A Proposal to add Explicit Congestion Notification (ECN) to IPv6 and to TCP

K. K. Ramakrishnan,
Sally Floyd

References: Floyd, S., TCP and Explicit Congestion Notification. ACM Computer Communication Review, V. 24 N. 5, October 1994, p. 10-23. http://www-nrg.ee.lbl.gov/

Braden et al., Recommendations on Queue Management and Congestion Avoidance in the Internet, Internet draft draft-irtf-e2e-queue-mgt-00.txt, March 1997.

What is new?

1. The deployment of active queue management (e.g., RED) in the Internet.

2. An increasing amount of best-effort traffic where the user is sensitive to the delay (due to retransmission) or drop of an individual packet. E.g., telnet, web browsing, best-effort audio and video, etc.

What is needed from IPv6?

Two bits:

1. An ECN-capable bit set by the origin transport protocol.

2. An ECN-bit set by the router (instead of dropping the packet).
(When the buffer has not overflowed,
and the ECN-capable bit is set,
and the router would otherwise drop the packet because of the RED algorithm, based on the average queue size.

There is a single-bit version of this, described in Floyd94, that overloads a single bit.

What does the ECN bit indicate?

Incipient congestion as indicated by the **average** queue size exceeding a threshold, using the RED algorithms [FJ93, RED-ietf-draft].

The ECN bit should **not** be set in response to an unfiltered signal such as the instantaneous queue size.

A **single** packet with the ECN bit set should be treated by the end-nodes as an indication of congestion, just as would a **single** dropped packet.

How do transport protocols respond to the ECN bit?

The congestion control response should be the same as that to a dropped packet.

The details depend on the transport protocol:
Reliable unicast (e.g., TCP).
Reliable multicast (e.g., SRM).
Unreliable unicast.
Unreliable multicast (e.g., vic).

What modifications are needed for TCP?

Negotiation between the endpoints during setup to determine if they are both ECN-capable.

A TCP option with an ECN-Notify bit so that the data receiver can inform the data sender when a packet has been received with the ECN bit set.

The data sender halves its congestion window `cwnd' in response to an ECN-notify. This is done only once per window of data.