The Addition of Explicit Congestion Notification (ECN) to IP

Sally Floyd and K. K. Ramakrishnan

December 11, 2000

Transport Area Working Group, IETF

The ECN Web Page: “http://www.aciri.org/floyd/ecn.html”

Applications that might benefit from ECN:

- Short web transfers:
  - Avoiding the wait for retransmit timers to expire.
  - Avoiding the delay of retransmitting a dropped packet.

- Larger TCP connections:
  - ECN is somewhat more efficient than Fast Retransmit.

- Unreliable, realtime traffic:
  - Avoiding unnecessary packet losses.

- Reliable multicast traffic:
  - Avoiding costs of detecting and recovering dropped packets.

- Low-bandwidth telnet connections:
  - Avoiding unnecessary waits for retransmit timeouts.

Performance results:

- ECN performance has been well-studied:
  - Jamal Hadi Salim and Uvaiz Ahmed, Performance Evaluation of Explicit Congestion Notification (ECN) in IP Networks, RFC 2884.
  - Prasad Bagal, Shivkumar Kalyanaraman, Bob Packer, Comparative Study of RED, ECN and TCP Rate Control, 1999.
  - Chris Chen, Hariharan Krishnan, Steven Leung, Nelson Tang, Implementing Explicit Congestion Notification (ECN) in TCP for IPv6.
  - And more...

Research on non-default ECN semantics:

- Non-default semantics could be signalled with the diffserv field.

  – S. Kalyanaraman, S. Arora, K. Wanglee, G. Guarriello and D. Harrison, A One-bit Feedback Enhanced Differentiated Services Architecture.

  – R.J. Gibbens and F.P. Kelly, Resource Pricing and the Evolution of Congestion Control.

  – Koenraad Lavens, Peter Key and Derek McAuley, An ECN-based End-to-end Congestion-control Framework: Experiments and Evaluation.

  – Steven Low and others, Optimization Flow Control.

  – Teunis J. Ott, ECN Protocols and the TCP Paradigm.

  – And others...

ECN Implementations:

- Routers:
  - Kenjiro Cho's ALTQ, which runs on FreeBSD/NetBSD/OpenBSD.
  - Linux 2.3, 2.4.
  - Nortel's Open IP Environment 2.1.
  - Others are planned...

- TCP Implementations:
  - Linux 2.4.
  - Experimental prototype from UCLA-ECN.
  - Others are planned...

”The Addition of Explicit Congestion Notification (ECN) to IP”

- Draft-ietf-tsvwg-ecn-00.txt is intended to obsolete:
  - ”A Proposal to add Explicit Congestion Notification (ECN) to IP”:
    RFC 2481, Experimental.
  - ”IPsec with ECN”:
    Approved for Informational.
  - ”ECN Interactions with IP Tunnels”:
    Internet-draft draft-ietf-tsvwg-ecn-tunnels-00.txt.
  - ”TCP with ECN: The Treatment of Retransmitted Data Packets”:
    Internet-draft draft-ietf-tsvwg-ecn-00.txt.

- Section 6.1.1.2 on optional SYN procedures will be removed from the next revision and submitted as an informational RFC.

- A Proposal to Incorporate ECN in MPLS, internet-draft draft-mpls-ecn-00.txt, is not included.

TCP with ECN: The Treatment of Retransmitted Data Packets

- The problem: a potential for denial-of-service attacks.
  – An attacker capable of spoofing the IP source address could send a TCP packet with an arbitrary sequence number and both the ECT and CE bits set in the IP header.

- The fix:
  – Don't set the ECT (or CWR) bit on retransmitted TCP packets, and ignore the ECN field on arriving packets outside the receiver's current window.

- TCP window probes:
  Don't set the ECT (or CWR) bit on window probe packets, but do respond to ECN indications on received window probe packets.

Broken Equipment, and Options for SYN Packets:

- For TCP, ECN-Capability is negotiated between the two ends with ECN-setup SYN or SYN/ACK packets (using the ECN-Echo and CWR flags).

- TCP's ECN-setup SYN packets have been cited as a signature for a port-scanning tool.

- Some broken implementations respond to TCP's ECN-setup SYN packet with a RST.

- Other broken implementations do not respond TCP's ECN-setup SYN packets.

- Detailed results are on the TBIT web page at "http://www.aciri.org/tbit/".

The Optional Procedure for SYN Packets:

- Section 6.1.1.2 on ”Robust TCP Initialization with no response to the SYN” is being removed from this internet-draft, and will be submitted as a separate, informational RFC.

- Bug-fixes have been announced for the broken implementations that send a RST in response to an ECN-setup SYN packet, but so far deployment of this bug-fix seems to be minimal.

- Some work still has to be done to identify the bugs for the broken implementations that ignore ECN-setup SYN packets.

Summary:

- We are asking for this internet-draft to be considered for Proposed Standard. We think it is time.