перегруженный маршрутизатор может начать отбрасывание поступающих пакетов вместо их маркировки. Для потоков, которые не изолированы с помощью механизмов независимого планирования или аналогичных методов и при этом агрегируются с другими потоками в одной очереди без дифференциации, такое отбрасывание потоков на перегруженном маршрутизаторе будет применяться ко всем потокам, использующим общую очередь. В результате будет расти уровень перегрузки в сети.

В некоторых случаях рост перегрузки будет приводить к существенному росту заполненности буферов на загруженных очередях, которого будет достаточно для перевода очередей из режима маркировки пакетов в режим их отбрасывания. Этот переход будет происходить в результате переполнения буферов или в соответствии с политикой активного управления очередями, описанного выше, когда пакеты начинают отбрасываться при достижении заданного порога RED. С этого момента все потоки, включая поток с нарушениями, будут видеть отбрасывание пакетов вместо их маркировки и все враждебные или некорректно настроенные маршрутизаторы больше не смогут удалять такую индикацию перегрузки в сети. Если конечные узлы используют подходящую технологию сквозного контроля перегрузки, в нарушенном потоке скорость передачи будет снижена в ответ на перегрузку. Когда уровень перегрузки существенно снизится, перегруженная очередь может вернуться из режима отбрасывания пакетов в режим маркировки. В установившемся состоянии могут наблюдаться осцилляции очереди с переходом из одного режима в другой.

В других случаях последствия нарушения сквозного контроля перегрузки не будут достаточно велики для того, чтобы перевести нагруженный канал в состояние столь сильной перегрузки, что начинается отбрасывание пакетов вместо их маркировки. В такой ситуации для конкурирующих пакетов в сети будет слегка возрастать скорость маркировки или отбрасывания пакетов с соответствующим снижением доступной этим потокам полосы. Это состояние может быть стабильным, если скорость прибытия пакетов нарушенного потока достаточно мала по сравнению с полосой канала и средний размер очереди на загруженных маршрутизаторах остаётся под контролем. В частности, нарушенный поток может иметь ограниченные потребности в полосе канала на данном маршрутизаторе, но при этом запрашивать существенно больше. Ограниченность запросов может быть обусловлена потребностями отправителя данных, ограничениями анонсированного окна TCP, малой полосой канала доступа или другими факторами. Таким образом, нарушение контроля перегрузки на основе ECN может вести к утрате беспристрастности разделения полосы, которая была отмечена выше.

Опасность для сети, порождаемая нарушением основанного на ECN контроля перегрузки в сети, не отличается существенно от опасности, вносимой возникающими время от времени случаями отказов при организации сквозного контроля перегрузки. Развёртывание в маршрутизаторах механизмов, позволяющих решить эту проблему, является предметом другого исследования и рассмотрено в разделе 10.

Вернёмся к примеру, описанному в параграфе 18.1.1, где установленный в пакете код CE удалялся: {11 -> 10 или 11-> 01}. Последствием этого для перегруженного маршрутизатора восходящего направления, который установил код CE, является то, что эта индикация не достигнет конечного узла данного потока. Отправитель (даже из числа полностью кооперированных и не враждебных) в результате отсутствия индикации может продолжать увеличение скорости передачи (для TCP путём расширения окна перегрузки). Поток может получить более высокую пропускную способность на перегруженном маршрутизаторе по сравнению с другими потоками, особенно при отсутствии на маршрутизаторе механизмов реализации политики или независимого распределения очередей по потокам. Рассмотрим поведение других потоков (особенно, кооперированных), которые не подверглись нарушению сквозного контроля перегрузки. Ясно, что эти потоки снизят свою нагрузку на данный маршрутизатор (например, за счёт сокращения окна), отдавая преимущества нарушенному потоку. Таким образом утрачивается беспристрастность. Как обсуждалось выше, эта пристрастность может быть кратковременной (поскольку перегруженная очередь находится в режиме маркировки пакетов), осциллирующей (очередь переключается между режимами маркировки и отбрасывания) или более сдержанной, но постоянной (поскольку очередь никогда не переключается в режим отбрасывания пакетов).

Результат для нарушенного потока похож на случай преднамеренного отказа от сквозного контроля перегрузки. Единственное отличие заключается в том, что поток, преднамеренно отказавшийся от сквозного контроля перегрузки на конечных узлах, может избежать контроля перегрузки даже при переходе перегруженной очереди в режим отбрасывания пакетов, отвергая снижение скорости передачи в ответ на отбрасывание пакетов в сети. Таким образом, проблема для сетей с нарушенным контролем перегрузки на базе ECN менее значима, чем проблема, вызываемая преднамеренным отказом от контроля перегрузки на конечных узлах. Важно отметить, что контролировать поведение конечных узлов существенно сложнее, чем контролировать поведение инфраструктуры. Отмеченное выше не говорит о малозначимости проблем, вносимых в сеть нарушением системы контроля перегрузки на базе ECN, просто приводится сравнение значимости этих проблем с другими нарушениями контроля перегрузки на конечных узлах.

19.2. Влияние на нарушенный поток

Когда отправитель указывает поддержку ECN, предполагается, что маршрутизаторы в сети, способные участвовать в ECN, будут использовать код CE для индикации перегрузки. Потенциальным преимуществом использования ECN

является снижение числа теряемых пакетов (в дополнение к снижению задержки в очередях за счёт активного управления). Когда пакет передаётся через туннель IPsec, проходящий через узлы, которые не пользуются доверием по тем или иным причинам, предполагается, что IPsec будет обеспечивать защиту, предотвращающую возникновение нежелательных последствий.

Во многих случаях нарушенный поток будет получать преимущества перед конкурирующими потоками данных в результате нарушения сквозного контроля перегрузки в сети за счёт увеличения доступной ему полосы. Если перегруженная очередь достигает порога отбрасывания пакетов, нарушение сквозного контроля перегрузки может давать или не давать преимущества нарушенному потоку, в зависимости от относительных потребностей потока в части пропускной способности, потерь и задержек.

Одной из форм нарушения сквозного контроля перегрузки является ложная индикация поддержки ECN путём установки кода ECT. Установка ложного кода CE будет оказывать влияние на перегруженные маршрутизаторы нисходящего направления. Однако, как было указано в параграфе 9.1.2, если код ECT меняется в туннеле IP, это может быть обнаружено на выходе туннеля, поскольку внутренний заголовок не будет изменён.

Другим вариантом нарушения сквозного контроля перегрузки является удаление индикации перегрузки путём исключения кода CE. В этом случае воздействие будет оказываться на перегруженные маршрутизаторы восходящего направления, которые устанавливают код CE.

Если код ECT удаляется внутри туннеля IP, это может быть обнаружено на выходе туннеля, поскольку внутренние заголовки не будут изменены. Если код CE устанавливается в восходящем направлении к туннелю IP, удаление кода CE в туннеле из внешнего заголовка не будет оказывать никакого влияния, поскольку установленный код CE сохранится во внутреннем заголовке. Однако, если код CE установлен внутри туннеля и удалён в туннеле или в нисходящем направлении от туннеля, такое удаление может остаться незамеченным на выходе из туннеля.

При таком нарушении сквозного контроля перегрузки транспортный протокол конечных узлов не реагирует на индикацию перегрузки. Вместе с утратой беспристрастности по отношению к незатронутым нарушением потокам, описанной в предыдущем параграфе, может продолжиться заполнение очереди перегруженного маршрутизатора, приводящее в результате к отбрасыванию пакетов, которое будет означать безусловную индикацию перегрузки транспортному протоколу. В переходном периоде поток может столкнуться существенным ростом задержки в очереди по сравнению с другими (ненарушенными) потоками. Но транспортные протоколы не принимают допущений о согласованности управления очередями на пути доставки. Мы полагаем, что эти формы нарушения сквозного контроля перегрузки не представляют большей опасности, чем простое отбрасывание пакетов потока.

19.3. Не связанные с ECN методы нарушения сквозного контроля перегрузки

Мы показали, что во многих случаях враждебный или некорректно настроенный маршрутизатор, который может изменять биты поля ECN, не способен внести больших помех, чем он мог бы доставить простым отбрасыванием пакетов. Однако это верно не во всех случаях - в частности, такое допущение неверно для некорректно работающих маршрутизаторов, нарушающих сквозной контроль перегрузки путём ложной индикации поддержки ECN или удаления индикации перегрузки ECN (код CE). Несмотря на наличие множества способов нанесения маршрутизатором вреда потоку за счёт отбрасывания пакетов, такое отбрасывание не может нарушить сквозной контроль перегрузки. Например, маршрутизатор не способен нарушить контроль перегрузки TCP путём отбрасывания пакетов данных, подтверждений или пакетов управления.

Хотя отбрасывание пакетов само по себе не может использоваться для нарушения сквозного контроля перегрузки, существуют не связанные с ECN методы нарушения сквозного контроля, которые могут использоваться враждебными или некорректно настроенными маршрутизаторами. Например, некорректно настроенный маршрутизатор может дублировать пакеты данных, сводя на нет контроль перегрузки на неком участке пути (для маршрутизатора, дублирующего пакеты в туннеле IPsec, администратор безопасности может настроить отбрасывание дубликатов путём организации в туннеле защиты против повторов). Такое дублирование пакетов будет оказывать на сеть и нарушенные потоки такое же влияние, как в описанных выше (параграфы 18.1.1 и 18.1.4) случаях.

20. Обоснование для маркеров ECT

20.1. Обоснование для кодов ECT

Необходимость введения кода ECT обусловлена тем, что развёртывание ECN в сети Internet будет осуществляться поэтапно и не все транспортные протоколы и маршрутизаторы будут понимать ECN. При использовании кода ECT маршрутизатор может отбрасывать пакеты, которые не совместимы с ECN, но может взамен отбрасывания устанавливать код CE в пакетах, которые поддерживают ECN. Поскольку код ECT позволяет конечному узлу получать код CE вместо информации об отбрасывании пакета, это даёт стимул для внедрения ECN.

Если в пакете не было кода ECT, маршрутизатор будет устанавливать код CE, как для поддерживающих, так и для не поддерживающих ECN потоков. В этом случае для конечных узлов нет стимула развёртывать ECN, а также не обеспечивается путь постепенного перехода к повсеместному использованию ECN. Рассмотрим первый этап постепенного развёртывания ECN, когда только часть потоков поддерживает ECN. В начале перегрузки, когда скорость отбрасывания/маркировки пакетов мала, маршрутизаторы будут лишь устанавливать код CE, не отбрасывая пакетов. Однако понимать пакеты с кодом CE и должным образом реагировать на них будут только потоки, поддерживающие ECN. В результате поддерживающие ECN потоки будут снижать скорость, а не понимающие сигналов ECN будут работать с прежним размером окна перегрузки.

В этом случае возможны два варианта: (1) поддерживающий ECN поток снижает скорость, не поддерживающий ECN поток забирает освободившуюся полосу и перегрузка сохраняется или (2) поддерживающий ECN поток снижает скорость, а не поддерживающий - не снижает и перегрузка возрастает, пока маршрутизатор не переходит от маркировки пакетов кодом CE к отбрасыванию. Хотя второй вариант не вполне беспристрастен, поддерживающий ECN поток в этом получает некоторые преимущества, поскольку увеличившаяся перегрузка заставляет маршрутизатор перейти в режим отбрасывания пакетов.

Поток, анонсирующий свою поддержку ECN, но не отвечающий на код CE, функционально эквивалентен потоку, в котором выключен контроль перегрузки, как обсуждалось ранее в этом документе.

Таким образом, среда, где часть потоков поддерживает ECN, но эти потоки не имеют механизма для индикации такой поддержки маршрутизаторам, будет менее эффективной и более пристрастно реагировать на перегрузки, что явится стимулом для конечных узлов к развёртыванию ECN.

20.2. Обоснование для двух кодов ECT

Основной причиной использования двух кодов ECT является необходимость поддержки однобитовых сигналов ECN nonce. Эти сигналы позволяют разработать для отправителя механизмы вероятностной проверки того, что элементы сети не удаляют код CE, а получатели данных корректно уведомляют отправителя о получении пакетов с кодом CE.

Другим способом обнаружения некорректно ведущих себя сетевых элементов или получателей является случайная передача пакетов данных с кодом CE и наблюдение за реакцией получателей на такие пакеты. Если такой пакет сталкивается с перегрузкой в сети, маршрутизатор, естественно, уже не может изменить пакет, поскольку флаг CE в нем уже установлен. Таким образом, для пакетов, переданных с кодом CE, конечные узлы не могут определить, где установлен этот код. По этой причине передача пакетов с кодом CE является более экономной, но менее эффективной мерой обнаружения некорректно работающих сетевых элементов и конечных узлов по сравнению с ECN nonce.[1]

Выделение четвёртого кода ECN - ECT(1) преследовало иные цели. Для ясности эти цели кратко перечислены ниже.

Одной из целей может быть использование отправителем четвёртого кода ECN для индикации дополнительной семантики ECN. Однако для такой сигнализации нам представляется более целесообразным использование кодов дифференцированного обслуживания в поле DS.

Второй возможной целью является использование четвёртого кода ECN для предоставления маршрутизаторам двух разных кодов перегрузки - CE(0) и CE(1) для слабого и сильной перегрузки, соответственно. Это может быть полезно в некоторых случаях, однако такое требование не представляется целесообразным. Если внести это требование, сложность постепенного развёртывания существенно возрастёт по сравнению со случаем использования одного кода.

Третьим неформально предложенным вариантом использования четвёртого кода ECN является его применение в некоторых формах контроля перегрузки для групповых приложений на основе процедур случайного дублирования маркированных пакетов в маршрутизаторах. Некоторые из предложенных процедур дублирования multicast-пакетов на маршрутизаторах основаны на новом коде ECN, который (1) переносит информацию о перегрузке в восходящем направлении от точки дублирования, промаркировавшей пакет этим кодом, и (2) может детектировать перегрузку в нисходящем направлении от точки дублирования. ECT(1) можно использовать для этой цели, поскольку этот код отличается от ECT(0) и заменяется кодом CE при маркировке ECN в ответ на перегрузку или начало перегрузки. Описание этой расширенной версии использования ECN для контроля перегрузки в групповых приложениях выходит за рамки документа, как и процедуры дублирования групповых пакетов, совместимые с ECN, или обработка поля ECN получателями группового трафика во всех случаях (независимо от процедур дублирования multicast-пакетов).

Спецификация изменения туннелей IP для ECN в этом документе предполагает, что единственным изменением, которое вносится в поле ECN внешнего заголовка IP между конечными точками туннеля, является установка кода CE для индикации перегрузки. Это не согласуется с предложениями по использованию ECT(1) процедурами дублирования multicast-пакетов, рассмотренными в предыдущем параграфе, и такие процедуры не следует развёртывать до разрешения противоречия между процедурами дублирования и туннелями IP с полной функциональностью ECN. Взамен может использоваться ограниченная функциональность ECN, хотя на практике многие протоколы туннелирования (включая IPsec) не будут корректно работать при дублировании группового трафика в туннеле.

21. Зачем использовать два бита в заголовке IP?

Необходимость индикации ECT в заголовке IP понятна, но остаётся вопрос о возможности использования для кодов ECT (транспорт с поддержкой ECN) и CE (обнаружена перегрузка) одного бита в заголовке. Такое однобитовое представление предложено в работе [Floyd94]. Одно значение «ECT, но без CE» будет представлять поддерживающий ECN транспорт, а другое - «CE или без ECT» будет представлять факт перегрузки или транспорт без поддержки ECN.

Различие между однобитовой и двухбитовой реализацией возникает для пакетов, проходящих через множество перегруженных маршрутизаторов. Рассмотрим пакет с кодом CE, который приходит на второй перегруженный маршрутизатор и выбирается системой активного управления очередью на маршрутизаторе для маркировки или отбрасывания. В однобитовом варианте второму перегруженному маршрутизатору остаётся только один вариант - отбросить пакет с кодом CE, поскольку этот маршрутизатор не может отличить пакет с кодом CE от пакета без ECT. В двухбитовом варианте второй перегруженный маршрутизатор может отбросить пакет с кодом CE или переслать его дальше, сохранив код CE.

Другое различие между однобитовыми и двухбитовыми реализациями заключается в том, что в однобитовом случае получатели не могут различить пакеты CE и non-ECT в одном потоке. Таким образом, в однобитовой реализации поддерживающий ECN отправитель будет давать получателю однозначную индикацию поддержки ECN. У отправителя остаётся возможность показать поддержку ECN в заголовке транспортного уровня. Другим вариантом является функциональное ограничение для однобитовых реализаций, когда отправитель трактует все переданные им пакеты как поддерживающие или не поддерживающие ECN. Для транспортных протоколов с групповой адресацией такая однозначная индикация будет видная получателям, присоединившимся к действующей multicast-сессии.

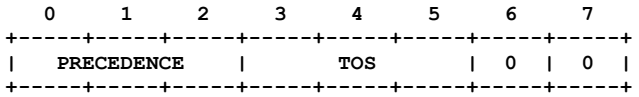
Другой, рассмотренный выше вопрос (и рекомендация) касается того, что транспортному протоколу (в частности, TCP) не следует маркировать чистые пакеты ACK и пакеты, передаваемые повторно, как поддерживающие ECN. Чистый пакет ACK из неподдерживающего ECN транспорта может быть отброшен без воздействия на транспорт с точки зрения контроля перегрузки (поскольку подтверждения кумулятивны). Поддерживающий ECN транспорт реагирует на код CE в чистом пакете ACK снижением размера окна перегрузки и такое поведение является проигрышным по сравнению с транспортом без поддержки ECN. По этой причине (и по описанным выше причинам в части повторно передаваемых пакетов), желательно устанавливать код ECT независимо для каждого пакета.

[1] В соответствии с RFC 8311 этот и предыдущий абзац исключены. *Прим. перев.*

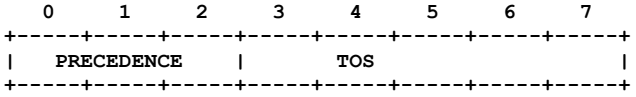
Ещё одним преимуществом двухбитового кода является повышенная отказоустойчивость. Наиболее критический момент, описанный в разделе 8, заключается в том, что по умолчанию следует указывать не поддерживающий ECN транспорт. В двухбитовом варианте для реализации этого требования достаточно просто устанавливать по умолчанию код not-ECT. В однобитовом варианте для выполнения этого требования следует устанавливать код CE или ECT. Этот вариант менее понятен и возможно более открыт для некорректных реализаций на конечных узлах или маршрутизаторах.

Хотя в целом 1-битовая реализация вполне допустима, она имеет ряд существенных недостатков по сравнению с двухбитовым вариантом. Во-первых, функциональность однобитового варианта существенно ограничена в плане трактовки пакетов с кодом CE на втором перегруженном маршрутизаторе. Во-вторых, однобитовый вариант требует передачи дополнительной информации в заголовке транспортного уровня пакетов из поддерживающего ECN потока (функциональность двухбитового варианта просто переносится на транспортный уровень) или понимания со стороны отправителей поддерживающих ECN потоков того, что получатели должны быть способны a-priori определить какие пакеты поддерживают ECN, а какие не поддерживают. В-третьих, однобитовая реализация потенциально более открыта для ошибок со стороны некорректных реализаций, которые могут по умолчанию устанавливать неверное значение бита ECN. Мы полагаем, что перечисленные ограничения обеспечивают достаточные основания использования дополнительного бита в заголовке IP для кодов ECT.

22. Ретроспектива использования октета IPv4 TOS

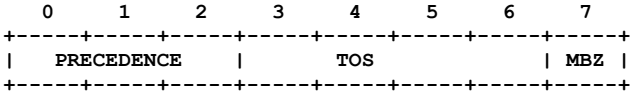


RFC 791 [RFC791] определяет октет ToS[1] в заголовке IP. В RFC 791 биты 6 и 7 октета ToS отмечены, как резервные (Reserved for Future Use), и указано, что они имеют нулевое значение. Первые два поля октета ToS определены в документе, как Precedence (предпочтения) и Type of Service (тип обслуживания - TOS).



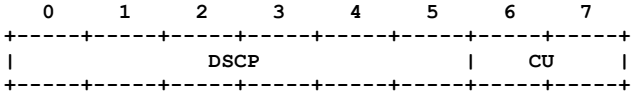
RFC 1122 включает биты 6 и 7 в поле ToS, не обсуждая конкретного их использования (см. рисунок выше).

Октет ToS заголовка IPv4 был заново определён в RFC 1349 [RFC1349], как показано ниже.



Бит 6 поля TOS был определён в RFC 1349, как «Minimize Monetary Cost[2]». В дополнение к полям Precedence и TOS было определено поле MBZ[3], которое в настоящее время не используется. В RFC 1349 отмечено, что отправитель дейтаграмм устанавливает в поле MBZ нулевое значение, если не используется экспериментальных протоколов с иной трактовкой этого бита.

RFC 1455 [RFC 1455] определяет экспериментальный стандарт, использующий все четыре бита поля TOS для запроса гарантированного уровня защиты канала.



Документы RFC 1349 и RFC 1455 были отменены RFC 2474 «Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers[4]» [RFC2474], в котором биты 6 и 7 поля DS были указаны, как неиспользуемые (CU[5]). В RFC 2780 [RFC2780] содержится спецификация ECN для экспериментального использования двухбитового поля CU. RFC 2780 обновляет определение поля DS, оставляя в нем лишь первых шесть битов, которые трактуются, как коды дифференцированного обслуживания (DSCP[6]). Формат поля показан на рисунке выше.

Поскольку история ещё не завершена, определение поля ECN в данном документе не гарантирует совместимости со всеми предшествующими вариантами использования этих двух битов.

До RFC 2474, маршрутизаторам не разрешалось менять биты в полях DSCP и ECN пересылаемых пакетов и, следовательно, маршрутизаторы, соответствующие только требованиям RFC до 2474, не оказывают влияния на ECN. Для конечных узлов бит 7 (второй бит ECN) должен передаваться с нулевым значением всеми реализациями, совместимыми только с RFC до 2474. Такие узлы могут передавать в бите 6 (первый бит ECN) единицу для указания необходимости экономной передачи (Minimize Monetary Cost) в соответствии с RFC 1349 или экспериментами, разрешёнными RFC 1455, однако оба эти варианта не получили широкого распространения. Помехи, которые могут создавать некорректно работающие маршрутизаторы, включают удаление кода CE для поддерживающих ECN пакетов, которые поступили на маршрутизатор с установленным флагом CE, и установку кода CE при отсутствии перегрузок. Эти вопросы рассмотрены в разделе 8. Неподатливость в сети.

Нарушения в работе поддерживающей ECN среды, которые могут создаваться несовместимыми с ECN конечными узлами, передающими пакеты с установленным кодом ECT, описаны в разделе 7. Неподатливость конечных узлов.

23. Взаимодействие с IANA

В этом разделе описаны пространства имён, создаваемые в соответствии с данной спецификацией или выделяемые для данной спецификации значения в существующих пространствах имён, которые управляются IANA.

---

[1]Type of Service - тип обслуживания.
[2]Минимизация финансовых расходов.
[3]Must be zero - должно иметь нулевое значение.
[4]Определение поля дифференцированного обслуживания (DS) в заголовках IPv4 и IPv6.
[5]Currently Unused - в настоящее время не используются.
[6]Differentiated Services CodePoint.

23.1. Байт IPv4 TOS и октет IPv6 Traffic Class

Коды для поля ECN в заголовке IP задаются по процедуре Standards Action данным RFC в соответствии с RFC 2780.

Когда этот документ будет опубликован в качестве RFC, агентству IANA следует создать новый реестр IPv4 TOS Byte and IPv6 Traffic Class Octet[1] с пространством имён IPv4 TOS Byte и IPv6 Traffic Class Octet.

Описание: регистрация идентична для IPv4 и IPv6.

Биты 0-5: см. реестр кодов поля Differentiated Services (http://www.iana.org/assignments/dscp-registry)

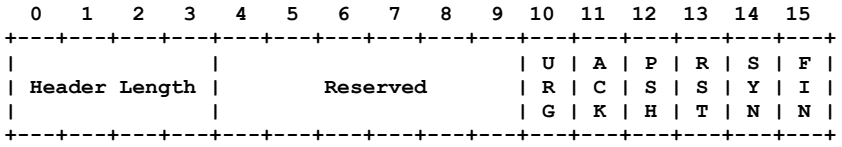
Биты 6-7: поле ECN.

Значение	*Ключевое слово и описание*	*Источник*
00	Not-ECT (не поддерживающий ECN транспорт)	[RFC 3168]
01	ECT(1) (совместимый с ECN транспорт (1))	[RFC 3168]
10	ECT(0) (совместимый с ECN транспорт (0))	[RFC 3168]
11	CE (Наблюдается перегрузка)	[RFC 3168]

23.2. Флаги заголовка TCP

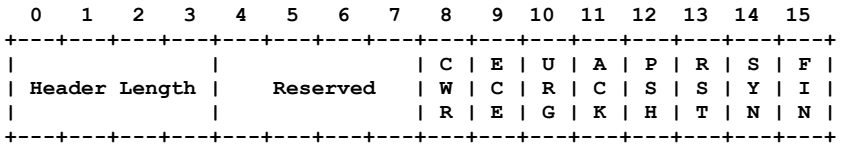
Коды для флагов CWR и ECE в заголовке TCP задаются по процедуре Standards Action в данном RFC, в соответствии с RFC 2780.

При публикации этого документа в качестве RFC агентству IANA следует создать новый реестр TCP Header Flags с пространством имён TCP Header Flags (Флаги заголовка TCP).



Заголовок TCP включает 6-битовое резервное поле, определённое в RFC 793, в байтах 13 и 14, как показано на рисунке выше. Остальные шесть битов управления определены в RFC 793.

RFC 3168 определяет два из шести битов резервного (Reserved) поля для ECN, как показано на рисунке.



Флаги заголовка TCP

Бит	*Имя*	*Источник*
8	CWR (Congestion Window Reduced)	[RFC 3168]
9	ECE (ECN-Echo)	[RFC 3168]

23.3. Атрибуты IPSEC Security Association

Агентство IANA выделило значение атрибута IPSEC Security Association = 10 для ECN Tunnel, как описано в параграфе 9.2.1.2 по запросу David Black в ноябре 1999. В настоящее время агентство IANA сменило ссылку (Reference) для этого выделения с David Black на данный RFC.

24. Адреса авторов

K. K. Ramakrishnan
TeraOptic Networks, Inc.
Phone: +1 (408) 666-8650
EMail: kk@teraoptic.com

Sally Floyd
ACIRI
Phone: +1 (510) 666-2989
EMail: floyd@aciri.org
URL: http://www.aciri.org/floyd/

David L. Black
EMC Corporation
42 South St.
Hopkinton, MA 01748
Phone: +1 (508) 435-1000 x75140
EMail: black_david@emc.com

Перевод на русский язык

Николай Малых
nmalykh@protokols.ru

25. Полное заявление авторских прав

[1] Байт IPv4 TOS и октет IPv6 Traffic Class.

Подтверждение

Финансирование функций RFC Editor обеспечено Internet Society.