Transport Area Working Group
Internet-Draft
Obsoletes: 3540 (if approved)
Updates: 3168, 6679 (if approved)
Intended status: Standards Track
Expires: April 3, 2017

D. Black
Dell EMC

September 30, 2016

# Explicit Congestion Notification (ECN) Experimentation
## draft-black-tsvwg-ecn-experimentation-00

Abstract

Multiple protocol experiments have been proposed that involve changes to Explicit Congestion Notification (ECN) as specified in RFC 3168. This memo summarizes the proposed areas of experimentation to provide an overview to the Internet community and updates RFC 3168, a Proposed Standard RFC, to allow the experiments to proceed without requiring a standards process exception for each Experimental RFC to update RFC 3168. This memo also makes related updates to the ECN specification for RTP in RFC 6679 for the same reason. Each experiment is still required to be documented in an Experimental RFC. This memo also records the conclusion of the ECN Nonce experiment in RFC 3540, obsoletes RFC 3540 and reclassifies it as Historic.

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 http://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 3, 2017.

Table of Contents



## 1. Introduction

Multiple protocol experiments have been proposed that involve changes to Explicit Congestion Notification (ECN) as specified in RFC 3168 [RFC3168]. This memo summarizes the proposed areas of experimentation to provide an overview to the Internet community and updates RFC 3168 to allow the experiments to proceed without requiring a standards process exception for each Experimental RFC to update RFC 3168, a Proposed Standard RFC. This memo also makes related updates to the ECN specification for RTP in RFC 6679 [RFC6679] 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 RFC 3168 or RFC 6679 when the modification falls within the bounds established by this memo.

One of these areas of experimentation involves use of the ECT(1) codepoint that was dedicated to the ECN Nonce experiment as described

in RFC 3540 [RFC3540]. This memo records the conclusion of the ECN Nonce experiment, obsoletes RFC 3540 and reclassifies it as Historic.

### 1.1. Requirements Language

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

## 2. Scope of ECN Experiments

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

Alternative Backoff: For congestion indicated by ECN, use a different TCP sender response (e.g., backoff by a smaller amount) by comparison to congestion indicated by loss, e.g., as specified in [I-D.khademi-tcpm-alternativebackoff-ecn]. This is at variance with RFC 3168's requirement that a TCP sender's congestion control response to ECN congestion indications be the same as to drops.

ECT Differences: Use ECT(1) to request ECN congestion marking behavior in the network that differs from ECT(0), e.g., as specified in [I-D.briscoe-tsvwg-ecn-l4s-id]. This 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.

Generalized ECN: Use ECN for TCP control packets (i.e., send control packets such as SYN with ECT marking) and for retransmitted packets, e.g., as specified in [I-D.bagnulo-tsvwg-generalized-ecn]. This is at variance with RFC 3168's prohibition of use of ECN for TCP control packets and retransmitted packets

The scope of this memo is limited to these three areas of experimentation.

## 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), with the second codepoint used to support ECN nonce functionality to discourage receivers from exploiting ECN to improve their throughput at the expense of other network users, as specified in experimental RFC 3540 [RFC3540].

While the ECN Nonce works as specified, and has been deployed in limited environments, widespread usage in the Internet has not materialized, as the potential for this sort of receiver ECN exploitation has not turned out to be a significant concern in practice. 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 is no longer justified.

Therefore, in support of ECN experimentation with the ECT(1) codepoint, this memo:

- Declares that the ECN Nonce experiment [RFC3540] has concluded, and notes the absence of widespread deployment.

- Obsoletes RFC 3540 in order to facilitate experimental use of the ECT(1) codepoint.

- Reclassifies RFC 3540 as Historic to document the ECN Nonce experiment and discourage further implementation of the ECN Nonce.

- Updates RFC 3168 [RFC3168] to remove discussion of the ECN Nonce and use of ECT(1) for that Nonce. The specific text updates are omitted for brevity.

The following guidance on ECT codepoint usage in Section 5 of RFC 3168 is relevant when the ECN Nonce is not implemented:

> Protocols and senders that only require a single ECT codepoint SHOULD use ECT(0).

## 4. Updates to RFC 3168

In support of these areas of experimentation, this memo updates RFC 3168 [RFC3168] to allow changes in the following areas, provided that the changes are documented by an Experimental RFC. It is also possible to change RFC 3168 via a standards track RFC.

### 4.1. Alternative Backoff

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."

In support of Alternative Backoff experimentation, this memo updates RFC 3168 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. The specific change to RFC 3168 is to insert the words "unless otherwise specified by an Experimental RFC" 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.

### 4.2. ECT Differences

Section 5 of RFC 3168 specifies that:

> "Routers treat the ECT(0) and ECT(1) codepoints as equivalent. Senders are free to use either the ECT(0) or the ECT(1) codepoint to indicate ECT, on a packet-by-packet basis."

In support of ECT Differences experimentation, this memo updates RFC 3168 to allow routers to treat the ECT(0) and ECT(1) codepoints differently, and allow requirements to be imposed on sender usage of ECT(0) and ECT(1), provided that the changes from RFC 3168 are documented in an Experimental RFC. The specific change to RFC 3168 is to insert the words "Unless otherwise specified by an Experimental RFC" and combine the two sentences into a single sentence with this result:

> "Unless otherwise specified by an Experimental RFC, routers treat the ECT(0) and ECT(1) codepoints as equivalent, and senders are free to use either the ECT(0) or the ECT(1) codepoint to indicate ECT, on a packet-by-packet basis."

As ECT(0) was the original codepoint used to signal ECN capability, it is preferable for ECT Differences experiments to modify the behavior of ECT(1) rather than ECT(0) if behavior of only one ECT codepoint is modified.

In support of ECT Differences experimentation, this memo also updates RFC 3168 to remove discussion of the ECN Nonce, as noted in Section 3 above.

### 4.3. Generalized ECN

RFC 3168 prohibits use of ECN for TCP control packets and retransmitted packets in a number of places:

- "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)

- "A host MUST NOT set ECT on SYN or SYN-ACK packets." (Section 6.1.1)

- "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)

- "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)

In support of Generalized ECN experimentation, this memo updates RFC 3168 to allow the use of ECT codepoints on SYN and SYN-ACK packets, pure acknowledgement 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. The specific change to RFC 3168 is to insert the words "unless otherwise specified by an Experimental RFC" at the end of each sentence quoted above.

### 4.4. Effective Congestion Control is Required

Congestion control remains an important aspect of the Internet architecture [RFC2914]. Any Experimental RFC that takes advantage of this memo's updates to RFC 3168 or RFC 6679 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.

## 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 ECT 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 sentence:

> Use of ECT(1) and random ECT values is discouraged, as that may expose RTP to differences in network treatment of ECT(1) and ECT(0), e.g., as proposed in [I-D.briscoe-tsvwg-ecn-l4s-id].

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 Alternative Backoff 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. The specific change to RFC 3168 is to insert the words "Unless otherwise specified by an Experimental RFC" 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:
>
> * 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 alternative 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" for consistency with the first update.

## 6. 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 and provided guidance on where to make the updates.

## 7. IANA Considerations

This memo includes no request to IANA.

## 8. Security Considerations

As a process memo that makes no changes to existing protocols, there are no protocol security considerations.

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, as specified in Section 4.4.

## 9. References

### 9.1. 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, <http://www.rfc-editor.org/info/rfc2119>.

[RFC2914] Floyd, S., "Congestion Control Principles", BCP 41, RFC 2914, DOI 10.17487/RFC2914, September 2000, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc3540>.

[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, <http://www.rfc-editor.org/info/rfc6679>.

### 9.2. Informative References

[I-D.bagnulo-tsvwg-generalized-ecn]
Bagnulo, M. and B. Briscoe, "Adding Explicit Congestion Notification (ECN) to TCP control packets", draft-bagnulo-tsvwg-generalized-ecn-01 (work in progress), July 2016.

[I-D.briscoe-tsvwg-ecn-l4s-id]
Schepper, K., Briscoe, B., and I. Tsang, "Identifying Modified Explicit Congestion Notification (ECN) Semantics for Ultra-Low Queuing Delay", draft-briscoe-tsvwg-ecn-l4s-id-01 (work in progress), March 2016.

[I-D.khademi-tcpm-alternativebackoff-ecn]
Khademi, N., Welzl, M., Armitage, G., and G. Fairhurst, "TCP Alternative Backoff with ECN (ABE)", draft-khademi-tcpm-alternativebackoff-ecn-00 (work in progress), May 2016.

[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, <http://www.rfc-editor.org/info/rfc4774>.

Author's Address

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

Email: david.black@dell.com