6MAN	B. Carpenter	
Internet-Draft	Univ. of Auckland	
Intended status: Informational	S. Jiang	
Expires: November 8, 2010	Huawei Technologies Co., Ltd	
	May 7, 2010	

TOC

Update to the IPv6 flow label specification
draft-carpenter-6man-flow-update-03

Abstract

Various published proposals for use of the IPv6 flow label are incompatible with its existing specification in RFC 3697. This document proposes changes to the specification that permit additional use cases. The concept of flow label domains is introduced, with the label possibly being rewritten at domain boundaries.

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 November 8, 2010.

Copyright Notice

Table of Contents

1. Introduction

TOC

The flow label field in the IPv6 header is reserved but left experimental by [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.) and is specified by [RFC3697] (Rajahalme, J., Conta, A., Carpenter, B., and S. Deering, “IPv6 Flow Label Specification,” March 2004.). We quote three rules from that RFC:

> **a.** "The Flow Label value set by the source MUST be delivered unchanged to the destination node(s)."
>
> **b.** "IPv6 nodes MUST NOT assume any mathematical or other properties of the Flow Label values assigned by source nodes."
>
> **c.** "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 rule appears to forbid a usage in which the bits of the flow label are encoded with a specific semantic meaning. If the word "alone" is overlooked, the third rule has sometimes been interpreted to forbid the use of the flow label by load balancing mechansims. 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. An analysis is presented in [I-D.hu-flow-label-cases] (Hu, Q. and B. Carpenter, “Survey of proposed

use cases for the IPv6 flow label," April 2010.), and examples are [I-D.conta-ipv6-flow-label] (Conta, A. and B. Carpenter, "A proposal for the IPv6 Flow Label Specification," July 2001.), [I-D.conta-diffserv-ipv6-fl-classifier] (Conta, A. and J. Rajahalme, "Amodel for Diffserv use of the IPv6 Flow Label Specification," November 2001.), [I-D.chakravorty-6lsa] (Chakravorty, S., Bush, J., and J. Bound, "IPv6 Label Switching Architecture," July 2008.), [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," April 2002.), [I-D.metzler-ipv6-flowlabel] (Metzler, J. and S. Hauth, "An end-to-end usage of the IPv6 flow label," November 2000.), [LeeKim] (Lee, I. and S. Kim, "A QoS Improvement Scheme for Real-Time Traffic Using IPv6 Flow Labels," 2004.), [LinTseng] (Lin, C., Tseng, P., and W. Hwang, "End-to-End QoS Provisioning by Flow Label in IPv6," 2006.), and [Prakash] (Prakash, B., "Using the 20 bit flow label field in the IPv6 header to indicate desirable quality of service on the internet," 2004.). These authors propose use cases in which some combination of the following options apply:

*The flow label may be changed by intermediate systems.

*It doesn't matter if the flow label is changed, because the receiver doesn't use it.

*Some or all bits of the flow label are coded: they have specific meanings understood by routers and switches along the path.

*The coding is related to the required quality of service, as well as identifying a flow.

*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 appear to infringe the rules from RFC 3697 quoted above.

Although [I-D.roberts-inband-qos-ipv6] (Roberts, L. and J. Harford, "In-Band QoS Signaling for IPv6," July 2005.) 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-ipv6flowswitching] (Beckman, M., "IPv6 Dynamic Flow Label Switching (FLS)," March 2007.), [I-D.martinbeckman-ietf-ipv6-amp-ipv6hcamp] (Beckman, M., "IPv6 Header Compression via Addressing Mitigation Protocol (IPv6 AMP)," March 2007.), [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," October 2009.), and [I-D.carpenter-flow-ecmp] (Carpenter, B. and S. Amante, "Using the IPv6 flow label for equal cost multipath routing and link aggregation in tunnels," April 2010.) 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.

If our goal is for the flow label to be used in practice, the conflict between these two approaches creates a dilemma. There appear to be two viable approaches:

1. Definitively forbid locally defined use of the flow label. Strengthen RFC 3697 to say that hosts SHOULD set a pseudo-random label value, which would clarify and limit its possible uses. In particular, its use for load balancing and possibly as a nonce would be encouraged.

2. Encourage locally defined use of the flow label. This approach would make the flow label mutable and would exclude any use case depending on end-to-end immutability. It would encourage applications of a pseudo-random flow label, such as load balancing, on a local basis, but it would exclude end-to-end applications such as [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," October 2009.).

This document is in the form of a set of proposed modifications to the standard, expressing approach 2 and written in normative form. It is suggested that if the proposal is generally accepted, a revised version of RFC 3697 should be produced including these changes. Alternatively, a much simpler revision to express approach 1 above could be chosen.

---

2. Normative Notation

TOC

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] (Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," March 1997.).

---

3. Proposed changes to specification

TOC

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] (Kent, S., “IP Authentication Header,” December 2005.). Both RFC 2460 and RFC 3697 define the default flow label to be zero. At the time of writing, this is the observed value 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.
In particular, the facts that the label is not checksummed and not used mean that the current immutability of the label can be changed without any operational consequences.
The purpose of the proposed change is that the flow label should be available for domain-specific use, with locally defined semantics, without preventing a default type of generic usage. The proposed generic usage is to enourage pseudo-random flow labels that can be used to assist load balancing. There should be no impact on specifications other than RFC 3697 and no impact on currently operational software and hardware.
Firstly we define a "Flow Label Domain" by direct analogy with a Differentiated Services Domain [RFC2474] (Nichols, K., Blake, S., Baker, F., and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” December 1998.):

> Flow Label Domain (also FL domain): a contiguous portion of the Internet over which a consistent scheme of flow label mechanisms is administered in a coordinated fashion. A flow label domain can represent different administrative domains or autonomous systems, different trust regions, different network layer technologies, hosts and routers, etc.
>
> Flow Label Boundary (also FL boundary): the edge of an FL domain. A flow label boundary can be further sub-divided into ingress and egress nodes, where the ingress/egress nodes are the downstream/ upstream nodes of a boundary link in a given traffic direction. A flow label boundary is typically found at the ingress to the first-hop flow label router (or network node) that a host's packets traverse, or at the egress of the last-hop flow label router (or network node) that packets traverse before arriving at a host. A flow label boundary may be co-located with a host, subject to local policy.
>
> Flow Label Router (also FL router): a router that sets or interprets the flow label according to the mechanisms used in a given FL domain.

The rules of RFC 3697 are modified as follows:

1. An FL domain implements a local scheme of flow label mechanisms. The RECOMMENDED scheme is that, whether set by the source host according to RFC 3697, or by an FL router according

to the rules below, the label contains a pseudo-random value between 1 and 0xFFFFF. This recommendation constrains the choice of flow label value more than RFC 3697. An FL domain MAY define an alternative scheme.

2. If and only if the flow label in an IPv6 packet has the default value of zero, then an FL router MAY set it to a value between between 1 and 0xFFFFF. This option modifies the rule that the flow label must be delivered unchanged, by allowing a router in an FL domain to set it if the source host did not set it.

3. If this is done, all packets in a given flow MUST be given the same flow label value. A flow is defined in this case as all packets with the same source and destination IPv6 addresses and port numbers and the same transport protocol number, i.e., the same final Next Header value [RFC2460] (Deering, S. and R. Hinden, “Internet Protocol, Version 6 (IPv6) Specification,” December 1998.). This rule constrains the definition of a flow in RFC 3697 for the specific case that a router sets the flow label. It should be noted that an FL router applying this rule will be obliged to inspect the IPv6 header of every packet, including finding the last "next header" field in the packet, at full line speed.

4. Hosts connected to an FL domain MUST be configured either to set a default (zero) flow label in all IPv6 packets, or to apply the locally defined scheme (which, by rule 1, SHOULD be the pseudo-random scheme).

5. When a locally defined scheme other than the pseudo-random scheme is used, packets entering the FL domain from outside might contain an invalid label according to that scheme. Therefore, boundary ingress FL routers MUST treat all packets entering such an FL domain as if they had a default (zero) flow label.

6. When a locally defined scheme other than the pseudo-random scheme is used, packets leaving the FL domain might contain a label that would be misinterpreted elsewhere. Therefore, the boundary egress FL router SHOULD set the label according to the pseudo-random mechanism defined in rule 1. If not, it MUST set the label to the default value of zero.

The following are the consequences of the above rules combined with those in RFC 3697:

*Sending hosts that are not updated will in practice continue to send all-zero labels. If there is no locally defined scheme in

use along the path taken by a packet, the label will be delivered as zero.

*Sending hosts conforming to this specification will by default choose pseudo-random labels between 1 and 0xFFFFF.

*Locally defined behaviour of the flow label will be limited to consistent administratively defined domains.

*Sending hosts wishing to use locally defined behaviour may continue to send all-zero labels, relying on a router in the local flow label domain to set a value according to the rules above. Alternatively, they may set a label according to locally defined rules.

*Routers wishing to implement a locally defined behaviour will set a label according to the rules above, if and only if the incoming flow label is all-zero, according to rule 1 above.

*The flow label is no longer immutable if it crosses a FL domain boundary. This will allow a wide range of uses cases previously forbidden, and will allow the ECMP/LAG usage defined in [I-D.carpenter-flow-ecmp] (Carpenter, B. and S. Amante, "Using the IPv6 flow label for equal cost multipath routing and link aggregation in tunnels," April 2010.). However, it will break the usage proposed in [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," October 2009.).

---

4. Discussion

TOC

Hosts that set a default (zero) flow label and ignore the flow label on receipt will be unaffected by implementations of this specification. In general, it is assumed that hosts will ignore the flow label on receipt; it cannot be safely used as an end-to-end transport or application layer signal of any kind.
Routers that ignore the flow label will be unaffected by implementations of this specification.
Hosts that set a default (zero) flow label and are in an FL domain where routers adopt a locally defined scheme, or the pseudo-random mechanism in Section 3 (Proposed changes to specification), will benefit from whatever flow label handling is used in the local domain. Clearly, the rules b and c quoted from RFC 3697 in Section 1 (Introduction) have no effect within the local domain, where the locally defined rules (whatever they are) replace them.

Hosts and routers that adopt the pseudo-random mechanism will enhance the performance of any load balancing devices that include the flow label in the hash used to select a particular path or server, even when packets leave the local FL domain. Again, rules b and c have no effect. The rules defined in this proposal are intended to allow encourage the adoption of pseudo-random flow labels in the general case, but also allow a wide variety of locally defined schemes. Such schemes do not need any global assignments of bits in the flow label, and should not have noticeable impact on backwards compatibility or on domains not using them.

---

5. Security Considerations

TOC

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

TOC

This document requests no action by IANA.

---

7. Acknowledgements

TOC

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 Shane Amante, Fred Baker, Steve Blake, Remi Despres, Joel Halpern, Chris Morrow, Mark Smith, and other participants in the 6man working group.
This document was produced using the xml2rfc tool [RFC2629] (Rose, M., “Writing I-Ds and RFCs using XML,” June 1999.).

---

8. Change log

TOC

draft-carpenter-6man-flow-update-03: futher simplified according to WG discussion, 2010-05-07
draft-carpenter-6man-flow-update-02: revised and simplified according to WG discussion, 2010-04-13

draft-carpenter-6man-flow-update-01: revised according to mail list discussion, 2010-03-05
draft-carpenter-6man-flow-update-00: original version, 2010-02-18

---

9. References

TOC

9.1. Normative References

TOC

[RFC2119]	Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," BCP 14, RFC 2119, March 1997 (TXT, HTML, XML).
[RFC2460]	Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification," RFC 2460, December 1998 (TXT, HTML, XML).
[RFC3697]	Rajahalme, J., Conta, A., Carpenter, B., and S. Deering, "IPv6 Flow Label Specification," RFC 3697, March 2004 (TXT).

9.2. Informative References

TOC

[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 (TXT, PS, PDF).
[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 (TXT).
[I-D.carpenter-flow-ecmp]	Carpenter, B. and S. Amante, "Using the IPv6 flow label for equal cost multipath routing and link aggregation in tunnels," draft-carpenter-flow-ecmp-02 (work in progress), April 2010 (TXT).
[I-D.chakravorty-6lsa]	Chakravorty, S., Bush, J., and J. Bound, "IPv6 Label Switching Architecture," draft-chakravorty-6lsa-03 (work in progress), July 2008 (TXT).

[I-D.conta-diffserv-ipv6-fl-classifier]	Conta, A. and J. Rajahalme, "Amodel 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.hu-flow-label-cases]	Hu, Q. and B. Carpenter, "Survey of proposed use cases for the IPv6 flow label," draft-hu-flow-label-cases-00 (work in progress), April 2010 (TXT).
[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 (TXT).
[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 (TXT).
[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 (TXT).
[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 (TXT, HTML, XML).
[RFC2629]	

	Rose, M., “Writing I-Ds and RFCs using XML,” RFC 2629, June 1999 (TXT, HTML, XML).
[RFC4302]	Kent, S., “IP Authentication Header,” RFC 4302, December 2005 (TXT).

Appendix A. Alternative Approaches

TOC

Two more complex alternative approaches were considered and rejected. The first was to distinguish locally significant flow labels from those conforming to RFC 3697 by setting or clearing the most significant bit (MSB) of the flow label. This led to quite complicated rules, seems impossible to make fully self-consistent, and was not considered practical.

The second was to use a specific differentiated services code point (DSCP)[RFC2474] (Nichols, K., Blake, S., Baker, F., and D. Black, “Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers,” December 1998.) in the Traffic Class octet instead of the MSB of the flow label itself, to flag a locally defined behaviour. A more elaborate version of this was proposed in [I-D.martinbeckman-ietf-ipv6-fls-ipv6flowswitching] (Beckman, M., “IPv6 Dynamic Flow Label Switching (FLS),” March 2007.). There are two issues with this approach. One is that DSCP values are themselves only locally significant, inconsistent 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, with the unknown future semantics of flow label usage. However, this approach, while not recommended, does not appear to violate any basic principles if applied strictly within a single differentiated services domain that is also a flow label domain.

Authors' Addresses

TOC

	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 Xinxi Rd.,
	Shang-Di Information Industry Base, Hai-Dian District, Beijing

	P.R. China
Email:	shengjiang@huawei.com