

Network Working Group
Internet-Draft
Updates: [3697](#) (if approved)
Intended status: Experimental
Expires: August 22, 2010

B. Carpenter
Univ. of Auckland
S. Jiang
Huawei Technologies Co., Ltd
February 18, 2010

Update to the IPv6 flow label specification
draft-carpenter-6man-flow-update-00

Abstract

Various uses proposed for the IPv6 flow label are incompatible with its existing specification. This document describes changes to the specification that permit additional use cases as well as allowing continued use of the previous specification.

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

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>.

This Internet-Draft will expire on August 22, 2010.

Copyright Notice

Copyright (c) 2010 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must 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 BSD License.

Table of Contents

| | | |
|----------------------|------------------------------------|-------------------|
| 1. | Introduction | 3 |
| 2. | Normative Notation | 4 |
| 3. | Changes to specification | 4 |
| 4. | Alternative Approach | 5 |
| 5. | Security Considerations | 6 |
| 6. | IANA Considerations | 6 |
| 7. | Acknowledgements | 6 |
| 8. | Change log | 6 |
| 9. | References | 6 |
| 9.1. | Normative References | 6 |
| 9.2. | Informative References | 7 |
| | Authors' Addresses | 8 |

1. Introduction

The flow label field in the IPv6 header is reserved but left experimental by [\[RFC2460\]](#) and is specified by [\[RFC3697\]](#). We quote three rules from that RFC:

1. "The Flow Label value set by the source MUST be delivered unchanged to the destination node(s)."
2. "IPv6 nodes MUST NOT assume any mathematical or other properties of the Flow Label values assigned by source nodes."
3. "Router performance SHOULD NOT be dependent on the distribution of the Flow Label values. Especially, the Flow Label bits alone make poor material for a hash key."

The second two rules essentially forbid a usage in which the bits of the flow label are encoded with a specific semantic meaning, or are assumed to have any particular property such as randomness. However, both before and after these rules were laid down, a considerable number of proposals for use of the flow label have been published that seem incompatible with them. Examples are

[\[I-D.conta-ipv6-flow-label\]](#), [\[I-D.conta-diffserv-ipv6-fl-classifier\]](#), [\[I-D.chakravorty-6lsa\]](#), [\[I-D.banerjee-flowlabel-ipv6-qos\]](#), [\[I-D.metzler-ipv6-flowlabel\]](#), [\[LeeKim\]](#), [\[LinTseng\]](#), and [\[Prakash\]](#).

These authors propose use cases in which some combination of the following options apply:

- o The flow label may be changed by intermediate systems.
- o It doesn't matter if the flow label is changed, because the receiver doesn't use it.
- o Some or all bits of the flow label are coded: they have specific meanings understood by routers and switches along the path.
- o The coding is related to the required quality of service, as well as identifying a flow.
- o The label is used to control forwarding or switching in some way.

These proposals all require either some form of encoding of semantics in the bits of the flow label, or the ability for routers to modify the flow label, or both. Thus they infringe the rules from [RFC 3697](#) quoted above.

Although [\[I-D.roberts-inband-qos-ipv6\]](#) does not explicitly consider the flow label, it requests hop-by-hop functionality in IPv6 packets very similar to what is needed by the above proposals.

We can conclude that a considerable number of researchers and designers are stymied by [RFC 3697](#). On the other hand, proposals such as [\[I-D.martinbeckman-ietf-ipv6-fls-ipv6flows switching\]](#), [\[I-D.martinbeckman-ietf-ipv6-amp-ipv6hcamp\]](#), [\[I-D.blake-ipv6-flow-label-nonce\]](#), and [\[I-D.carpenter-flow-ecmp\]](#) appear to be compatible with [RFC 3697](#). The latter two are based on

the originator of a packet choosing a pseudo-random flow label for each flow. Thus, we can also conclude that there is a useful role for this approach too. The proposal below is intended to resolve this dilemma by allowing both approaches to co-exist.

2. Normative Notation

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. Changes to specification

We note that although [RFC 3697](#) requires the flow label to be delivered unchanged, it is not included in any transport layer pseudo-header checksums nor in IPsec authentication [[RFC4302](#)]. We also note that at the time of writing, the flow label is observed to be set to zero in an overwhelming proportion of IPv6 packets; neither operating systems nor applications currently set it, and routers do not rely on it. Thus there is no reason to expect operational difficulties if a careful change is made to the rules of [RFC 3697](#).

The purpose of the proposed change is that some flow label values should be available for domain-specific use, with locally defined semantics, and that other flow label values should be available for uses essentially compatible with [RFC 3697](#). There should be no impact on specifications other than [RFC 3697](#) and no impact on currently operational software and hardware.

The proposal is as follows:

- o If the most significant bit (MSB) of the flow label is 0, then the remaining 19 bits MUST obey the rules of [[RFC3697](#)]. (Note that this does not change the meaning of an all-zero flow label or the requirement to deliver it unchanged.)
- o If the MSB of the flow label is 1, the remaining 19 bits MAY obey a locally defined set of rules and those bits MAY be changed en route.

The locally defined set of rules will apply within a given Flow Label Domain, analagous to a Differentiated Services Domain [[RFC2474](#)]. A "boundary router" is defined as any router at the boundary between a Flow Label Domain and other parts of the Internet. The following rules define the consequences for compatibility:

- o Sending hosts that are not updated will in practice continue to send zero labels, which MUST be delivered unchanged.

- o Sending hosts wishing to rely on [RFC 3697](#) behaviour MUST choose labels with MSB = 0.
- o Sending hosts wishing to use locally defined behaviour MUST choose labels with MSB = 1 and whatever other rules apply locally.
- o Receiving hosts that are not updated will continue to ignore labels.
- o Receiving hosts wishing to rely on [RFC 3697](#) behaviour MUST verify that MSB = 0.
- o Receiving hosts wishing to use locally defined behaviour MUST verify that MSB = 1.
- o Routers wishing to implement or rely on locally defined behaviour MUST verify that MSB = 1; if MSB = 0 they MUST NOT change the flow label.
- o Considering packets outbound from the Flow Label Domain, if MSB = 0, a boundary router MUST NOT change the flow label. If MSB = 1, it MUST set all 20 bits of the flow label to zero, so that the locally defined behaviour is not exported from the domain.
- o Considering packets inbound to the Flow Label Domain, if MSB = 0, a boundary router MUST NOT change the flow label. If an inbound packet has MSB = 1, it has originated from a source not following the current specification. This is considered to be an extremely unlikely case, and the boundary router MUST set all 20 bits of the flow label to zero, as the choice least likely to cause unwanted behaviour. (Note that this means the rules for inbound and outbound packets at the boundary router are identical.)

With the ability to define local semantics for 19 bits of the flow label, and the above provisions for compatibility, we add a further recommendation. Its intention is to encourage load balancing solutions based on the flow label, or to enable the behaviour defined in [\[I-D.blake-ipv6-flow-label-nonce\]](#).

- o Sending hosts that do not use a locally defined flow label behaviour SHOULD choose flow labels with MSB = 0 followed by a pseudo-random 19 bit number between 1 and 0x7FFFF.

4. Alternative Approach

Note that an alternative approach would be possible, using a specific differentiated services code point (DSCP)[\[RFC2474\]](#) in the Traffic Class octet instead of the MSB of the flow label itself, to flag a locally defined behaviour. In this model, the above rules would be modified by replacing the condition "MSB = 1" by the condition "DSCP = xxxxxx" (for a specific value xxxxxx) and other fairly straightforward changes. A more elaborate version of this was proposed in [\[I-D.martinbeckman-ietf-ipv6-fls-ipv6flowsnitching\]](#). However, there are two issues with this approach. One is that DSCP values are themselves only locally significant, whereas the

specification above makes the MSB a globally significant flag, consistent with the end-to-end nature of the original flow label definition. Secondly, it seems unwise to meld the semantics of differentiated services, which are currently deployed to some extent, with the unknown future semantics of flow label usage.

5. Security Considerations

The flow label is not protected in any way and can be forged by an on-path attacker. On the other hand, a pseudo-random flow label cannot be readily guessed by an off-path attacker. See [RFC 3697](#) for further discussion.

6. IANA Considerations

This document requests no action by IANA.

7. Acknowledgements

The authors are grateful to Qinwen Hu for general discussion about the flow label and for his work in searching the literature. Valuable comments and contributions were made by ..., and others.

This document was produced using the xml2rfc tool [[RFC2629](#)].

8. Change log

[draft-carpenter-6man-flow-update-00](#): original version, 2010-02-18

9. References

9.1. 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.

9.2. Informative References

- [I-D.banerjee-flowlabel-ipv6-qos]
Banerjee, R., "A Modified Specification for use of the IPv6 Flow Label for providing An efficient Quality of Service using hybrid approach", [draft-banerjee-flowlabel-ipv6-qos-03](#) (work in progress), April 2002.
- [I-D.blake-ipv6-flow-label-nonce]
Blake, S., "Use of the IPv6 Flow Label as a Transport-Layer Nonce to Defend Against Off-Path Spoofing Attacks", [draft-blake-ipv6-flow-label-nonce-02](#) (work in progress), October 2009.
- [I-D.carpenter-flow-ecmp]
Carpenter, B., "Using the IPv6 flow label for equal cost multipath routing in tunnels", [draft-carpenter-flow-ecmp-01](#) (work in progress), February 2010.
- [I-D.chakravorty-6lsa]
Chakravorty, S., Bush, J., and J. Bound, "IPv6 Label Switching Architecture", [draft-chakravorty-6lsa-03](#) (work in progress), July 2008.
- [I-D.conta-diffserv-ipv6-fl-classifier]
Conta, A. and J. Rajahalme, "A model for Diffserv use of the IPv6 Flow Label Specification", [draft-conta-diffserv-ipv6-fl-classifier-01](#) (work in progress), November 2001.
- [I-D.conta-ipv6-flow-label]
Conta, A. and B. Carpenter, "A proposal for the IPv6 Flow Label Specification", [draft-conta-ipv6-flow-label-02](#) (work in progress), July 2001.
- [I-D.martinbeckman-ietf-ipv6-amp-ipv6hcamp]
Beckman, M., "IPv6 Header Compression via Addressing Mitigation Protocol (IPv6 AMP)", [draft-martinbeckman-ietf-ipv6-amp-ipv6hcamp-01](#) (work in progress), March 2007.
- [I-D.martinbeckman-ietf-ipv6-fls-ipv6flowswitching]
Beckman, M., "IPv6 Dynamic Flow Label Switching (FLS)", [draft-martinbeckman-ietf-ipv6-fls-ipv6flowswitching-03](#) (work in progress), March 2007.

[I-D.metzler-ipv6-flowlabel]

Metzler, J. and S. Hauth, "An end-to-end usage of the IPv6 flow label", [draft-metzler-ipv6-flowlabel-00](#) (work in progress), November 2000.

[I-D.roberts-inband-qos-ipv6]

Roberts, L. and J. Harford, "In-Band QoS Signaling for IPv6", [draft-roberts-inband-qos-ipv6-00](#) (work in progress), July 2005.

[LeeKim] Lee, I. and S. Kim, "A QoS Improvement Scheme for Real-Time Traffic Using IPv6 Flow Labels", Lecture Notes in Computer Science Vol. 3043, 2004.

[LinTseng]

Lin, C., Tseng, P., and W. Hwang, "End-to-End QoS Provisioning by Flow Label in IPv6", JCIS , 2006.

[Prakash] Prakash, B., "Using the 20 bit flow label field in the IPv6 header to indicate desirable quality of service on the internet", University of Colorado (M.Sc. Thesis), 2004.

[RFC2474] Nichols, K., Blake, S., Baker, F., and D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers", [RFC 2474](#), December 1998.

[RFC2629] Rose, M., "Writing I-Ds and RFCs using XML", [RFC 2629](#), June 1999.

[RFC4302] Kent, S., "IP Authentication Header", [RFC 4302](#), December 2005.

Authors' Addresses

Brian Carpenter
Department of Computer Science
University of Auckland
PB 92019
Auckland, 1142
New Zealand

Email: brian.e.carpenter@gmail.com

Sheng Jiang
Huawei Technologies Co., Ltd
KuiKe Building, No.9 Xinxu Rd.,
Shang-Di Information Industry Base, Hai-Dian District, Beijing
P.R. China

Email: shengjiang@huawei.com