

tewg/mp1s
Internet Draft
Document: [draft-shayman-mp1s-ecn-00.txt](#)
Category: Informational

M. Shayman
R. Jaeger
University
of Maryland

November 15, 2000
Expires: May 2001

Using ECN to Signal Congestion Within an MPLS Domain

Status of this Memo

This document is an Internet Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

1. Abstract

We propose the addition of Explicit Congestion Notification (ECN) together with congestion signaling back to the ingress in order to provide notification to the ingress label switching router (LSR) if congestion is experienced along a label switched path (LSP). This information could be used by the ingress LSR to mitigate congestion by employing dynamic traffic engineering techniques such as shifting flows to alternate paths.

While extending ECN to enable its use for congestion control by the ingress LSR, the proposal retains (with modification) the capability to signal ECN end-to-end as proposed in an earlier IETF draft authored by others. The current proposal incorporates ECN into MPLS in a manner that is coordinated with the use of ECN in IP, is backwards compatible, and provides MPLS with a means of alleviating congestion of flows traversing MPLS tunnels.

2. Conventions used in this document

Shayman-Jaeger Informational - Expires May 2001

1

Using ECN to Signal Congestion within an MPLS Domain

Nov. 2000

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](#).

3. Introduction

The ECN (Explicit Congestion Notification) concept proposed an alternative to packet drops to indicate congestion to endpoints of a flow. Routers capable of active queue management, e.g., RED [[RFC2309](#), [FJ93](#)], can detect congestion before the queues overflow. In response to congestion, routers typically drop packets to indicate congestion to endpoints. Alternatively, they can mark the packets selected by RED by setting a Congestion Experienced (CE) bit in the IP header of packets from ECN-capable transport connections. By separating the queuing policies from those for congestion indications, endpoints may be able to react to congestion indicated in the packet header rather than suffering packet loss for that indication.

If an end-to-end path traverses an MPLS tunnel, the IP header, and hence the CE bit, is not accessible to the LSRs for marking. To permit ECN to be used on such paths, it was proposed in [draft-ietf-mpls-ecn-00](#) [[RFD99](#)], and summarized in [[DR2000](#)], to use one experimental bit in the MPLS shim header for marking. The proposal in that draft considered end-to-end congestion notification; it did not address the distinct issue of how congestion encountered along an LSP can be signaled back to the ingress LSR. If such information were made available, it could permit the ingress to take actions to alleviate the congestion. Such actions might include shifting flows to alternate LSPs or setting up new LSPs. Furthermore, if the ingress determines that some dropping is unavoidable, the dropping could be done at the ingress, thereby conserving network resources.

The purpose of the present draft is to propose a mechanism whereby ECN can be used to provide congestion notifications to ingress LSRs when MPLS tunnels become congested. At the same time, the ability to provide end-to-end congestion notifications is retained. These two goals are accomplished by (1) modifying the proposal made in [draft-ietf-mpls-ecn-00](#) for the use of a single experimental bit in the shim header; (2) proposing a new RSVP TUNNEL CONGESTION message which is sent to the ingress LSR and ignored by transit LSRs.

The use of ECN to signal congestion to the ingress is complementary to the use of flooding of link loading information via the IGP as proposed in OSPF-OMP [[V99a](#)] and MPLS-OMP [[V99b](#)]. IGP flooding primarily supports load balancing on a time-scale of minutes or longer; the time interval between flooding is envisioned to be at least 15 seconds. Such an approach can be viewed as proactive in that it seeks to reduce the likelihood of congestion occurring. However, even with load balancing, events such as link failures in

Shayman-Jaeger Informational - Expires May 2001

2

Using ECN to Signal Congestion within an MPLS Domain

Nov. 2000

adjacent domains may cause abrupt shifts in traffic patterns that lead to the sudden onset of congestion. Use of ECN can potentially facilitate rapid reaction to such events.

4. ECN Overview

[RFC 2481](#) specifies the addition of ECN to the IP protocol. Two bits in the IP header form the ECN field. The first bit is referred to as the ECN-capable Transport (ECT) bit. It is set by the source and indicates whether the transport connection can support ECN. The second bit is the Congestion Experienced (CE) bit. This bit is set by a router as a marking to indicate that congestion has been encountered. When the destination host receives a packet with the CE bit set, it echoes this information back to the source by setting a bit in the TCP header. The sender then reacts by decreasing its congestion window.

5. MPLS Support for End-to-end ECN

The [draft-ietf-mpls-ecn-00](#) [[RFD99](#)] offered a proposal for making end-to-end ECN possible when the route crosses an MPLS domain. The proposal is briefly summarized as follows:

It is assumed that the protocol used for label distribution (LDP or RSVP) is extended to permit the ingress and egress to signal to all LSRs on the LSP whether or not the LSP is itself ECN-capable. Thus, the ECN-capability of the LSP does not have to be recorded in the shim header. Denote an ECN-capable LSP by ECL; in contrast, ECT denotes an ECN-capable end-to-end transport connection. Let CELSP denote that congestion has been experienced at an LSR along the LSP; in contrast, CE refers to the congestion experienced bit in the IP header which indicates whether congestion has been experienced on the end-to-end path.

To make end-to-end ECN possible when packets traverse an MPLS

tunnel, one of the three experimental bits in the shim header is used. This bit is overloaded in the sense that one value of the bit indicates [ECT && not CELSP] while the other value indicates [not ECT || CELSP]. If a packet marked as [not ECT || CELSP] is picked out for ECN marking at a congested LSR, it will be dropped. This is done because such a packet may belong to a flow for which end-to-end ECN is not possible. A consequence of this is that if a packet has been marked as having experienced congestion at an LSR and then is picked out for ECN marking at a second congested LSR, the packet will be dropped by the second LSR since it cannot be determined whether the packet has previously experienced congestion or if ECN is not possible end-to-end.

The egress LSR folds the indication of congestion experienced along the LSP into the IP header of the packet: If the shim header indicates [not ECT or CELSP] and the IP header indicates [ECT and

Shayman-Jaeger Informational - Expires May 2001 3

Using ECN to Signal Congestion within an MPLS Domain Nov. 2000

not CE], this means that in fact ECN is possible end-to-end and congestion was indeed experienced along the LSP. Thus, the CE bit in the IP header is set.

6. MPLS ECN with Ingress LSR Notification

The use of the experimental bit as specified in [draft-ietf-mpls-ecn-00](#) does not facilitate congestion notification to the ingress LSR when packets encounter congestion in an LSP. This is illustrated by the following two cases:

(1) Suppose the LSP is ECN-capable but ECN is not possible end-to-end i.e., [not ECT && ECL]. In this case if a packet experiences congestion at an LSR and is selected by the active queue management algorithm, it is dropped.

(2) Suppose the LSP is ECN-capable and ECN is possible end-to-end i.e., [ECT && ECL]. In this case if a packet experiences congestion at two LSR's and is selected (e.g., by RED) at both, it will be dropped.

In each of these cases, it is not possible for the ingress LSR to obtain notification that the packet has experienced congestion.

We propose modifying the way the single experimental bit is used so that ECN notifications can be echoed to the ingress LSR if congestion is experienced along an LSP. Instead of overloading the experimental bit to take into account end-to-end ECN-capability, we use it solely to indicate whether or not congestion has been

experienced along the LSP. Thus, the value of the experimental bit indicates [CELSP] or [not CELSP]. (In particular, a packet that experiences congestion and is selected by RED at multiple LSRs is not dropped.)

At the penultimate LSR where the shim header is popped, both the experimental bit and the two-bit ECN field in the IP header are examined. If the experimental bit indicates that congestion was experienced along the LSP, the LSR can then notify the ingress LSR. This notification can occur after some configurable number of ECN packets arrive at the penultimate LSR. The mechanism for notification is a new RSVP TUNNEL CONGESTION message that is sent to the ingress LSR and ignored by transit LSRs.

This proposal retains the capability to do ECN end-to-end. When a packet arrives at the penultimate LSR with the experimental bit indicating that congestion was experienced along the LSP--i.e., the shim header indicates [CELSP]--there are three cases to consider:

(1) If the ECN field in the IP header indicates that the transport connection between the source and destination is not ECN-capable, then the packet is dropped.

Shayman-Jaeger	Informational - Expires May 2001	4
Using ECN to Signal Congestion within an MPLS Domain		Nov. 2000

(2) If the transport connection is ECN-capable and the packet has not yet been marked as having experienced congestion (prior to entering the MPLS domain), then it is remarked as having experienced congestion--i.e., the CE bit in the IP header is set.

(3) If the source and destination are ECN-capable and the packet has already been marked as having experienced congestion (prior to entering the MPLS domain), then it is forwarded to the egress LSR without change.

Note that in case (1) the packet is dropped at the penultimate node, while in [draft-ietf-mpls-ecn-00](#), such a packet would be dropped at the first LSR at which it experienced congestion (and was selected for marking by RED).

7. Conclusions / Further Work

In this draft we have proposed modifying the way a single experimental bit is used so that ECN notifications can be echoed, via a new RSVP TUNNEL_CONGESTION message, to the ingress LSR if congestion is experienced along an LSP. This information could be used by the ingress LSR to shift flows to alternate LSPs or to set

up new LSPs. If the ingress is not able to mitigate the congestion in this way, it can resort to dropping packets. In this case, the notification still has the beneficial effect of enabling the drops to be pushed to ingress, thereby conserving resources within the domain. While enabling ECN to be used to provide notification to the ingress LSRs, the proposal retains the capability to signal ECN end-to-end (provided the transport connection between the source and destination is ECN-capable).

The authors of this draft believe that portions of this material should be incorporated into working group drafts within the scope of the MPLS and possibly other working groups.

8. References

[DR2000] B. Davie and Y. Rekhter, MPLS Technology and Applications, Morgan Kaufmann, San Francisco, 2000.

[ECN] The ECN Web Page, URL <http://www.aciri.org/floyd/ecn.html>".

[F94] S. Floyd, "TCP and Explicit Congestion Notification", ACM Computer Communication Review, Vol. 24, October 1994, pp. 10-23. URL "ftp://ftp.ee.lbl.gov/papers/tcp_ecn.4.ps.Z".

[FBR99] S. Floyd, D. Black and K. K. Ramakrishnan, IPsec Interactions with ECN, Internet Draft, [draft-ipsec-ecn-00.txt](#), work in progress, April 1999. URL "http://www.ietf.cnri.reston.va.us/internet-drafts/draft-ipsec-ecn-00.txt".

Shayman-Jaeger Informational - Expires May 2001 5

Using ECN to Signal Congestion within an MPLS Domain Nov. 2000

[FJ93] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, Vol. 1, August 1993, pp. 397-413. URL <http://www-nrg.ee.lbl.gov/nrg-papers.html>

[LeF99] F. Le Faucheur et al., "MPLS Support of Differentiated Services over PPP Links", Internet Draft, [draft-lefaucheur-mpls-diff-ppp-00.txt](#), work in progress, June 1999. URL "http://www.ietf.cnri.reston.va.us/internet-drafts/draft-lefaucheur-mpls-diff-ppp-00.txt"

[RFC2309] B. Braden et al., "Recommendations on Queue Management and Congestion Avoidance in the Internet," [RFC 2309](#), April 1998.

[RFC2481] K. K. Ramakrishnan and S. Floyd, "A Proposal to Add

Explicit Congestion Notification (ECN) to IP," [RFC 2481](#), January 1999. URL "[ftp://ftp.isi.edu/in-notes/rfc2481.txt](http://ftp.isi.edu/in-notes/rfc2481.txt)".

[RFD99] K. K. Ramakrishnan, S. Floyd and B. Davie, "A Proposal to Incorporate ECN in MPLS," Internet Draft, [draft-ietf-mpls-ecn-00.txt](#), work in progress, June 1999. URL <http://www.ietf.org/proceedings/99jul/I-D/draft-ietf-mpls-ecn-00.txt>.

[R99] E. Rosen, et al., "MPLS Label Stack Encoding", Internet Draft, [draft-ietf-mpls-label-encaps-04.txt](#), work in progress, April 1999. URL "<http://www.ietf.cnri.reston.va.us/internet-drafts/draft-ietf-mpls-label-encaps-04.txt>".

[V99a] C. Villamizar, "OSPF Optimized Multipath," Internet Draft, [draft-ietf-ospf-omp-02](#), work in progress, February 1999, URL <http://www.ietf.org/proceedings/99jul/I-D/draft-ietf-ospf-omp-02.txt>.

[V99b] C. Villamizar, "MPLS Optimized Multipath," Internet Draft, [draft-villamizar-mpls-omp-01](#), work in progress, February 1999, URL <http://www.fictitious.org/internet-drafts/mpls-omp/mpls-omp.html>.

9. Security Considerations

Security considerations with respect to MPLS ECN are equivalent to those for normal IP ECN and ECN with IPsec, which are discussed in [[FBR99](#)].

10. Author's Addresses

Mark Shayman
Department of Electrical and Computer Engineering
University of Maryland
College Park, MD 20742
Phone +1 (301) 405-3667
Email: shayman@glue.umd.edu

Shayman-Jaeger Informational - Expires May 2001

6

Using ECN to Signal Congestion within an MPLS Domain

Nov. 2000

Rob Jaeger
Department of Computer Science
University of Maryland
College Park, MD 20742
Phone +1 (301) 237-7623
Email: rfj@cs.umd.edu

Full Copyright Statement

Copyright (C) The Internet Society (2000). All Rights Reserved. This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.