conex Working Group   S. Krishnan
Internet-Draft   Ericsson
Intended status: Standards Track   M. Kuehlewind
Expires: September 15, 2011   IKR University of Stuttgart
                                                          C. Ucendo
                                                         Telefonica
                                                     March 14, 2011

Options for Conex marking in IPv6 packets
draft-krishnan-conex-ipv6-01

Abstract

   Conex is a mechanism by which senders inform the network about the
   congestion encountered by packets earlier in the same flow. This
   document describes the requirements for conex markings in IPv6
   datagrams and describes the various options for performing conex
   markings in IPv6.

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 September 15, 2011.

include Simplified BSD License text as described in Section 4.e of
the Trust Legal Provisions and are provided without warranty as
described in the Simplified BSD License.

Table of Contents

[1](). Introduction

   Conex is a mechanism by which senders inform the network about the
   congestion encountered by packets earlier in the same flow. This
   document describes the requirements for conex markings in IPv6
   datagrams and describes the various options for performing conex
   markings in IPv6.

[2](). Conventions used in this document

   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 [RFC2119].

[3](). Requirements for marking IPv6 packets

   R-1: The marking mechanism needs to be visible to all conex-capable
   nodes on the path.

   R-2: The mechanism needs to be able to traverse nodes that do not
   understand the markings. This is required to ensure that conex can
   be incrementally deployed over the Internet.

   R-3: The presence of the marking mechanism should not significantly
   alter the processing of the packet. This is required to ensure that
   conex marked packets do not face any undue delays or drops due to a
   badly chosen mechanism.

   R-4: The markings should be immutable once set by the sender. At the
   very least, any tampering should be detectable.

[4](). Possible Solutions

[4.1](). Hop-by-hop options

   The base IPv6 standard [RFC2460] defines hop-by-hop options. These
   options are processed by every node on the path. Hence they meet
   R-1. The options have variable semantics based on the 3 MSB of the
   option code. The state of these bits controls the behavior of nodes
   to either ignore unknown options or drop packets containing them. It
   also defines the ICMPv6 error message sending behavior and the
   mutability of the options en-route. This means that it is possible
   for hop-by-hop options to satisfy R-2 and R-4. In most commercial
   router implementations the mere presence of hop-by-hopoptions rResult
   in the packet being punted to the Slow path instead of being accorded

regular forwarding behavior (Fast Path). This means that R-3 is not satisifed.

4.2. Destination options

The base IPv6 standard [RFC2460] defines the destination options.
These options are processed only by the ultimate receiver of the
packet (as specified in the Destination Address field) and not by
nodes on the path. Hence they do not meet R-1. The options have the
same variable semantics based on the 3 MSBs as the hop-by-hop option
which means that they can satisfy R-2 and R-4. As intermediate nodes
currently do not process destination options R-3 is easily satisifed.

4.3. Header bits

The IPv6 header has no free bits. The only bits in the IPv6 header
that are not widely used are the flow label bits [RFC3697]. There
are some initiatives to redefine the use of the flow label for other
purposes (e.g. Load balancing, nonce). It may be possible (but
highly unlikeley) to save a few bits from the flow label for
alternate purposes to end up with a shorter flow label. The use of
IPv6 header bits can satisfy all the requirements for conex markings
but using valuable header bits for experimental purposes (such as
conex) may not be acceptable.

4.4. Extension Headers

The base IPv6 standard [RFC2460] defines extension headers as an
expansion mechanism to carry optional internet layer information.
Extension headers, with the exception of the hop-by-hop options
header, are not usually processed on intermediate nodes. This means
that R-1 cannot be met. Unknown extension headers cause the packet
to be dropped and hence such mechanism is not incrementally
deployable. Hence R-3 is not met either.

5. ConEx Encoding

The decision about where to code the ConEx inform might also
influence the decision on how to code congestion information itself.
Of course, a ConEx capable transport has to inform the network that
it is actually ConEx enabled. Thus, as a minimum, every packet has
to carry the information that the sender is ConEx enabled and, also
whether it is ConEx marked. Moreover, the abstract conex mechanism
[draft-mathis-conex-abstract-mech-00] requires that that a
distinction between loss or ECN marks as congestion signal is needed
in addition to the so-called 'congestion credits'. This implies that
a minimum of 4 bits is needed if bit-wise encoding is used , and a

minimum of 3 bits is needed if codepoints are used. Further ideas on additional ConEx information are currently discussed on the mailing list. Moreover, the ConEx information could be represented in a more sophisticated manner than a binary signal (Yes/No), if additional bits are available for use.

6. Acknowledgements

The authors would like to thank Marcelo Bagnulo, Bob Briscoe, Ingemar Johansson, Joel Halpern and John Leslie for the discussions that led to this document.

7. Security Considerations

This document does not bring up any new security issues.

8. IANA Considerations

This document does not require any IANA action.

9. Normative References

[RFC2119]   Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC2460]   Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[RFC3697]   Rajahalme, J., Conta, A., Carpenter, B., and S. Deering, "IPv6 Flow Label Specification", RFC 3697, March 2004.

Authors' Addresses

   Suresh Krishnan
   Ericsson
   8400 Blvd Decarie
   Town of Mount Royal, Quebec
   Canada

   Email: suresh.krishnan@ericsson.com

   Mirja Kuehlewind
   IKR University of Stuttgart

   Email: mirja.kuehlewind@ikr.uni-stuttgart.de

   Carlos Ralli Ucendo
   Telefonica

   Email: ralli@tid.es