

Relaxing Restrictions on Explicit Congestion Notification (ECN) Experimentation

Смягчение требований к явным уведомлениям о перегрузке в экспериментах

Аннотация

Этот документ обновляет RFC 3168, где заданы явные уведомления о перегрузке (Explicit Congestion Notification или ECN) как альтернатива отбрасыванию пакетов для информирования конечных точек о перегрузке в сети. Он смягчает ограничения RFC 3168, препятствующие экспериментам для получения преимуществ за счёт устранения потерь. Документ обобщает предполагаемые направления экспериментов и обновляет RFC 3168, разрешая эксперименты в этих областях. Для получения преимуществ от любого из обновлений требуется выпуск документа Experimental RFC в потоке IETF. Кроме того, этот документ содержит соответствующие обновления спецификаций ECN для RTP в RFC 6679 и протокола контроля перегрузок для дейтаграмм (Datagram Congestion Control Protocol или DCCP) в RFC 4341, 4342 и 5622. Документ также содержит заключение о результатах эксперимента с ECN nonce в RFC 3540 и предоставляет обоснование для перевода RFC 3540 из статуса Experimental в Historic, что позволяет применять в новых экспериментах код ECT(1).

Статус документа

Этот документ является проектом стандарта (Internet Standards Track).

Документ является результатом работы IETF¹ и представляет собой согласованное мнение сообщества IETF. Документ был вынесен на публичное рассмотрение и одобрен для публикации IESG². Дополнительная информация о документах BCP представлена в разделе 2 документа RFC 7841.

Информация о текущем статусе этого документа, обнаруженных ошибках и способах обратной связи может быть найдена по ссылке <https://www.rfc-editor.org/info/rfc8311>.

Авторские права

Авторские права (Copyright (c) 2018) принадлежат IETF Trust и лицам, указанным в качестве авторов документа. Все права защищены.

К документу применимы права и ограничения, указанные в BCP 78 и IETF Trust Legal Provisions и относящиеся к документам IETF (<http://trustee.ietf.org/license-info>), на момент публикации данного документа. Прочтите упомянутые документы внимательно. Фрагменты программного кода, включённые в этот документ, распространяются в соответствии с упрощённой лицензией BSD, как указано в параграфе 4.e документа IETF Trust Legal Provisions, без каких-либо гарантий (как указано в Simplified BSD License).

Документ может содержать материалы из IETF Document или IETF Contribution, опубликованных или публично доступных до 10 ноября 2008 года. Лица, контролирурующие авторские права на некоторые из таких документов, могли не предоставить IETF Trust права разрешать внесение изменений в такие документы за рамками процессов IETF Standards. Без получения соответствующего разрешения от лиц, контролирующих авторские права этот документ не может быть изменён вне рамок процесса IETF Standards, не могут также создаваться производные документы за рамками процесса IETF Standards за исключением форматирования документа для публикации или перевода с английского языка на другие языки.

Оглавление

1. Введение.....	2
1.1. Термины ECN.....	2
1.2. Уровни требований.....	2
2. Обзор экспериментов с ECN.....	2
2.1. Требуется эффективный контроль перегрузок.....	3
2.2. Сетевые соображения для экспериментов с ECN.....	3
2.3. Вопросы эксплуатации и управления.....	3
3. ECN Nonce и RFC 3540.....	3
4. Обновления RFC 3168.....	4
4.1. Различия в реакции на перегрузку.....	4
4.2. Различия в маркировке при перегрузке.....	4
4.3. Пакеты управления и повторная передача TCP.....	5
5. Обновления ECN для RTP.....	6
6. Обновления ECN для DCCP в RFC 4341, 4342 и 5622.....	6
7. Взаимодействие с IANA.....	7
8. Вопросы безопасности.....	7
9. Литература.....	7
9.1. Нормативные документы.....	7

¹Internet Engineering Task Force - комиссия по решению инженерных задач Internet.

²Internet Engineering Steering Group - комиссия по инженерным разработкам Internet.

9.2. Дополнительная литература.....	7
Благодарности.....	8
Адрес автора.....	8

1. Введение

Этот документ обновляет RFC 3168, где заданы явные уведомления о перегрузке (Explicit Congestion Notification или ECN) как альтернатива отбрасыванию пакетов для информирования конечных точек о перегрузке в сети. Он смягчает ограничения RFC 3168, препятствующие экспериментам для получения преимуществ за счёт устранения потерь. Документ обобщает предполагаемые направления экспериментов и обновляет RFC 3168, разрешая эксперименты в этих областях. Для получения преимуществ от любого из обновлений требуется выпуск документа Experimental RFC в потоке IETF [RFC4844]. Включение этих обновлений в один документ позволяет выполнять эксперименты без изменения стандартного процесса для каждого Experimental RFC, требующего изменять RFC 3168.

Этому документу не требуется обновлять RFC 3168 для упрощения стандартизации протоколов и механизмов, описанных в Standards Track RFC, поскольку такие RFC могут обновлять RFC 3168 напрямую, не полагаясь на обновления из этого документа или исключения из процесса стандартизации.

Кроме того, этот документ содержит соответствующие обновления спецификаций ECN для RTP [RFC6679] и 3 профилей DCCP ([RFC4341], [RFC4342] [RFC5622]) по тем же причинам. Каждый эксперимент по-прежнему нужно документировать в одном или нескольких RFC, но использование Experimental RFC для этой цели не требует исключения из процесса для изменения любого из этих Proposed Standard RFC, когда изменения остаются в рамках, заданных этим документом (RFC 5622 является Experimental RFC и обновляется этим документом для согласования с другими RFC, относящимися к DCCP).

Некоторые из предполагаемых экспериментов включают использование кода ECT(1), который был выделен для экспериментов с ECN поспе в RFC 3540 [RFC3540]. Этот документ информирует о завершении экспериментов с ECN и меняет статус RFC 3540 с Experimental на Historic, чтобы разрешить новые эксперименты с кодом ECT(1).

1.1. Термины ECN

ECT

Транспорт с поддержкой ECN (ECN-Capable Transport). Один из кодов ECT(0) или ECT(1) в поле ECN [RFC3168] заголовка IP (v4 или v6). Отправитель с поддержкой ECN устанавливает такой код для индикации поддержки ECN обеими транспортными конечными точками.

Not-ECT

Код ECN, устанавливаемый отправителями для указания отсутствия поддержки ECN в транспорте.

CE

Возникла перегрузка (Congestion Experienced). Этот код ECN устанавливают промежуточные узлы для индикации перегрузки. Узел устанавливает больше маркеров CE в пакетах ECT по мере роста перегрузки.

1.2. Уровни требований

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **не рекомендуется** (NOT RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с BCP 14 [RFC2119] [RFC8174] тогда и только тогда, когда они выделены шрифтом, как показано здесь.

2. Обзор экспериментов с ECN

В этом документе рассматриваются три типа экспериментов с ECN. Для получения более подробных сведений о целях и обосновании каждого эксперимента следует обращаться к цитируемым документам.

Различия в реакции на перегрузку

Уведомление о перегрузке ECN сообщает о более высокой вероятности наличия коротких очередей в узких местах сети (bottleneck), нежели отбрасывания пакетов [TCP-ABE]. Это предполагает, что для перегрузки, указываемой ECN, может быть уместна реакция отправителя, отличная от отклика на потерю пакетов (например, более медленное снижение скорости). Два примера предлагаемого изменения реакции отправителей описаны в [TCP-ABE] и [ECN-L4S]. Второе предложение связывает изменение отклика отправителя на перегрузку с функциональностью различий в маркировке при перегрузке (Congestion Marking Differences, см. ниже). Эти изменения противоречат требованию RFC 3168 считать индикацию перегрузки ECN эквивалентом отбрасывания пакета. Требуется одобрение IETF (например, через Experimental RFC в потоке документов IETF для любых откликов отправителя, проводимых в этом направлении экспериментов (см. 4.1. Различия в реакции на перегрузку).

Различия в маркировке при перегрузке

Маркировку перегрузки на узлах сети можно настроить на поддержку очень небольших очередей в сочетании с другой реакцией отправителя на индикацию перегрузки (CE), например, как предложено в [ECN-L4S]. Вовлечённый трафик нужно идентифицировать отправителю, чтобы избежать ущерба для другого трафика, отправители которого не ожидают более частых маркеров перегрузки, служащих для поддержки очень малых очередей. Использование различных кодов ECN, в частности ECT(0) и ECT(1), является многообещающим способом идентификации, но это противоречит требованиям RFC 3168 не предоставлять трафику с маркером ECT(0) обработку в сети, отличающуюся от обработки трафика с маркером ECT(1). Требуется одобрение IETF (например, через Experimental RFC в потоке документов IETF для любых откликов отправителя, проводимых в этом направлении экспериментов (см. 4.2. Различия в маркировке при перегрузке).

Пакеты управления и повтор передачи TCP

RFC 3168 ограничивает применение ECN для TCP пакетами данных, исключая повторы передачи. Успешное внедрение ECN в больших частях Internet вызвало интерес к добавлению достигаемых преимуществ ECN для пакетов управления TCP (например, SYN) и повторно передаваемых пакетов, например, как предложено в [ECN-TCP]. Это противоречит запрету RFC 3168 на применение ECN и пакетам управления и повторной передаче TCP (см. 4.3. Пакеты управления и повторная передача TCP).

Действие этого документа ограничено тремя направлениями экспериментирования. Документ не выражает какого-либо мнения о вероятных результатах и не задаёт деталей экспериментов. Возможны дополнительные эксперименты в этих направлениях, например, по использованию ECN для поддержки внедрения протоколов, подобных Data Center TCP (DCTCP) [RFC8257] за пределами текущего применения DCTCP в среде ЦОД. Целью документа является устранение ограничения Standards Track RFC, стоящих на пути экспериментов в отмеченных областях.

2.1. Требуется эффективный контроль перегрузок

Контроль перегрузок остаётся важным аспектом архитектуры Internet [RFC2914]. Любые Experimental RFC в потоке документов IETF, использующие обновления этого документа для какого-либо RFC должны включать рассмотрение влияния экспериментов по контролю перегрузки для обеспечения уверенности в том, что развёртывание эксперимента не создаст связанных с перегрузкой проблем в работе Internet.

2.2. Сетевые соображения для экспериментов с ECN

Функциональность ECN [RFC3168] получила широкое распространение в Internet и реализуется также в форме дополнительных протоколов, таких как TRILL¹ [ECN-TRILL]. Предполагается сосуществование экспериментов с ECN и развёрнутой функциональности ECN и ответственность за это ложится в первую очередь на разработчиков экспериментальных изменений в ECN. Кроме того, разработчики и создатели протоколов, а также операторы сетей могут запланировать или выполнить эксперименты с ECN. Приведённые ниже рекомендации помогут избежать конфликтов при экспериментах в областях, разрешённых этим документом.

1. Описанное в RFC 3168 поведение при пересылке остаётся предпочтительным для маршрутизаторов, не участвующих в экспериментах с ECN, продолжая, в частности, считать эквивалентными коды ECT(0) и ECT(1), как указано в параграфе 4.2. Различия в маркировке при перегрузке.
2. Пересылающим пакеты узлам сети **не следует** считать, что код ECN CE говорит об отбрасывании пакета, если бы не было ECN. Это обусловлено тем, что эксперименты по различию в откликах на перегрузку (Congestion Response Differences) предполагают разную реакцию на отбрасывание пакетов и приём пакетов с маркером CE (4.1. Различия в реакции на перегрузку), поэтому пакеты с маркером CE **не следует** произвольно отбрасывать. Соответствующее различие в откликах на перегрузку уже имеется при использовании поля ECN для индикации начинающейся перегрузки (Pre-Congestion Notification или PCN) [RFC6660].
3. Узлу сети **недопустимо** создавать трафик с маркировкой ECT(1), пока этот узел не участвует в эксперименте по различной маркировке (Congestion Marking Differences) с использованием ECT(1), как указано в параграфе 4.2. Различия в маркировке при перегрузке.

В некоторых экспериментах ECN применяется для пакетов, где раньше уведомления ECN не использовались, в частности, это пакеты управления и повторные пакеты TCP (см. 4.3. Пакеты управления и повторная передача TCP). Новые требования к поведению промежуточных устройств, в данном разделе имеют особое значение. В общем случае любым системам и протоколам, отслеживающим и инспектирующим сетевой трафик **следует** быть готовыми встретить ECN в пакетах и трафике, которые в настоящее время не используют ECN.

Требования к обработке полей ECN при туннельной инкапсуляции и декапсуляции заданы в [RFC6040], обновлению которого посвящён документ [ECN-SHIM]. Руководство для инкапсуляции с внешними заголовками, отличными от IP, приведено в [ECN-ENCAP]. Эти требования и рекомендации применимы ко всему трафику, включая связанный с любыми экспериментами ECN.

2.3. Вопросы эксплуатации и управления

Изменения в поведении сетевого трафика в результате экспериментов с ECN, вероятно, повлияют на работу сети и управление ею. Предполагается, что разработчики экспериментов с ECN предвидят возможное влияние и продумают, как иметь с ним дело. Конкретные темы для рассмотрения включают возможные изменения и расширения для управления сетью, мониторинг экспериментального развёртывания, сбор данных для оценки экспериментов, возможные взаимодействия с другими протоколами, в частности, с протоколами инкапсуляции сетевого трафика.

В [RFC5706] рассматриваются эти вопросы, а Приложение А к RFC 5706 содержит краткий обзор некоторых важных аспектов.

3. ECN Nonce и RFC 3540

Как задано в RFC 3168, ECN использует 2 кода транспорта с поддержкой ECN (ECT) - ECT(0) и ECT(1) - для индикации поддержки ECN для пакета. В RFC 3168 код ECT(1) выделен для поддержки функциональности ECN nonce с целью предотвращения использования ECN для повышения своей пропускной способности за счёт других пользователей сети. Функциональность ECN nonce полностью задана в RFC 3540 [RFC3540]. В этом разделе указаны причины смены статуса RFC 3540 с Experimental на Historic (устаревший) и внесены соответствующие обновления для RFC 3168.

Хотя ECN nonce работает в соответствии со спецификацией и применяется в ограниченных средах, широкого распространения в Internet этот подход не получил. Изучение поведения ECN для 1 миллиона web-серверов с использованием данных 2014 г. [Trammell15] показало, что после согласования ECN ни один из 581711 протестированных серверов IPv4 не использовал оба кода ECT, что говорило бы о применении ECN nonce. Из 17028 протестированных серверов IPv6 только 4 устанавливали коды ECT(0) и ECT(1) в пакетах данных. Это могло говорить о применении ECN этими 4 серверами, но в равной мере могло быть связано с ошибочной перемаркировкой поля ECN промежуточным устройством или маршрутизатором.

С появлением новых экспериментальных функций, зависящих от применения кода ECT(1), дальнейшее резервирование этого кода для экспериментов с ECN nonce перестало быть оправданным. Кроме того, появились другие подходы для предотвращения злоупотреблений ECN (см. Приложение С.1² к [ECN-L4S]). Поэтому для поддержки экспериментов ECN с кодом ECT(1) этот документ вносит указанные ниже изменения.

¹Transparent Interconnection of Lots of Links - прозрачные соединения большого числа каналов.

²В оригинале ошибочно указано Приложение В.1, см. <https://www.rfc-editor.org/errata/eid5649>. Прим перев.

- Эксперимент с ECN nonce [RFC3540] считать завершённым и отметить отсутствие широкого внедрения.
- Обновить RFC 3168 [RFC3168] для исключения обсуждения ECN nonce и применение для этого кода ECT(1).

Ниже приведены 4 основных обновления RFC 3168, исключающие обсуждение ECN nonce и использование ECT(1).

1. Исключён абзац из раздела 5 после рисунка 1, где обсуждается ECN nonce как мотив для 2 кодов ECT.
2. Полностью удалён параграф 11.2 «Обсуждение ECN nonce».
3. Удалён последний абзац раздела 12, где сказано, что ECT(1) может быть частью реализации ECN nonce.
4. Удалены два первых абзаца параграфа 20.2, где обсуждаются ECN nonce и альтернативы. Остальная часть параграфа 20.2 с обсуждением использования 4 кодов ECN сохранена.

Кроме того, в RFC 3168 нужно внести менее важные изменения, удалив все упоминания ECN nonce и устранив последствия выделения ECT(1) для ECN nonce. Для краткости они не включены в документ.

4. Обновления RFC 3168

В последующих параграфах приведены обновления RFC 3168, позволяющие экспериментировать в трёх областях, указанных в разделе 2. Обзор экспериментов с ECN.

4.1. Различия в реакции на перегрузку

RFC 3168 задаёт идентичную реакцию отправителя на отбрасывание пакетов и индикация перегрузки ECN. Маркеры перегрузки ECN устанавливают в основном механизмы активного управления очередями (Active Queue Management или AQM) в промежуточных буферах. Обычно эти механизмы настраиваются на поддержку более коротких очередей по сравнению с механизмами без AQM, в частности, основанные на отбрасывании механизмы без AQM (такие как «обрубание хвостов» - tail-drop), поскольку механизмы AQM указывают перегрузку до переполнения очереди. Хотя наличие потерь не позволяет отправителю легко понять, применяется ли AQM, получение маркера ECN CE указывает большую вероятность использования AQM для управления очередью в узком месте. Поэтому индикация перегрузки ECN с большей вероятностью (нежели отбрасывание пакета) говорит о короткой очереди в узком месте сети [TCP-ABE]. Это различие предполагает, что на указанную ECN перегрузку отправитель может реагировать иначе (например, выполнять меньший откат), нежели на указание перегрузки потерей пакета. Однако в разделе 5 RFC 3168 сказано:

«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 отбрасывания одного пакета.»

Этот документ обновляет приведённый текст RFC 3168, чтобы отклик контроля перегрузки (включая отклик отправителя TCP) на пакет с маркером CE мог отличаться от реакции на потерю пакета, при условии, что отличия от RFC 3168 документированы в Experimental RFC потока документов IETF. Конкретным изменением RFC 3168 является вставка слов «если иное не указано в Experimental RFC потока документов IETF» в конце приведённого выше предложения.

В RFC 4774 [RFC4774] включён приведённый выше текст из RFC 3168 в качестве основы, но не задаётся требований на базе этого фрагмента. Поэтому для разрешения экспериментов обновлять RFC 4774 не требуется.

В параграфе 6.1.2 RFC 3168 сказано:

«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 насыщение. Индикацию насыщения следует трактовать просто, как потерю в результате насыщения для TCP без поддержки ECN. Т. е. отправитель TCP снижает вдвое размер окна насыщения cwnd и уменьшает порог замедленного старта ssthresh.»

Этот документ обновляет приведённый текст из RFC 3168, чтобы разрешить отклик контроля перегрузок (включая отправителя TCP) на пакет с маркером CE, отличный от реакции на отбрасывание пакета, при условии документирования отличий от RFC 3168 в Experimental RFC потока документов IETF. Конкретным изменением RFC 3168 является вставка слов: «Если не указано иное в Experimental RFC потока документов IETF» в конце второго предложения, процитированного выше.

4.2. Различия в маркировке при перегрузке

Доведённый до предела алгоритм AQM с индикацией контроля перегрузки ECN можно настроить на поддержку очень мелких очередей, снижая тем самым задержку в сети по сравнению с поддержкой больших очередей. Для эффективного использования мелких очередей нужны значительно более энергичные отклики отправителей на ECN. Примером такого подхода является протокол Datacenter TCP (DCTCP) [RFC8257]. В этом случае нужна отдельная обработка узлами сети для предотвращения препятствий со стороны энергичного трафика с малыми задержками для обычного трафика (при наличии) и препятствий со стороны обычного трафика любым службам с малыми задержками, использующим мелкие очереди. Использование разных кодов ECN является многообещающим способом идентификации двух отмеченных классов трафика на узлах сети, поэтому данная область исследований основана на применении кода ECT(1) для запроса поведения маркировки ECN в сети, отличающегося от случая ECT(0). Важно, чтобы такое изменение поведения маркировки ECN уравнивалось использованием иного одобренного IETF отклика отправителя на маркировку CE, например, как предложено в [ECN-L4S].

В разделе 5 RFC 3168 сказано: «Маршрутизаторы трактуют коды ECT(0) и ECT(1) одинаково.»

Этот документ обновляет RFC 3168, чтобы разрешить маршрутизаторам разную трактовку кодов ECT(0) и ECT(1) при условии, что отличия от RFC 3168 документированы в Experimental RFC потока документов IETF. Конкретным изменением RFC 3168 является вставка слов: «Если не указано иное в Experimental RFC потока документов IETF» в конце предложения, процитированного выше.

При настройке AQM на использование контроля перегрузки ECN для поддержки очень мелкой очереди, маркеры перегрузки устанавливаются для пакетов, которые не обязательно отбрасывались бы при отсутствии ECN.

В разделе 5 RFC 3168 сказано:

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

Этот документ обновляет RFC 3168, чтобы разрешить индикацию перегрузки, не эквивалентную отбрасыванию, при условии, что отличия от RFC 3168 документированы в Experimental RFC потока документов IETF. Конкретным изменением RFC 3168 является замена слова «Маршрутизаторам» в начале предложения словами: «Если не указано иное в Experimental RFC потока документов IETF»

Требуется более масштабное обновление RFC 3168, чтобы разрешить отправителю использовать код ECT(1) для запроса поведения маркировки перегрузки, поддерживающего очень мелкие очереди на узлах сети. При использовании потери в качестве сигнала перегрузки число передаваемых сигналов следует сводить к минимуму, поскольку передаются лишь сведения о наличии или отсутствии перегрузки. ECN может обеспечить более детальную сигнализацию, например, для указания текущего уровня перегрузки, без недостатка, связанного с потерей большого числа пакетов. Предложенный в этой области эксперимент L4S¹ [ECN-L4S] существенно повышает вероятность маркировки CE для трафика, помеченного кодом ECT(1), способом, который будет негативно влиять на имеющуюся функциональность реагирования на перегрузку, поскольку эта функциональность предполагает, что сеть маркирует пакеты ECT с частотой, с которой бы отбрасывались пакеты без ECT. Если сетевой трафик, использующий традиционные отклики отправки на перегрузку столкнется с ростом вероятности (и частоты) маркировки L4S в очереди узкого места сети, пропускная способность трафика будет, вероятно, гораздо меньше, чем предполагалось для уровня перегрузки в узком месте.

Этот документ обновляет RFC 3168, исключая такое взаимодействие для ECT(1). Конкретное обновление раздела 5 в RFC 3168 заключается в замене текста:

«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 пакетом с маркером CE в 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)».

Новый текст:

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.

Экспериментам по различной маркировке перегрузки **следует** менять поведение сети для трафика с маркерами ECT(1), а не ECT(0), если поведение сети меняется лишь для одного кода ECT. Экспериментам по различной маркировке перегрузки **недопустимо** менять поведение сети для трафика с маркерами ECT(0) так, чтобы требовалось менять реакцию отправителя на перегрузку для обеспечения желаемого поведения сети. Если эксперимент по различной маркировке перегрузки меняет поведение сети для трафика с маркерами ECT(1), например, поведение маркировки CE-marking так, что требуется менять реакцию отправителя на перегрузку для обеспечения желаемого поведения сети, документ Experimental RFC в потоке IETF для такого эксперимента **должен** задавать:

- отклик отправителя на маркировку CE в сети;
- изменения в поведении маршрутизаторов (или их отсутствие) при пересылке пакетов с маркером CE, являющихся частью эксперимента.

Кроме того, этот документ исключает из RFC 3168 обсуждение ECN nonce, как отмечено выше в разделе 3.

4.3. Пакеты управления и повторная передача TCP

Успешное применение ECN для трафика во многих частях Internet вызвало интерес к распространению преимуществ ECN на пакеты управления TCP (например, SYN) и повторно передаваемые пакеты, как предложено, например, в ECN++ [ECN-TCP].

RFC 3168 запрещает использовать ECN для пакетов управления TCP и повторно передаваемых пакетов во многих местах.

- Параграф 5.2: «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.»
- Параграф 6.1.1: «Для хоста **недопустимо** устанавливать ECT в пакетах SYN или SYN-ACK.»
- Параграф 6.1.4: «...чистые пакеты подтверждения (т. е. пакеты, содержащие только подтверждение без дополнительных данных) **должны** передаваться с кодом not-ECT.»

¹Low Latency Low Loss Scalable throughput - малые задержки и потери, расширяемая пропускная способность.

- Параграф 6.1.5: «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.»
- Параграф 6.1.6: «...отправителю TCP **недопустимо** устанавливать код ECT или флаг CWR в пробных пакетах.»

Этот документ обновляет RFC 3168, чтобы разрешить использование кодов ECT в пакетах SYN и SYN-ACK, «чистых» пакетах подтверждения, пакетах зондирования окна и повторно передаваемых пакетах, которые изначально были переданы с кодом ECT, если отличия от RFC 3168 документированы в Experimental RFC потока IETF. Конкретным изменением RFC 3168 является вставка в конце приведённых выше предложений слов: «если не указано иное в Experimental RFC потока документов IETF».

Помимо требования к отправителям TCP не устанавливать ECT для пакетов управления TCP и передаваемых повторно пакетов в RFC 3168 ничего не сказано об уместности отбрасывания такого пакета элементами сети (например, межсетевыми экранами) как непригодного. Чтобы эксперименты с ECN в этом направлении были полезны, промежуточные устройства не должны отбрасывать такие пакеты. Поэтому в конце параграфа 6.1.1.1 RFC 3168 добавляется приведённый ниже текст.

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 тому, 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.

5. Обновления ECN для RTP

RFC 6679 [RFC6679] определяет применение ECN для трафика RTP. Документ позволяет использовать оба кода ECT(0) и ECT(1), а также содержит рекомендации по их применению в параграфе 7.3.1.

Отправителю **следует** помечать пакеты кодом ECT(0), если получатель не указал предпочтение ECT(1) или случайному значению ECT в параметре `ect` атрибута `a=ecn-capable-rtcp`.

Эксперименты с разной маркировкой перегрузки расширяют возможные последствия применения ECT(1) вместо ECT(0), поэтому приведённая выше рекомендация обновляется двумя указанными ниже предложениями.

Недопустимо применять случайные значения ECT, поскольку это может вызывать для RTP разную обработку в сети пакетов с маркерами ECT(1) и ECT(0) и различной реакции конечных точек на перегрузку. **Требуется** использовать код ECT(0), если не указано иное в Experimental RFC потока документов IETF.

В параграфе 7.3.3 RFC 6679 задана реакция RTP на получение пакетов с маркером CE как идентичная реакции на отбрасывание пакетов.

Получение пакетов RTP с маркером ECN-CE в заголовке IP является уведомлением о возникновении перегрузки. Принятой по умолчанию реакцией на пакеты с маркером ECN-CE **должно** быть уведомление алгоритма контроля перегрузки, вызывающее такой же отклик, как в случае потери пакета. Не следует по-разному обрабатывать маркировку ECN-CE и обнаружение отбрасывания пакета.

Для поддержки экспериментов с различными откликами на перегрузку этот документ обновляет текст аналогично RFC 3168, чтобы алгоритм контроля перегрузки RTP мог по-разному отвечать на пакеты с маркером CE и отбрасывание пакетов при условии, что отличия от RFC 6679 указаны в Experimental RFC потока документов IETF. Конкретным изменением RFC 6679 является вставка слов: «Если не указано иное в Experimental RFC потока документов IETF» и реформатирование двух приведённых выше предложений в соответствии с этим условием.

Получение пакетов RTP с маркером ECN-CE в заголовке IP является уведомлением о возникновении перегрузки. Если не указано иное в Experimental RFC потока документов IETF:

- принятой по умолчанию реакцией на пакеты с маркером ECN-CE **должно** быть уведомление алгоритма контроля перегрузки, вызывающее такой же отклик, как в случае потери пакета;
- не следует по-разному обрабатывать маркировку ECN-CE и обнаружение отбрасывания пакета.

Во втором предложении следующего абзаца параграфа 7.3.3 в RFC 6679:

«В будущем может быть задана иная реакция на ECN-CE после IETF Review. Подробное описание таких реакций **должно** приводиться в Standards Track RFC и рецензироваться для гарантий безопасного внедрения при соблюдении заданных ограничений.»

Слова «Standards Track RFC» заменяются фразой «Standards Track RFC или Experimental RFC в потоке документов IETF» для согласованности с первым обновлением.

6. Обновления ECN для DCCP в RFC 4341, 4342 и 5622

Спецификации трёх идентификаторов контроля перегрузки DCCP (DCCP Congestion Control ID или CCID) - 2 [RFC4341], 3 [RFC4342], 4 [RFC5622] - содержат в целом общую формулировку:

«каждый пакет DCCP-Data и DCCP-DataAck передаётся с полем ECN Capable, имеющим значение ECT(0) или ECT(1).»

Этот документ обновляет такие предложения во всех 3 RFC, как показано ниже:

«каждый пакет DCCP-Data и DCCP-DataAck передаётся с полем ECN Capable. Если не указано иное в Experimental RFC потока документов IETF, такие отправители DCCP **должны** указывать код ECT(0).»

Для поддержки экспериментов с различными откликами на перегрузку (как указано в разделе 3) этот документ обновляет упомянутые RFC, удаляя обсуждение ECN nonce. Для краткости текст здесь не приводится.

7. Взаимодействие с IANA

Чтобы отразить смену статуса RFC 3540 на Historic, агентство IANA обновило реестр Transmission Control Protocol (TCP) Header Flags <<https://www.iana.org/assignments/tcp-header-flags>> удалив регистрацию бита 7 как NS (Nonce Sum) и добавив в реестр слова о том, что бит 7 использовался устаревшим (Historic) RFC 3540 как NS (Nonce Sum), а сейчас является резервным (Reserved).

Для отражения экспериментального использования ECT(1), предложенного в этом документе, агентство IANA добавило для поля ECN в реестре <<https://www.iana.org/assignments/dscp-registry/dscp-registry.xhtml#ecn-field>> примечание:

Код ECT(1) предназначен лишь для экспериментов [RFC8311, параграф 4.2]¹

8. Вопросы безопасности

Документ лишь смягчает требования к экспериментам и не меняет состояния безопасности протоколов, поэтому соображения безопасности рассматриваются в документах, предлагающих эксперименты.

Однако эффективный контроль перегрузок критически важен для работы Internet, поэтому данный документ передаёт ответственность за сохранение контроля перегрузки в Internet экспериментаторам и предложившим эксперимент разработчикам. Эта ответственность включает требование рассматривать влияние на контроль перегрузок в Experimental RFC потока IETF для каждого такого эксперимента, как указано в параграфе 2.1, а также обзор (рецензирование) сообществом IETF и IESG до публикации RFC для обеспечения гарантии того, что эксперимент не нарушает контроль перегрузок в Internet.

Альтернативы ECN nonce рассмотрены в Приложении С.1 к [ECN-L4S].

9. Литература

9.1. Нормативные документы

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, [RFC 2119](#), DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, [RFC 2914](#), DOI 10.17487/RFC2914, September 2000, <<https://www.rfc-editor.org/info/rfc2914>>.
- [RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", [RFC 3168](#), DOI 10.17487/RFC3168, September 2001, <<https://www.rfc-editor.org/info/rfc3168>>.
- [RFC3540] Spring, N., Wetherall, D., and D. Ely, "Robust Explicit Congestion Notification (ECN) Signaling with Nonces", [RFC 3540](#), DOI 10.17487/RFC3540, June 2003, <<https://www.rfc-editor.org/info/rfc3540>>.
- [RFC4341] Floyd, S. and E. Kohler, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 2: TCP-like Congestion Control", RFC 4341, DOI 10.17487/RFC4341, March 2006, <<https://www.rfc-editor.org/info/rfc4341>>.
- [RFC4342] Floyd, S., Kohler, E., and J. Padhye, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 3: TCP-Friendly Rate Control (TFRC)", RFC 4342, DOI 10.17487/RFC4342, March 2006, <<https://www.rfc-editor.org/info/rfc4342>>.
- [RFC5622] Floyd, S. and E. Kohler, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion ID 4: TCP-Friendly Rate Control for Small Packets (TFRC-SP)", RFC 5622, DOI 10.17487/RFC5622, August 2009, <<https://www.rfc-editor.org/info/rfc5622>>.
- [RFC6679] Westerlund, M., Johansson, I., Perkins, C., O'Hanlon, P., and K. Carlberg, "Explicit Congestion Notification (ECN) for RTP over UDP", RFC 6679, DOI 10.17487/RFC6679, August 2012, <<https://www.rfc-editor.org/info/rfc6679>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, [RFC 8174](#), DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

9.2. Дополнительная литература

- [ECN-ENCAP] Briscoe, B., Kaippallimalil, J., and P. Thaler, "Guidelines for Adding Congestion Notification to Protocols that Encapsulate IP", Work in Progress, draft-ietf-tsvwg-ecn-encap-guidelines-09, July 2017.
- [ECN-EXPERIMENT] Khademi, N., Welzl, M., Armitage, G., and G. Fairhurst, "Updating the Explicit Congestion Notification (ECN) Specification to Allow IETF Experimentation", Work in Progress, draft-khademi-tsvwg-ecn-response-01, July 2016.
- [ECN-L4S] Schepper, K. and B. Briscoe, "Identifying Modified Explicit Congestion Notification (ECN) Semantics for Ultra-Low Queuing Delay", Work in Progress², draft-ietf-tsvwg-ecn-l4s-id-01, October 2017.
- [ECN-SHIM] Briscoe, B., "Propagating Explicit Congestion Notification Across IP Tunnel Headers Separated by a Shim", Work in Progress, draft-ietf-tsvwg-rfc6040update-shim-05, November 2017.
- [ECN-TCP] Bagnulo, M. and B. Briscoe, "ECN++: Adding Explicit Congestion Notification (ECN) to TCP Control Packets", Work in Progress, draft-ietf-tcpm-generalized-ecn-02, October 2017.

¹В оригинале это примечание и предыдущий абзац отсутствуют, см. <https://www.rfc-editor.org/errata/eid5399>. Прим. перев.

²Опубликовано в [RFC 9331](#). Прим. перев.

- [ECN-TRILL] Eastlake, D. and B. Briscoe, "TRILL: ECN (Explicit Congestion Notification) Support", Work in Progress, draft-ietf-trill-ecn-support-04, November 2017.
- [RFC4774] Floyd, S., "Specifying Alternate Semantics for the Explicit Congestion Notification (ECN) Field", BCP 124, RFC 4774, DOI 10.17487/RFC4774, November 2006, <<https://www.rfc-editor.org/info/rfc4774>>.
- [RFC4844] Daigle, L., Ed. and Internet Architecture Board, "The RFC Series and RFC Editor", RFC 4844, DOI 10.17487/RFC4844, July 2007, <<https://www.rfc-editor.org/info/rfc4844>>.
- [RFC5706] Harrington, D., "Guidelines for Considering Operations and Management of New Protocols and Protocol Extensions", RFC 5706, DOI 10.17487/RFC5706, November 2009, <<https://www.rfc-editor.org/info/rfc5706>>.
- [RFC6040] Briscoe, B., "Tunnelling of Explicit Congestion Notification", RFC 6040, DOI 10.17487/RFC6040, November 2010, <<https://www.rfc-editor.org/info/rfc6040>>.
- [RFC6660] Briscoe, B., Moncaster, T., and M. Menth, "Encoding Three Pre-Congestion Notification (PCN) States in the IP Header Using a Single Diffserv Codepoint (DSCP)", RFC 6660, DOI 10.17487/RFC6660, July 2012, <<https://www.rfc-editor.org/info/rfc6660>>.
- [RFC8257] Bensley, S., Thaler, D., Balasubramanian, P., Eggert, L., and G. Judd, "Data Center TCP (DCTCP): TCP Congestion Control for Data Centers", RFC 8257, DOI 10.17487/RFC8257, October 2017, <<https://www.rfc-editor.org/info/rfc8257>>.
- [TCP-ABE] Khademi, N., Welzl, M., Armitage, G., and G. Fairhurst, "TCP Alternative Backoff with ECN (ABE)", Work in Progress¹, draft-ietf-tcpm-alternativebackoff-ecn-05, December 2017.
- [Trammell15] Trammell, B., Kuehlewind, M., Boppart, D., Learmonth, I., Fairhurst, G., and R. Scheffenegger, "Enabling Internet-Wide Deployment of Explicit Congestion Notification", In Conference Proceedings of Passive and Active Measurement (PAM), pp. 193-205, March 2015, <https://doi.org/10.1007/978-3-319-15509-8_15>.

Благодарности

Содержимое этой спецификации, включая обновлённые части RFC 3168, в значительной мере заимствовано из [ECN-EXPERIMENT] - спасибо авторам этого документа. Авторы документов, описывающих эксперименты, дали мотивы для создания этого документа, спасибо им за интерес к инновациям. Colin Perkins предложил обновление RFC 6679 для RTP и предоставил рекомендации по обновляемым фрагментам.

Эта спецификация стала лучше, благодаря комментариям множества рецензентов, включая Ben Campbell, Brian Carpenter, Benoit Claise, Spencer Dawkins, Gory Fairhurst, Sue Hares, Ingemar Johansson, Naeem Khademi, Mirja Kuehlewind, Karen Nielsen, Hilarie Orman, Eric Rescorla, Adam Roach, Michael Welzl. Внимательный обзор Bob Briscoe для множества черновых вариантов этого документа привёл к многочисленным улучшениям, включая добавление изменений в RFC для протокола DCCP.

Адрес автора

David Black
Dell EMC
176 South Street
Hopkinton, MA 01748
United States of America
Email: david.black@dell.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

¹Опубликовано в RFC 8511. Прим. перев.