Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 10, 2015
M. Welzl
University of Oslo
G. Fairhurst
University of Aberdeen
March 09, 2015

The Benefits to Applications of using Explicit Congestion Notification (ECN)
draft-welzl-ecn-benefits-02

Abstract

This document describes the potential benefits to applications when they enable Explicit Congestion Notification (ECN). It outlines the principal gains in terms of increased throughput, reduced delay and other benefits when ECN is used over network paths that include equipment that supports ECN-marking. The focus of this document is on usage of ECN, not its implementation in hosts, routers and other network devices.

1. Introduction

Internet Transports (such as TCP and SCTP) have two ways to detect congestion: the loss of a packet and, if Explicit Congestion Notification (ECN) [RFC3168] is enabled, by reception of a packet with a Congestion Experienced (CE)-marking in the IP header. Both of these are treated by transports as indications of (potential) congestion. ECN may also be enabled by other transports: UDP applications may enable ECN when they are able to correctly process the ECN signals [RFC5405] (e.g. ECN with RTP [RFC6679]).

A network device (router, middlebox, or other device that forward packets through the network) that does not support AQM, typically uses a drop-tail policy to discard excess IP packets when its queue becomes full. The discard of this packet serves as a signal to the end-to-end transport that there may be congestion on the network path being used. This triggers a congestion control reaction to reduce the maximum rate permitted by the sending endpoint.

When an application uses a transport that enables the use of ECN, the transport layer sets the ECT(0) or ECT(1) codepoint in the IP header of packets that it sends. This indicate to network devices that they may mark, rather than drop, packets in periods of congestion. This marking is generally performed by Active Queue Management (AQM) [RFC2309.bis] and may be the result of various AQM algorithms, where the exact combination of AQM/ECN algorithms is generally not known by the transport endpoints. The focus of this document is on usage of ECN, not its implementation in hosts, routers and other network devices.

ECN makes it possible for the network to signal congestion without packet loss. This lets the network deliver some packets to an application that would otherwise have been dropped. This reduction in packet loss is the most obvious benefit of ECN, but it is often relatively modest. However, enabling ECN can also result in a number of beneficial side-effects, some of which may be much more significant than the immediate reduction in packet loss from ECN-marking instead of dropping packets.

The remainder of this document discusses the potential for ECN to positively benefit an application without making specific assumptions about configuration or implementation.

[RFC3168] describes a method in which a router, sets the CE codepoint of an ECN-Capable packet at the time that the network device would otherwise have dropped the packet. While it has often been assumed that network devices mark packets at the same level of congestion at which they would otherwise have dropped them (e.g., when a queue reaches an AQM threshold), separate configuration of the drop and mark thresholds is known to be supported in some network devices and this is recommended in [RFC2309.bis]. Some benefits of ECN that are discussed rely upon routers marking packets at a lower level of congestion before they would otherwise drop packets [KH13].

Some benefits are also only realised when the transport endpoint behaviour is also updated, this is discussed further in Section 5.

2. ECN Deployment

For an application to use ECN requires that the endpoint first enables ECN within the transport. This requires network devices along the path to at least pass IP packets that set ECN codepoints, and do not drop packets because these codepoints are used. This is the recommended behaviour for network devices [RFC2309.bis]. Applications and transports (such as TCP or SCTP) can be designed to fall-back to not using ECN when they discover they are using a path that does not allow use of ECN (e.g., a firewall or other network device configured to drop the ECN codepoint).

XXX NOTE: A future revision could include some words and reference a paper on the current state of network support for transparently passing the ECN codepoints.

For an application at an endpoint to gain benefit from ECN, network devices need to enable ECN marking. Not all network devices along the path need to enable ECN, for the application to benefit. Any network devices that does not set a CE-codepoint can be expected to drop packets under congestion. Applications that experience congestion in such endpoints do not see any benefit from using ECN, but would see benefit if the congestion were to occur within a network device that did support ECN.

ECN can be incrementally deployed in the general Internet. The IETF has provided guidance on configuration and usage in [RFC2309.bis].

ECN may also be deployed within a controlled environment, for example within a data centre or within a well-managed private network. In this case, the use of ECN may be tuned to the specific use-case. An example is Datacenter TCP (DCTCP) [AL10].

Deployment needs also to consider the requirements for processing ECN at tunnel endpoints of network tunnels, and guidance on the treatment of ECN is provided in [RFC6040]. Further guidance on the encapsulation and use of ECN by non-IP network devices is provided in [ID.ECN-Encap].

3. Benefit of using ECN to avoid congestion loss

When packet loss is a result of (mild) congestion, an ECN-enabled router may CE-mark, rather than drop an ECN-enabled packet. An application can benefit from this marking in several ways:

3.1. Improved Throughput

ECN can improve the throughput performance of an application, although this increase in throughput offered by ECN is often not the most significant gain.

When an application uses a light to moderately loaded network path, the number of packets that are dropped due to congestion is small. Using an example from Table 1 of [RFC3649], for a standard TCP sender with a Round Trip Time, RTT, of 0.1 seconds, a packet size of 1500 bytes and an average throughput of 1 Mbps, the average packet drop ratio is 0.02. This translates into an approximate 2% throughput gain if ECN is enabled. In heavy congestion, packet loss may be unavoidable with, or without, ECN [RFC2309.bis].

3.2. Reduced Head-of-Line Blocking

Many transports provide in-order delivery of received data segments to the applications they support. This requires that the transport stalls (or waits) for all data that was sent ahead of a particular segment to be correctly received before it can forward any later data. This is the usual requirement for TCP and SCTP. PR-SCTP [RFC3758], UDP, and DCCP [RFC4340] provide a transport that does not have this requirement.

Delaying data to provide in-order transmission to an application results in latency when segments are dropped as indications of congestion. The congestive loss creates a delay of at least one RTT for a loss event before data can be delivered to an application. We call this Head-of-Line (HOL) blocking.

In contrast, using ECN can remove the resulting delay following a loss that is a result of congestion:

o First, the application receives the data normally - this also avoids dropping data that has already made it across the network

path. It avoids the additional delay of waiting for recovery of the lost segment when using a reliable transport.

- Second, the transport receiver notes that it has received CE-marked packets, and then requests the sender to make an appropriate congestion-response to reduce the maximum transmission rate for future traffic.

3.3. Reduced Probability of RTO Expiry

In some situations, ECN can help reduce the chance of a retransmission timer expiring (e.g., expiry of the TCP or SCTP retransmission timeout, RTO [RFC5681]). When an application sends a burst of segments and then becomes idle (either because the application has no further data to send or the network prevents sending further data - e.g., flow or congestion control at the transport layer), the last segment of the burst may be lost. It is often not possible to recover this last segment (or last few segments) using standard methods such as Fast Recovery [RFC5681], since the receiver generates no feedback because it is unaware that the lost segments were actually sent.

In addition to avoiding HOL blocking, this allows the transport to avoid the consequent loss of state about the network path it is using, which would have arisen had there been a retransmission timeout. Typical impacts of a transport timeout are to reset path estimates such as the RTT, the congestion window, and possibly other transport state that can reduce the performance of the transport until it again adapts to the path.

Avoiding timeouts can hence improve the throughput of the application. This benefits applications that send intermittent bursts of data, and rely upon timer-based recovery of packet loss. It can be especially significant when ECN is used on TCP SYN/ACK packets as specified in [RFC5562] where the RTO interval may be large because in this case TCP cannot base the timeout period on prior RTT measurements from the same connection.

3.4. Applications that do not retransmit lost packets

Some latency-critical applications use transports that do not retransmit lost packets, yet these applications may be able to adjust the sending rate in the presence of congestion. Examples of such applications include UDP-based services that carry Voice over IP (VoIP), interactive video, or real-time data. The performance of many such applications degrades rapidly with increasing packet loss, and many therefore employ loss-hiding mechanisms (e.g., packet forward error correction, or data duplication) to mitigate the effect

of congestion loss on the application. However, such mechanisms add complexity and can themselves consume additional network capacity reducing the capacity for application data and contributing to the path latency when congestion is experienced.

By decoupling congestion control from loss, ECN can allow the transports supporting these applications to reduce their rate before the application experiences loss from congestion, especially when the congestion is mild and the application/transport can react promptly to reception of a CE-marked packet. Because this reduces the negative impact of using loss-hiding mechanisms, ECN can have a direct positive impact on the quality experienced by the users of these applications.

4. Benefit from Early Congestion Detection

An application can further benefit from using ECN, when the network devices are configured such that they mark packets at a lower level of congestion before they would otherwise have dropped packets from queue overflow:

4.1. Avoiding Capacity Overshoot

Internet transports do not know apriori how much capacity exists along a network path. Transports therefore try to measure the capacity available to an application by probing the network path with increasing traffic to the point where they detect the onset of congestion (such as TCP or SCTP Slow Start).

ECN can help capacity probing algorithms from significantly exceeding the bottleneck capacity of a network path. Since a transport that enables ECN can receive congestion signals before there is significant congestion, an early-marking method in network devices can help a transport respond before it induces significant congestion with resultant loss to itself or other applications sharing a common bottleneck. For example, an application/transport can avoid incurring significant congestion during Slow Start, or a bulk application that tries to increase its rate as fast as possible, may quickly detect the presence of congestion, causing it to promptly reduce its rate.

Use of ECN is more effective than transport mechanisms such as Limited Slow-Start [RFC3742] because it provides direct information about the state of the network path. An ECN-enabled application/ transport that probes for capacity can reduce its rate as soon as it discovers CE-marked packets are received, and before the applications increases its rate to the point where it builds a queue in a network device that induces congestion loss. This benefits an application

seeking to increase its rate - but perhaps more significantly, it eliminates the often unwanted loss and queueing delay that otherwise may be inflicted on flows that share a common bottleneck.

4.2. Making Congestion Visible

A characteristic of using ECN is that it exposes the presence of congestion on a network path to the transport and network layers. This information could be used for monitoring the performance of the path, and could be used to directly meter the amount of congestion that has been encountered upstream on a path; metering packet loss is harder. This is used by Congestion Exposure (CoNex) [RFC6789].

A network flow that only experiences CE-marks and no loss implies that the sending endpoint is experiencing only congestion and not other sources of packet loss (e.g., link corruption or loss in middleboxes). The converse is not true - a flow may experience a mixture of ECN-marks and loss when there is only congestion or when there is a combination of packet loss and congestion [RFC2309.bis]. Recording the presence of CE-marked packets can therefore provide information about the performance of the network path.

5. Other forms of ECN-Marking/Reactions

The ECN mechanism defines both how packets are CE-marked and how transports need to react to reception of marked packets. This section describes the benefits when updated methods are used.

Benefit has been noted when packets are CE-marked earlier than they would otherwise be dropped, using an instantaneous queue, and if the receiver provides precise feedback about the number of packet marks encountered, a better sender behavior is possible. This has been shown by Datacenter TCP (DCTCP) [AL10].

Precise feedback about the number of packet marks encountered is supported by RTP over UDP [RFC6679] and proposed for SCTP [ST14] and TCP [KU13]. An underlying assumption of DCTCP is that it is deployed in confined environments such as a datacenter. It is currently unknown whether or how such behaviour could be introduced into the Internet.

6. Conclusion

Network devices should enable ECN and people configuring host stacks should also enable ECN. These are pre-requisites to allow applications to gain the benefits of ECN.

Application developers should where possible use transports that enable the benefits of ECN. Applications that directly use UDP need to provide support to implement the functions required for ECN. Once enabled, an application that uses a transport that supports ECN will experience the benefits of ECN as network deployment starts to enable ECN. The application does not need to be rewritten to gain these benefits.

Table 1 summarizes some of these benefits.

Section	Benefit
2.1	Improved Throughput
2.2	Reduced Head-of-Line
2.3	Reduced Probability of RTO Expiry
2.4	Applications that do not retransmit lost packets
3.1	Avoiding Capacity Overshoot
3.2	Making Congestion Visible

Table 1: Summary of Key Benefits from using ECN

7. Acknowledgements

The authors were part-funded by the European Community under its Seventh Framework Programme through the Reducing Internet Transport Latency (RITE) project (ICT-317700). The views expressed are solely those of the authors.

8. IANA Considerations

XXRFC ED - PLEASE REMOVE THIS SECTION XXX

This memo includes no request to IANA.

9. Security Considerations

This document introduces no new security considerations. Each RFC listed in this document discusses the security considerations of the specification it contains.

10. Revision Information

RFC-Ed please remove this section prior to publication.

Revision 00 was the first WG draft.

Revision 01 includes updates to complete sections and improve readability. Added section 2.

Comments are welcome to the authors or via the IETF AQM or TSVWG mailing lists.

11. References

11.1. Normative References

[RFC2309.bis]
Baker, F. and G. Fairhurst, "IETF Recommendations Regarding Active Queue Management", Internet-draft draft-ietf-aqm-recommendation-06, October 2014.

[RFC3168] Ramakrishnan, K., Floyd, S., and D. Black, "The Addition of Explicit Congestion Notification (ECN) to IP", RFC 3168, September 2001.

11.2. Informative References

[AL10] Alizadeh, M., Greenberg, A., Maltz, D., Padhye, J., Patel, P., Prabhakar, B., Sengupta, S., and M. Sridharan, "Data Center TCP (DCTCP)", SIGCOMM 2010, August 2010.

[ID.ECN-Encap]
Briscoe, B., Kaippallimalil, J., and P. Thaler, "Guidelines for Adding Congestion Notification to Protocols that Encapsulate IP , IETF work-in-progress: draft-ietf-tsvwg-ecn-encap-guidelines", .

[KH13] Khademi, N., Ros, D., and M. Welzl, "The New AQM Kids on the Block: Much Ado About Nothing?", University of Oslo Department of Informatics technical report 434, October 2013.

[KU13] Kuehlewind, M. and R. Scheffenegger, "Problem Statement and Requirements for a More Accurate ECN Feedback", Internet-draft draft-ietf-tcpm-accecn-reqs-04.txt, October 2013.

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

[RFC3649] Floyd, S., "HighSpeed TCP for Large Congestion Windows", RFC 3649, December 2003.

[RFC3742] Floyd, S., "Limited Slow-Start for TCP with Large Congestion Windows", RFC 3742, March 2004.

[RFC3758] Stewart, R., Ramalho, M., Xie, Q., Tuexen, M., and P. Conrad, "Stream Control Transmission Protocol (SCTP) Partial Reliability Extension", RFC 3758, May 2004.

[RFC4340] Kohler, E., Handley, M., and S. Floyd, "Datagram Congestion Control Protocol (DCCP)", RFC 4340, March 2006.

[RFC5405] Eggert, Lars. and Gorry. Fairhurst, "Unicast UDP Usage Guidelines for Application Designers", November 2008.

[RFC5562] Kuzmanovic, A., Mondal, A., Floyd, S., and K. Ramakrishnan, "Adding Explicit Congestion Notification (ECN) Capability to TCP's SYN/ACK Packets", RFC 5562, June 2009.

[RFC5681] "TCP Congestion Control", .

[RFC6040] "Tunnelling of Explicit Congestion Notification", .

[RFC6679] Westerlund, M., Johansson, I., Perkins, C., O'Hanlon, P., and K. Carlberg, "Explicit Congestion Notification (ECN) for RTP over UDP", RFC 6679, August 2012.

[RFC6789] Briscoe, B., Woundy, R., and A. Cooper, "Congestion Exposure (ConEx) Concepts and Use Cases", RFC 6789, December 2012.

[ST14] Stewart, R., Tuexen, M., and X. Dong, "ECN for Stream Control Transmission Protocol (SCTP)", Internet-draft draft-stewart-tsvwg-sctpecn-05.txt, January 2014.

Authors' Addresses

Michael Welzl
University of Oslo
PO Box 1080 Blindern
Oslo N-0316
Norway

Phone: +47 22 85 24 20
Email: michawe@ifi.uio.no

Godred Fairhurst
University of Aberdeen
School of Engineering, Fraser Noble Building
Aberdeen AB24 3UE
UK

Email: gorry@erg.abdn.ac.uk