

Transport Area Working Group  
Internet-Draft  
Updates: [3168](#), [4341](#), [4342](#), [5622](#), [6679](#)  
(if approved)  
Intended status: Standards Track  
Expires: April 23, 2018

D. Black  
Dell EMC  
October 20, 2017

**Relaxing Restrictions on Explicit Congestion Notification (ECN)  
Experimentation  
draft-ietf-tsvwg-ecn-experimentation-07**

**Abstract**

This memo updates [RFC 3168](#), which specifies Explicit Congestion Notification (ECN) as an alternative to packet drops for indicating network congestion to endpoints. It relaxes restrictions in [RFC 3168](#) that hinder experimentation towards benefits beyond just removal of loss. This memo summarizes the anticipated areas of experimentation and updates [RFC 3168](#) to enable experimentation in these areas. An Experimental RFC in the IETF document stream is required to take advantage of any of these enabling updates. In addition, this memo makes related updates to the ECN specifications for RTP in [RFC 6679](#) and for DCCP in [RFC 4341](#), [RFC 4342](#) and [RFC 5622](#). This memo also records the conclusion of the ECN nonce experiment in [RFC 3540](#), and provides the rationale for reclassification of [RFC 3540](#) as Historic; this reclassification enables new experimental use of the ECT(1) codepoint.

**Status of This Memo**

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 23, 2018.

## Copyright Notice

Copyright (c) 2017 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

## Table of Contents

<a href="#">1.</a>	Introduction . . . . .	<a href="#">3</a>
<a href="#">1.1.</a>	ECN Terminology . . . . .	<a href="#">3</a>
<a href="#">1.2.</a>	Requirements Language . . . . .	<a href="#">4</a>
<a href="#">2.</a>	ECN Experimentation: Overview . . . . .	<a href="#">4</a>
<a href="#">2.1.</a>	Effective Congestion Control is Required . . . . .	<a href="#">5</a>
<a href="#">2.2.</a>	Considerations for Other Protocols . . . . .	<a href="#">5</a>
<a href="#">2.3.</a>	Operational and Management Considerations . . . . .	<a href="#">6</a>
<a href="#">3.</a>	ECN Nonce and <a href="#">RFC 3540</a> . . . . .	<a href="#">7</a>
<a href="#">4.</a>	Updates to <a href="#">RFC 3168</a> . . . . .	<a href="#">8</a>
<a href="#">4.1.</a>	Congestion Response Differences . . . . .	<a href="#">8</a>
<a href="#">4.2.</a>	Congestion Marking Differences . . . . .	<a href="#">9</a>
<a href="#">4.3.</a>	TCP Control Packets and Retransmissions . . . . .	<a href="#">12</a>
<a href="#">5.</a>	ECN for RTP Updates to <a href="#">RFC 6679</a> . . . . .	<a href="#">13</a>
<a href="#">6.</a>	ECN for DCCP Updates to RFCs 4341, 4342 and 5622 . . . . .	<a href="#">14</a>
<a href="#">7.</a>	Acknowledgements . . . . .	<a href="#">15</a>
<a href="#">8.</a>	IANA Considerations . . . . .	<a href="#">15</a>
<a href="#">9.</a>	Security Considerations . . . . .	<a href="#">16</a>
<a href="#">10.</a>	References . . . . .	<a href="#">16</a>
<a href="#">10.1.</a>	Normative References . . . . .	<a href="#">16</a>

Black

Expires April 23, 2018

[Page 2]

<a href="#">10.2.</a>	Informative References . . . . .	<a href="#">17</a>
	Author's Address . . . . .	<a href="#">21</a>

## **[1.](#) Introduction**

This memo updates [RFC 3168](#) [[RFC3168](#)] which specifies Explicit Congestion Notification (ECN) as an alternative to packet drops for indicating network congestion to endpoints. It relaxes restrictions in [RFC 3168](#) that hinder experimentation towards benefits beyond just removal of loss. This memo summarizes the proposed areas of experimentation and updates [RFC 3168](#) to enable experimentation in these areas. An Experimental RFC in the IETF document stream [[RFC4844](#)] is required to take advantage of any of these enabling updates. Putting all of these updates into a single document enables experimentation to proceed without requiring a standards process exception for each Experimental RFC that needs changes to [RFC 3168](#), a Proposed Standard RFC.

There is no need for this memo to update [RFC 3168](#) to simplify standardization of protocols and mechanisms that are documented in Standards Track RFCs, as any Standards Track RFC can update [RFC 3168](#) directly without either relying on updates in this memo or using a standards process exception.

In addition, this memo makes related updates to the ECN specification for RTP [[RFC6679](#)] and for three DCCP profiles ([[RFC4341](#)], [[RFC4342](#)] and [[RFC5622](#)]) for the same reason. Each experiment is still required to be documented in one or more separate RFCs, but use of Experimental RFCs for this purpose does not require a process exception to modify any of these Proposed Standard RFCs when the modification falls within the bounds established by this memo ([RFC 5622](#) is an Experimental RFC; it is modified by this memo for consistency with modifications to the other two DCCP RFCs).

Some of the anticipated experimentation includes use of the ECT(1) codepoint that was dedicated to the ECN nonce experiment in [RFC 3540](#) [[RFC3540](#)]. This memo records the conclusion of the ECN nonce experiment and provides the explanation for reclassification of [RFC 3540](#) as Historic in order to enable new experimental use of the ECT(1) codepoint.

### **[1.1.](#) ECN Terminology**

ECT: ECN-Capable Transport. One of the two codepoints ECT(0) or ECT(1) in the ECN field [[RFC3168](#)] of the IP header (v4 or v6). An ECN-capable sender sets one of these to indicate that both transport end-points support ECN.



Not-ECT: The ECN codepoint set by senders that indicates that the transport is not ECN-capable.

CE: Congestion Experienced. The ECN codepoint that an intermediate node sets to indicate congestion. A node sets an increasing proportion of ECT packets to CE as the level of congestion increases.

## **1.2. Requirements Language**

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

## **2. ECN Experimentation: Overview**

Three areas of ECN experimentation are covered by this memo; the cited Internet-Drafts should be consulted for the detailed goals and rationale of each proposed experiment:

**Congestion Response Differences:** An ECN congestion indication communicates a higher likelihood that a shorter queue exists at the network bottleneck node by comparison to a longer queue that is more likely when a packet drop occurs that indicates congestion [[I-D.ietf-tcpm-alternativebackoff-ecn](#)]. This difference suggests that for congestion indicated by ECN, a different sender congestion response (e.g., sender backs off by a smaller amount) may be appropriate by comparison to the sender response to congestion indicated by loss. Two examples of proposed sender congestion response changes are described in [[I-D.ietf-tcpm-alternativebackoff-ecn](#)] and [[I-D.ietf-tsvwg-ecn-l4s-id](#)] - the proposal in the latter draft couples the sender congestion response change to Congestion Marking Differences changes (see next paragraph). This is at variance with [RFC 3168](#)'s requirement that a sender's congestion control response to ECN congestion indications be the same as to drops. IETF approval, e.g., via an Experimental RFC in the IETF document stream, is required for any sender congestion response used in this area of experimentation. See [Section 4.1](#) for further discussion.

**Congestion Marking Differences:** Congestion marking at network nodes can be configured to maintain very shallow queues in conjunction with a different sender response to congestion indications (CE marks), e.g., as proposed in [[I-D.ietf-tsvwg-ecn-l4s-id](#)]. The traffic involved needs to be identified by the senders to the network nodes in order to avoid damage to other network traffic whose senders do not expect the more frequent congestion marking

Black

Expires April 23, 2018

[Page 4]

used to maintain very shallow queues. Use of different ECN codepoints, specifically ECT(0) and ECT(1), is a promising means of traffic identification for this purpose, but that technique is at variance with [RFC 3168](#)'s requirement that ECT(0)-marked traffic and ECT(1)-marked traffic not receive different treatment in the network. IETF approval, e.g., via an Experimental RFC in the IETF document stream, is required for any sender congestion response used in this area of experimentation. See [Section 4.2](#) for further discussion.

TCP Control Packets and Retransmissions: [RFC 3168](#) limits the use of ECN with TCP to data packets, excluding retransmissions. With the successful deployment of ECN in large portions of the Internet, there is interest in extending the benefits of ECN to TCP control packets (e.g., SYNs) and retransmitted packets, e.g., as proposed in [[I-D.bagnulo-tcpm-generalized-ecn](#)]. This is at variance with [RFC 3168](#)'s prohibition of use of ECN for TCP control packets and retransmitted packets. See [Section 4.3](#) for further discussion.

The scope of this memo is limited to these three areas of experimentation. This memo expresses no view on the likely outcomes of the proposed experiments and does not specify the experiments in detail. Additional experiments in these areas are possible, e.g., on use of ECN to support deployment of a protocol similar to DCTCP [[I-D.ietf-tcpm-dctcp](#)] beyond DCTCP's current applicability that is limited to data center environments. The purpose of this memo is to remove constraints in standards track RFCs that stand in the way of these areas of experimentation.

### **[2.1.](#) Effective Congestion Control is Required**

Congestion control remains an important aspect of the Internet architecture [[RFC2914](#)]. Any Experimental RFC in the IETF document stream that takes advantage of this memo's updates to any RFC is required to discuss the congestion control implications of the experiment(s) in order to provide assurance that deployment of the experiment(s) does not pose a congestion-based threat to the operation of the Internet.

### **[2.2.](#) Considerations for Other Protocols**

ECN is widely deployed in the Internet and is being designed into additional protocols such as TRILL [[I-D.ietf-trill-ecn-support](#)]. While the responsibility for coexistence with other protocols and transition from current ECN functionality falls primary upon the designers of experimental changes to ECN, this subsection provides some general guidelines for designers and users of other protocols





that minimize the likelihood of interaction with the areas of ECN experimentation enabled by this memo.

1. [RFC 3168](#)'s forwarding behavior remains the preferred approach for routers that are not involved in ECN experiments, in particular continuing to treat the ECT(0) and ECT(1) codepoints as equivalent, as specified in [Section 4.2](#) below.
2. The ECN CE codepoint SHOULD NOT be assumed to indicate that the packet would have been dropped if ECN were not in use, as that is not the case for either Congestion Response Differences experiments (see [Section 4.1](#) below) or Congestion Marking Differences experiments (see [Section 4.2](#) below). This is already the case when the ECN field is used for Pre-Congestion Notification (PCN) [[RFC6660](#)].
3. Traffic marked with ECT(1) MUST NOT be originated, as specified in [Section 4.2](#) below.
4. ECN may now be used on packets where it has not been used previously, specifically TCP control packets and retransmissions, see [Section 4.3](#) below, and in particular its new requirements for middlebox behavior. In general, any system or protocol that inspects or monitors network traffic SHOULD be prepared to encounter ECN usage on packets and traffic that currently do not use ECN.
5. Requirements for handling of the ECN field by tunnel encapsulation and decapsulation are specified in [[RFC6040](#)]. Additional related guidance can be found in [[I-D.ietf-tsvwg-ecn-encap-guidelines](#)] and [[I-D.ietf-tsvwg-rfc6040update-shim](#)].

### **[2.3. Operational and Management Considerations](#)**

Changes in network traffic behavior that result from ECN experimentation are likely to impact network operations and management. Designers of ECN experiments are expected to anticipate possible impacts and consider how they may be dealt with. Specific topics to consider include possible network management changes or extensions, monitoring of the experimental deployment, collection of data for evaluation of the experiment and possible interactions with other protocols, particularly protocols that encapsulate network traffic.

For further discussion, see [[RFC5706](#)]; the questions in [Appendix A](#) provide a concise survey of some important aspects to consider.

Black

Expires April 23, 2018

[Page 6]

### 3. ECN Nonce and [RFC 3540](#)

As specified in [RFC 3168](#), ECN uses two ECN Capable Transport (ECT) codepoints to indicate that a packet supports ECN, ECT(0) and ECT(1). The second codepoint, ECT(1), is used to support ECN nonce functionality that discourages receivers from exploiting ECN to improve their throughput at the expense of other network users, as specified in Experimental [RFC 3540](#) [[RFC3540](#)]. This section explains why [RFC 3540](#) is being reclassified as Historic and makes associated updates to [RFC 3168](#).

While the ECN nonce works as specified, and has been deployed in limited environments, widespread usage in the Internet has not materialized. A study of the ECN behaviour of the top one million web servers using 2014 data [[Trammell15](#)] found that after ECN was negotiated, none of the 581,711 IPv4 servers tested were using both ECT codepoints, which would have been a possible sign of ECN nonce usage. Of the 17,028 IPv6 servers tested, 4 set both ECT(0) and ECT(1) on data packets. This might have been evidence of use of the ECN nonce by these 4 servers, but might equally have been due to erroneous re-marking of the ECN field by a middlebox or router.

With the emergence of new experimental functionality that depends on use of the ECT(1) codepoint for other purposes, continuing to reserve that codepoint for the ECN nonce experiment is no longer justified. In addition, other approaches to discouraging receivers from exploiting ECN have emerged, see [Appendix B.1](#) of [[I-D.ietf-tsvwg-ecn-l4s-id](#)]. Therefore, in support of ECN experimentation with the ECT(1) codepoint, this memo:

- o Declares that the ECN nonce experiment [[RFC3540](#)] has concluded, and notes the absence of widespread deployment.
- o Updates [RFC 3168](#) [[RFC3168](#)] to remove discussion of the ECN nonce and use of ECT(1) for that nonce.

The four primary updates to [RFC 3168](#) that remove discussion of the ECN nonce and use of ECT(1) for that nonce are:

1. Remove the paragraph in [Section 5](#) that immediately follows Figure 1; this paragraph discusses the ECN nonce as the motivation for two ECT codepoints.
2. Remove [Section 11.2](#) "A Discussion of the ECN nonce." in its entirety.
3. Remove the last paragraph of [Section 12](#), which states that ECT(1) may be used as part of the implementation of the ECN nonce.



4. Remove the first two paragraphs of [Section 20.2](#), which discuss the ECN nonce and alternatives. No changes are made to the rest of [Section 20.2](#), which discusses alternate uses for the fourth ECN codepoint.

In addition, other less substantive [RFC 3168](#) changes are required to remove all other mentions of the ECN nonce and to remove implications that ECT(1) is intended for use by the ECN nonce; these specific text updates are omitted for brevity.

#### **4. Updates to [RFC 3168](#)**

The following subsections specify updates to [RFC 3168](#) to enable the three areas of experimentation summarized in [Section 2](#).

##### **4.1. Congestion Response Differences**

[RFC 3168](#) specifies that senders respond identically to packet drops and ECN congestion indications. ECN congestion indications are predominately originated by Active Queue Management (AQM) mechanisms in intermediate buffers. AQM mechanisms are usually configured to maintain shorter queue lengths than non-AQM based mechanisms, particularly non-AQM drop-based mechanisms such as tail-drop, as AQM mechanisms indicate congestion before the queue overflows. While the occurrence of loss does not easily enable the receiver to determine if AQM is used, the receipt of an ECN Congestion Experienced (CE) mark conveys a strong likelihood that AQM was used to manage the bottleneck queue. Hence an ECN congestion indication communicates a higher likelihood that a shorter queue exists at the network bottleneck node by comparison to a packet drop that indicates congestion [[I-D.ietf-tcpm-alternativebackoff-ecn](#)]. This difference suggests that for congestion indicated by ECN, a different sender congestion response (e.g., sender backs off by a smaller amount) may be appropriate by comparison to the sender response to congestion indicated by loss. However, [section 5 of RFC 3168](#) specifies that:

Upon the receipt by an ECN-Capable transport of a single CE packet, the congestion control algorithms followed at the end-systems MUST be essentially the same as the congestion control response to a \*single\* dropped packet.

This memo updates this [RFC 3168](#) text to allow the congestion control response (including the TCP Sender's congestion control response) to a CE-marked packet to differ from the response to a dropped packet, provided that the changes from [RFC 3168](#) are documented in an Experimental RFC in the IETF document stream. The specific change to [RFC 3168](#) is to insert the words "unless otherwise specified by an



Experimental RFC in the IETF document stream" at the end of the sentence quoted above.

[RFC 4774](#) [[RFC4774](#)] quotes the above text from [RFC 3168](#) as background, but does not impose requirements based on that text. Therefore no update to [RFC 4774](#) is required to enable this area of experimentation.

[Section 6.1.2 of RFC 3168](#) specifies that:

If the sender receives an ECN-Echo (ECE) ACK packet (that is, an ACK packet with the ECN-Echo flag set in the TCP header), then the sender knows that congestion was encountered in the network on the path from the sender to the receiver. The indication of congestion should be treated just as a congestion loss in non-ECN-Capable TCP. That is, the TCP source halves the congestion window "cwnd" and reduces the slow start threshold "ssthresh".

This memo also updates this [RFC 3168](#) text to allow the congestion control response (including the TCP Sender's congestion control response) to a CE-marked packet to differ from the response to a dropped packet, provided that the changes from [RFC 3168](#) are documented in an Experimental RFC in the IETF document stream. The specific change to [RFC 3168](#) is to insert the words "Unless otherwise specified by an Experimental RFC in the IETF document stream" at the beginning of the second sentence quoted above.

#### **[4.2. Congestion Marking Differences](#)**

Taken to its limit, an AQM algorithm that uses ECN congestion indications can be configured to maintain very shallow queues, thereby reducing network latency by comparison to maintaining a larger queue. Significantly more aggressive sender responses to ECN are needed to make effective use of such very shallow queues; Datacenter TCP (DCTCP) [[I-D.ietf-tcpm-dctcp](#)] provides an example. In this case, separate network node treatments are essential, both to prevent the aggressive low latency traffic from starving conventional traffic (if present) and to prevent any conventional traffic disruption to any lower latency service that uses the very shallow queues. Use of different ECN codepoints is a promising means of identifying these two classes of traffic to network nodes, and hence this area of experimentation is based on the use of the ECT(1) codepoint to request ECN congestion marking behavior in the network that differs from ECT(0) counterbalanced by use of a different IETF-approved congestion response to CE marks at the sender, e.g., as proposed in [[I-D.ietf-tsvwg-ecn-l4s-id](#)].

[Section 5 of RFC 3168](#) specifies that:





Routers treat the ECT(0) and ECT(1) codepoints as equivalent.

This memo updates [RFC 3168](#) to allow routers to treat the ECT(0) and ECT(1) codepoints differently, provided that the changes from [RFC 3168](#) are documented in an Experimental RFC in the IETF document stream. The specific change to [RFC 3168](#) is to insert the words "unless otherwise specified by an Experimental RFC in the IETF document stream" at the end of the above sentence.

When an AQM is configured to use ECN congestion indications to maintain a very shallow queue, congestion indications are marked on packets that would not have been dropped if ECN was not in use. [Section 5 of RFC 3168](#) specifies that:

For a router, the CE codepoint of an ECN-Capable packet SHOULD only be set if the router would otherwise have dropped the packet as an indication of congestion to the end nodes. When the router's buffer is not yet full and the router is prepared to drop a packet to inform end nodes of incipient congestion, the router should first check to see if the ECT codepoint is set in that packet's IP header. If so, then instead of dropping the packet, the router MAY instead set the CE codepoint in the IP header.

This memo updates [RFC 3168](#) to allow congestion indications that are not equivalent to drops, provided that the changes from [RFC 3168](#) are documented in an Experimental RFC in the IETF document stream. The specific change is to change "For a router," to "Unless otherwise specified by an Experimental RFC in the IETF document stream" at the beginning of the first sentence of the above paragraph.

A larger update to [RFC 3168](#) is necessary to enable sender usage of ECT(1) to request network congestion marking behavior that maintains very shallow queues at network nodes. When using loss as a congestion signal, the number of signals provided should be reduced to a minimum and hence only presence or absence of congestion is communicated. In contrast, ECN can provide a richer signal, e.g., to indicate the current level of congestion, without the disadvantage of a larger number of packet losses. A proposed experiment in this area, Low Latency Low Loss Scalable throughput (L4S) [[I-D.ietf-tsvwg-ecn-l4s-id](#)] significantly increases the CE marking probability for ECT(1)-marked traffic in a fashion that would interact badly with existing sender congestion response functionality because that functionality assumes that the network marks ECT packets as frequently as it would drop Not-ECT packets. If network traffic that uses such a conventional sender congestion response were to encounter L4S's increased marking probability (and hence rate) at a network bottleneck queue, the resulting traffic throughput is likely

Black

Expires April 23, 2018

[Page 10]

to be much less than intended for the level of congestion at the bottleneck queue.

This memo updates [RFC 3168](#) to remove that interaction for ECT(1). The specific update to [Section 5 of RFC 3168](#) is to replace the following two paragraphs:

Senders are free to use either the ECT(0) or the ECT(1) codepoint to indicate ECT, on a packet-by-packet basis.

The use of both the two codepoints for ECT, ECT(0) and ECT(1), is motivated primarily by the desire to allow mechanisms for the data sender to verify that network elements are not erasing the CE codepoint, and that data receivers are properly reporting to the sender the receipt of packets with the CE codepoint set, as required by the transport protocol. Guidelines for the senders and receivers to differentiate between the ECT(0) and ECT(1) codepoints will be addressed in separate documents, for each transport protocol. In particular, this document does not address mechanisms for TCP end-nodes to differentiate between the ECT(0) and ECT(1) codepoints. Protocols and senders that only require a single ECT codepoint SHOULD use ECT(0).

with this paragraph:

Protocols and senders MUST use the ECT(0) codepoint to indicate ECT unless otherwise specified by an Experimental RFC in the IETF document stream. Protocols and senders MUST NOT use the ECT(1) codepoint to indicate ECT unless otherwise specified by an Experimental RFC in the IETF document stream. Guidelines for senders and receivers to differentiate between the ECT(0) and ECT(1) codepoints will be addressed in separate documents, for each transport protocol. In particular, this document does not address mechanisms for TCP end-nodes to differentiate between the ECT(0) and ECT(1) codepoints.

Congestion Marking Differences experiments SHOULD modify the network behavior for ECT(1)-marked traffic rather than ECT(0)-marked traffic if network behavior for only one ECT codepoint is modified. Congestion Marking Differences experiments MUST NOT modify the network behavior for ECT(0)-marked traffic in a fashion that requires changes to sender congestion response to obtain desired network behavior. If a Congestion Marking Differences experiment modifies the network behavior for ECT(1)-marked traffic, e.g., CE-marking behavior, in a fashion that requires changes to sender congestion response to obtain desired network behavior, then the Experimental RFC in the IETF document stream for that experiment MUST specify:



- o The sender congestion response to CE marking in the network, and
- o Router behavior changes, or the absence thereof, in forwarding CE-marked packets that are part of the experiment.

In addition, this memo updates [RFC 3168](#) to remove discussion of the ECN nonce, as noted in [Section 3](#) above.

#### **4.3. TCP Control Packets and Retransmissions**

With the successful use of ECN for traffic in large portions of the Internet, there is interest in extending the benefits of ECN to TCP control packets (e.g., SYNs) and retransmitted packets, e.g., as proposed by ECN++ [[I-D.bagnulo-tcpm-generalized-ecn](#)].

[RFC 3168](#) prohibits use of ECN for TCP control packets and retransmitted packets in a number of places:

- o "To ensure the reliable delivery of the congestion indication of the CE codepoint, an ECT codepoint MUST NOT be set in a packet unless the loss of that packet in the network would be detected by the end nodes and interpreted as an indication of congestion." ([Section 5.2](#))
- o "A host MUST NOT set ECT on SYN or SYN-ACK packets." ([Section 6.1.1](#))
- o "pure acknowledgement packets (e.g., packets that do not contain any accompanying data) MUST be sent with the not-ECT codepoint." ([Section 6.1.4](#))
- o "This document specifies ECN-capable TCP implementations MUST NOT set either ECT codepoint (ECT(0) or ECT(1)) in the IP header for retransmitted data packets, and that the TCP data receiver SHOULD ignore the ECN field on arriving data packets that are outside of the receiver's current window." ([Section 6.1.5](#))
- o "the TCP data sender MUST NOT set either an ECT codepoint or the CWR bit on window probe packets." ([Section 6.1.6](#))

This memo updates [RFC 3168](#) to allow the use of ECT codepoints on SYN and SYN-ACK packets, pure acknowledgement packets, window probe packets and retransmissions of packets that were originally sent with an ECT codepoint, provided that the changes from [RFC 3168](#) are documented in an Experimental RFC in the IETF document stream. The specific change to [RFC 3168](#) is to insert the words "unless otherwise specified by an Experimental RFC in the IETF document stream" at the end of each sentence quoted above.



In addition, beyond requiring TCP senders not to set ECT on TCP control packets and retransmitted packets, [RFC 3168](#) is silent on whether it is appropriate for a network element, e.g. a firewall, to discard such a packet as invalid. For this area of ECN experimentation to be useful, middleboxes ought not to do that, therefore [RFC 3168](#) is updated by adding the following text to the end of [Section 6.1.1.1](#) on Middlebox Issues:

Unless otherwise specified by an Experimental RFC in the IETF document stream, middleboxes SHOULD NOT discard TCP control packets and retransmitted TCP packets solely because the ECN field in the IP header does not contain Not-ECT. An exception to this requirement occurs in responding to an attack that uses ECN codepoints other than Not-ECT. For example, as part of the response, it may be appropriate to drop ECT-marked TCP SYN packets with higher probability than TCP SYN packets marked with not-ECT. Any such exceptional discarding of TCP control packets and retransmitted TCP packets in response to an attack MUST NOT be done routinely in the absence of an attack and SHOULD only be done if it is determined that the use of ECN is contributing to the attack.

## 5. ECN for RTP Updates to [RFC 6679](#)

[RFC 6679](#) [[RFC6679](#)] specifies use of ECN for RTP traffic; it allows use of both the ECT(0) and ECT(1) codepoints, and provides the following guidance on use of these codepoints in [section 7.3.1](#) :

The sender SHOULD mark packets as ECT(0) unless the receiver expresses a preference for ECT(1) or for a random ECT value using the "ect" parameter in the "a=ecn-capable-rtp:" attribute.

The Congestion Marking Differences area of experimentation increases the potential consequences of using ECT(1) instead of ECT(0), and hence the above guidance is updated by adding the following two sentences:

Random ECT values MUST NOT be used, as that may expose RTP to differences in network treatment of traffic marked with ECT(1) and ECT(0) and differences in associated endpoint congestion responses. In addition, ECT(0) MUST be used unless otherwise specified in an Experimental RFC in the IETF document stream.

[Section 7.3.3 of RFC 6679](#) specifies RTP's response to receipt of CE marked packets as being identical to the response to dropped packets:

The reception of RTP packets with ECN-CE marks in the IP header is a notification that congestion is being experienced. The default





reaction on the reception of these ECN-CE-marked packets MUST be to provide the congestion control algorithm with a congestion notification that triggers the algorithm to react as if packet loss had occurred. There should be no difference in congestion response if ECN-CE marks or packet drops are detected.

In support of Congestion Response Differences experimentation, this memo updates this text in a fashion similar to [RFC 3168](#) to allow the RTP congestion control response to a CE-marked packet to differ from the response to a dropped packet, provided that the changes from [RFC 6679](#) are documented in an Experimental RFC in the IETF document stream. The specific change to [RFC 6679](#) is to insert the words "Unless otherwise specified by an Experimental RFC in the IETF document stream" and reformat the last two sentences to be subject to that condition, i.e.:

The reception of RTP packets with ECN-CE marks in the IP header is a notification that congestion is being experienced. Unless otherwise specified by an Experimental RFC in the IETF document stream:

- \* The default reaction on the reception of these ECN-CE-marked packets MUST be to provide the congestion control algorithm with a congestion notification that triggers the algorithm to react as if packet loss had occurred.
- \* There should be no difference in congestion response if ECN-CE marks or packet drops are detected.

The second sentence of the immediately following paragraph in [RFC 6679](#) requires a related update:

Other reactions to ECN-CE may be specified in the future, following IETF Review. Detailed designs of such additional reactions MUST be specified in a Standards Track RFC and be reviewed to ensure they are safe for deployment under any restrictions specified.

The update is to change "Standards Track RFC" to "Standards Track RFC or Experimental RFC in the IETF document stream" for consistency with the first update.

## **6. ECN for DCCP Updates to RFCs 4341, 4342 and 5622**

The specifications of the three DCCP Congestion Control IDs (CCIDs) 2 [[RFC4341](#)], 3 [[RFC4342](#)] and 4 [[RFC5622](#)] contain broadly the same wording as follows:



each DCCP-Data and DCCP-DataAck packet is sent as ECN Capable with either the ECT(0) or the ECT(1) codepoint set.

This memo updates these sentences in each of the three RFCs as follows:

each DCCP-Data and DCCP-DataAck packet is sent as ECN Capable. Unless otherwise specified by an Experimental RFC in the IETF document stream, such DCCP senders MUST set the ECT(0) codepoint.

In support of Congestion Marking Differences experimentation (as noted in [Section 3](#)), this memo also updates all three of these RFCs to remove discussion of the ECN nonce. The specific text updates are omitted for brevity.

## **7. Acknowledgements**

The content of this draft, including the specific portions of [RFC 3168](#) that are updated draws heavily from [\[I-D.khademi-tsvwg-ecn-response\]](#), whose authors are gratefully acknowledged. The authors of the Internet Drafts describing the experiments have motivated the production of this memo - their interest in innovation is welcome and heartily acknowledged. Colin Perkins suggested updating [RFC 6679](#) on RTP and provided guidance on where to make the updates.

The draft has been improved as a result of comments from a number of reviewers, including Ben Campbell, Brian Carpenter, Benoit Claise, Spencer Dawkins, Gorrry Fairhurst, Sue Hares, Ingemar Johansson, Naeem Khademi, Mirja Kuehlewind, Karen Nielsen, Hilarie Orman, Eric Rescorla, Adam Roach and Michael Welzl. Bob Briscoe's thorough review of an early version of this memo resulted in numerous improvements including addition of the updates to the DCCP RFCs.

## **8. IANA Considerations**

To reflect the reclassification of [RFC 3540](#) as Historic, IANA is requested to update the Transmission Control Protocol (TCP) Header Flags registry (<https://www.iana.org/assignments/tcp-header-flags/tcp-header-flags.xhtml#tcp-header-flags-1>) to remove the registration of bit 7 as the NS (Nonce Sum) bit and add an annotation to the registry to state that bit 7 was used by Historic [RFC 3540](#) as the NS (Nonce Sum) bit.



## **9. Security Considerations**

As a process memo that only relaxes restrictions on experimentation, there are no protocol security considerations, as security considerations for any experiments that take advantage of the relaxed restrictions are discussed in the Internet-Drafts that propose the experiments.

However, effective congestion control is crucial to the continued operation of the Internet, and hence this memo places the responsibility for not breaking Internet congestion control on the experiments and the experimenters who propose them. This responsibility includes the requirement to discuss congestion control implications in an IETF document stream Experimental RFC for each experiment, as stated in [Section 2.1](#); review of that discussion by the IETF community and the IESG prior to RFC publication is intended to provide assurance that each experiment does not break Internet congestion control.

See [Appendix C.1](#) of [\[I-D.ietf-tsvwg-ecn-l4s-id\]](#) for discussion of alternatives to the ECN nonce.

## **10. References**

### **10.1. Normative References**

- [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>>.

## **10.2. Informative References**

- [I-D.bagnulo-tcpm-generalized-ecn]  
Bagnulo, M. and B. Briscoe, "ECN++: Adding Explicit Congestion Notification (ECN) to TCP Control Packets", [draft-bagnulo-tcpm-generalized-ecn-04](#) (work in progress), May 2017.
- [I-D.ietf-tcpm-alternativebackoff-ecn]  
Khademi, N., Welzl, M., Armitage, G., and G. Fairhurst, "TCP Alternative Backoff with ECN (ABE)", [draft-ietf-tcpm-alternativebackoff-ecn-01](#) (work in progress), May 2017.
- [I-D.ietf-tcpm-dctcp]  
Bensley, S., Thaler, D., Balasubramanian, P., Eggert, L., and G. Judd, "Datacenter TCP (DCTCP): TCP Congestion Control for Datacenters", [draft-ietf-tcpm-dctcp-10](#) (work in progress), August 2017.
- [I-D.ietf-trill-ecn-support]  
Eastlake, D. and B. Briscoe, "TRILL: ECN (Explicit Congestion Notification) Support", [draft-ietf-trill-ecn-support-03](#) (work in progress), May 2017.





- [I-D.ietf-tsvwg-ecn-encap-guidelines]  
Briscoe, B., Kaippallimalil, J., and P. Thaler,  
"Guidelines for Adding Congestion Notification to  
Protocols that Encapsulate IP", [draft-ietf-tsvwg-ecn-encap-guidelines-09](#) (work in progress), July 2017.
- [I-D.ietf-tsvwg-ecn-l4s-id]  
Schepper, K. and B. Briscoe, "Identifying Modified  
Explicit Congestion Notification (ECN) Semantics for  
Ultra-Low Queuing Delay", [draft-ietf-tsvwg-ecn-l4s-id-00](#)  
(work in progress), May 2017.
- [I-D.ietf-tsvwg-rfc6040update-shim]  
Briscoe, B., "Propagating Explicit Congestion Notification  
Across IP Tunnel Headers Separated by a Shim", [draft-ietf-tsvwg-rfc6040update-shim-04](#) (work in progress), July 2017.
- [I-D.khademi-tsvwg-ecn-response]  
Khademi, N., Welzl, M., Armitage, G., and G. Fairhurst,  
"Updating the Explicit Congestion Notification (ECN)  
Specification to Allow IETF Experimentation", [draft-khademi-tsvwg-ecn-response-01](#) (work in progress), July  
2016.
- [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>>.



[Trammell15]

Trammell, B., Kuehlewind, M., Boppart, D., Learmonth, I., Fairhurst, G., and R. Scheffenegger, "Enabling Internet-Wide Deployment of Explicit Congestion Notification".

In Proc Passive & Active Measurement (PAM'15) Conference (2015)

## **Appendix A. Change History**

[To be removed before RFC publication.]

Changes from [draft-ietf-tsvwg-ecn-experimentation-00](#) to -01:

- o Add mention of DCTCP as another protocol that could benefit from ECN experimentation (near end of [Section 2](#)).

Changes from [draft-ietf-tsvwg-ecn-experimentation-01](#) to -02:

- o Generalize to describe rationale for areas of experimentation, with less focus on individual experiments
- o Add ECN terminology section
- o Change name of "ECT Differences" experimentation area to "Congestion Marking Differences"
- o Add overlooked [RFC 3168](#) modification to [Section 4.1](#)
- o Clarify text for Experimental RFC exception to ECT(1) non-usage requirement
- o Add explanation of exception to "SHOULD NOT drop" requirement in 4.3
- o Rework [RFC 3540](#) status change text to provide rationale for a separate status change document that makes [RFC 3540](#) Historic. Don't obsolete [RFC 3540](#).
- o Significant editorial changes based on reviews by Mirja Kuehlewind, Michael Welzl and Bob Briscoe.

Changes from [draft-ietf-tsvwg-ecn-experimentation-02](#) to -03:

- o Remove change history prior to WG adoption.
- o Update L4S draft reference to reflect TSVWG adoption of draft.



- o Change the "SHOULD" for DCCP sender use of ECT(0) to a "MUST" (overlooked in earlier editing).
- o Other minor edits.

Changes from [draft-ietf-tsvwg-ecn-experimentation-03](#) to -04:

- o Change name of "Generalized ECN" experimentation area to "TCP Control Packets and Retransmissions."
- o Add IANA Considerations text to request removal of the registration of the NS bit in the TCP header.

Changes from [draft-ietf-tsvwg-ecn-experimentation-04](#) to -05:

- o Minor editorial changes from Area Director review

Changes from [draft-ietf-tsvwg-ecn-experimentation-05](#) to -06:

- o Add summary of [RFC 3168](#) changes to remove the ECN nonce, and use lower-case "nonce" instead of "Nonce" to match [RFC 3168](#) usage.
- o Add security considerations sentence to indicate that review of Experimental RFCs prior to publication approval is the means to ensure that congestion control is not broken by experiments.
- o Other minor editorial changes from IETF Last Call

Changes from [draft-ietf-tsvwg-ecn-experimentation-06](#) to -07:

- o Change draft title to make scope clear - this only covers relaxing of restrictions on ECN experimentation.
- o Any Experimental RFC that takes advantage of this memo has to be in the IETF document stream.
- o Added sections [2.2](#) and [2.3](#) on considerations for other protocols and O&M, relocated discussion of congestion control requirement to [section 2.1](#) from [section 4.4](#)
- o Remove text indicating that ECT(1) may be assigned to L4S - the requirement for an Experimental RFC suffices to ensure that coordination with L4S will occur.
- o Improve explanation of attack response exception to not dropping packets "solely because the ECN field in the IP header does not contain Not-ECT" in [Section 4.3](#)



- o Fix L4S draft reference for discussion of ECN Nonce alternatives - it's [Appendix C.1](#), not B.1.
- o Numerous additional editorial changes from IESG Evaluation

Author's Address

David Black  
Dell EMC  
176 South Street  
Hopkinton, MA 01748  
USA

Email: [david.black@dell.com](mailto:david.black@dell.com)



