Network Working Group   T. Dreibholz
Internet-Draft   Simula@OsloMET
Intended status: Standards Track   September 05, 2018
Expires: March 9, 2019

An IPv4 Flowlabel Option
draft-dreibholz-ipv4-flowlabel-28

Abstract

This draft defines an IPv4 option containing a flowlabel that is compatible to IPv6. It is required for simplified usage of IntServ and interoperability with IPv6.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

1.1. Terminology

This document uses the following terms:

- IntServ (Integrated Services): Reservation of network resources (bandwidth) on a per-flow basis. See [RFC1633], [RFC2205], [RFC2208], [RFC2209], [RFC2210], [RFC2211] and [RFC2212] for details.

- Flow: An IntServ reservation between two endpoints.

- Flow Label: The Flow Label field of the IPv6 header and the IPv4 option header defined in this draft. It is used for marking a packet to use a specific IntServ reservation. See [RFC6437], [RFC6436] for detailed descriptions.

1.2. Abbreviations

- RSVP: ReSource Reservation Protocol

- SCTP: Stream Control Transmission Protocol

o TCP: Transmission Control Protocol

o QoS: Quality of Service

o UDP: User Datagram Protocol

1.3. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. A Flow Label Option for IPv4

2.1. Motivation

This section describes the motivation to add a flow label option to the IPv4 protocol.

2.1.1. The Flow Label Field of IPv6

The Flow Label field (see [RFC6436] and [RFC6437]) of the IPv6 header (see [RFC2460]) is a 20-bit number. All packets from the same source address having the same flow label MUST contain the same destination address. Therefore, the flow label combined with the source address is a network- unique identification for a specific packet flow. The idea behind the flow label is marking specific flows for IntServ. That is, the routers on the path from source to destination keep e.g. reservation states for the flows. The flow label provides easy identification and utilizes efficient lookup, e.g. using a hash function on the 3-tuple (source address, destination address, flow label).

Using the IPv6 flow label, packets can be mapped easily to specific flows, with the following features:

o Transport Layer Protocol Independence: Since the mapping is directly specified in the IP header, all possible layer 4

protocols are supported, even protocols to be specified in a far future.

- o Support for Network Layer Encryption: The mapping is independent of payload encryption (e.g. by IPsec).

- o Support for Fragmentation: If fragmentation of a large IP packet is necessary, all fragments contain the same flow label. Therefore, fragmentation does not cause any flow-marking problem.

- o Flow Sharing: By marking packets with a flow label, it is possible to share a single flow (IntServ reservation) with several

communication associations from host A to host B. For example, a video stream via UDP and a HTTP download via TCP could share a single reservation. For the user, flow sharing has the advantage that if one of its communication associations temporarily requires lower bandwidth than expected, other associations sharing the same flow may use the remaining bandwidth. That is, his possibly expensive reservation is fully utilized. Flow sharing also helps keeping the total number of reservations a router has to handle small, reducing their CPU and memory requirements and therefore cost.

- o Multi-Flow Connections: One communication association can divide up its packets to several flows, simply by marking packets with different flow labels. This technique can be used for layered transmission. That is, a stream (e.g. a video) is divided up into several parts (called layers). For example, the first layer (base layer) of a video contains a low-quality version, the second (1st enhancement layer) the data to generate a higher-quality version, etc.. Now, the first layer can be mapped to a high-quality reservation (guaranteed bandwidth, low loss rate) at higher cost, but the following layers can be mapped to lower-quality reservations (e.g. higher loss rate) or even best effort at lower cost. Research shows that the total transmission cost can be highly reduced using layered transmission (see [Dre2001], [IJMUE2009] for details).

2.1.2. The Limitations of IntServ via IPv4

Using IntServ with IPv4, there are several problems that can only be

solved with high management effort:

o No Transport Layer Protocol Independence: It is necessary to mark the packets within the layer 4 protocol header. For example, the TCP, UDP or SCTP port numbers can be used to mark flows (with limitations, see below). But for new protocols (e.g. experimental, new standards, proprietary), software updates for *all* IntServ routers are necessary to recognize the packet flow!

o No Support for Network Layer Encryption: Since it is necessary to read fields of the layer 4 protocol header, it may not be encrypted. Therefore, e.g. the usage of IPsec is impossible.

o Support for Fragmentation: Only the first fragment of a large packet contains the layer 4 header necessary to map the packet to a flow. Mapping other fragments would require the hops to remember packet identities and try to map fragments to packet identities. Due to the management effort and memory requirements, this is not realistic for high-bandwidth backbone routers;

especially when packet reordering must be considered. Furthermore, load sharing or traffic distribution would be impossible.

o No Flow Sharing: It is usually impossible for two different communication associations to share the same flow, e.g. if TCP flows are recognized using port numbers. This makes it necessary to reserve an IntServ flow for each communication association. This implies an increased number of flow states for routers to keep and maintain. Furthermore, if one association temporarily uses a lower bandwidth, the free bandwidth of its flow cannot easily be borrowed to another association.

o No Multi-Flow Connections: To use layered transmission, e.g. a video via UDP, the transmission of every layer would require own port numbers. In the case of connection-oriented transmission protocols (e.g. TCP, SCTP), every layer would even require its own connection setup and management. Depending on the transport protocol, the number of communication associations and the number of flows, much more work is necessary compared to IPv6 using flow labels.

All in all, using IntServ flows with IPv4 requires much more work compared to IPv6, where simply the flow label can be used. It is therefore useful to add such a field to IPv4, too. An appropriate place to add such a field is an IPv4 option header.

2.2. Definition of the Flow Label Option

IPv4 (see [RFC0791]) already defines an option header for a 16-bit SATNET stream identifier. Since this identifier would be incompatible to the 20-bit IPv6 flow label, reuse of this existing option header is inappropriate. Therefore, a new one is defined as follows.

Flow Label Option

```
 0                   1                   2                   3
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |    Length     |0 0 0 0 0 0 0 0|0 0 0 0 0 0 0 0|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|0 0 0 0 0 0 0 0|0 0 0 0|              Flow Label               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

- Type: 143

- Length: 8 octets

- Flow Label: The 20-bit flow label. All definitions of [RFC6437] and [RFC6436] for the IPv6 flow label are also valid for this field. A value of zero denotes that no flow label is used. In this case, the flow label option is in fact unnecessary.

The Flow Label option SHOULD be copied on fragmentation. It MUST be the first option of the IP header and therefore MUST NOT appear more than once per IPv4 packet. The Router Alert option SHOULD NOT be used to mark the necessity for routers to examine the options. Placing the Flow Label option as first option allows for easy processing in hardware.

3. Translation between IPv6 and IPv4

Since the new IPv4 flow label is fully compatible to the IPv6 flow

label, the field MAY be translated in the other protocol's one during protocol translation. That is, a router can translate an IPv6 packet set from an IPv6-only host to an IPv4-mapped address of an IPv4-only host and the flow label may simply be copied. The same may also be applied in the backwards direction.

Note, that copying the flow label during protocol translation is not mandatory. There may be IntServ reservation reasons for not copying but setting the flow label to zero. But a router MUST NOT set the flow label to another value than the copy or 0, since the source is responsible to ensure that the source address combined with the flow label is network-unique.

4. Security Considerations

Security considerations are similar to the IPv6 flow label, see [RFC6437].

5. IANA Considerations

This document introduces no additional considerations for IANA.

6. Acknowledgments

The author would like to thank Brian E. Carpenter, Wes George, Perry Lorier, Christoph Reichert and Michael Tuexen for their comments.

7. References

7.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/info/rfc791>.

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

DOI 10.17487/RFC2119, March 1997,
<https://www.rfc-editor.org/info/rfc2119>.

[RFC2205] Braden, R., Ed., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, DOI 10.17487/RFC2205, September 1997, <https://www.rfc-editor.org/info/rfc2205>.

[RFC2210] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, DOI 10.17487/RFC2210, September 1997, <https://www.rfc-editor.org/info/rfc2210>.

[RFC2211] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, DOI 10.17487/RFC2211, September 1997, <https://www.rfc-editor.org/info/rfc2211>.

[RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, DOI 10.17487/RFC2212, September 1997, <https://www.rfc-editor.org/info/rfc2212>.

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <https://www.rfc-editor.org/info/rfc2460>.

[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme, "IPv6 Flow Label Specification", RFC 6437, DOI 10.17487/RFC6437, November 2011, <https://www.rfc-editor.org/info/rfc6437>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

7.2. Informative References

[Dre2001] Dreibholz, T., "Management of Layered Variable Bitrate

Multimedia Streams over DiffServ with Apriori Knowledge", Masters Thesis, February 2001, <https://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-29936/Dre2001.pdf>.

[IJMUE2009] Zhu, W., Dreibholz, T., Rathgeb, E., and X. Zhou, "A Scalable QoS Device for Broadband Access to Multimedia Services", SERSC International Journal of Multimedia and Ubiquitous Engineering (IJMUE) Number 2, Volume 4, Pages 157-172, ISSN 1975-0080, May 2009, <http://www.sersc.org/journals/IJMUE/vol4_no2_2009/14.pdf>.

[RFC1633] Braden, R., Clark, D., and S. Shenker, "Integrated Services in the Internet Architecture: an Overview", RFC 1633, DOI 10.17487/RFC1633, June 1994, <https://www.rfc-editor.org/info/rfc1633>.

[RFC2208] Mankin, A., Ed., Baker, F., Braden, B., Bradner, S., O'Dell, M., Romanow, A., Weinrib, A., and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Applicability Statement Some Guidelines on Deployment", RFC 2208, DOI 10.17487/RFC2208, September 1997, <https://www.rfc-editor.org/info/rfc2208>.

[RFC2209] Braden, R. and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules", RFC 2209, DOI 10.17487/RFC2209, September 1997, <https://www.rfc-editor.org/info/rfc2209>.

[RFC6436] Amante, S., Carpenter, B., and S. Jiang, "Rationale for Update to the IPv6 Flow Label Specification", RFC 6436, DOI 10.17487/RFC6436, November 2011, <https://www.rfc-editor.org/info/rfc6436>.

Author's Address

Thomas Dreibholz
Simula Centre for Digital Engineering
Martin Linges vei 17
1364 Fornebu, Akershus
Norway

Phone: +47-6782-8200
Fax: +47-6782-8201
Email: dreibh@simula.no
URI: https://www.uni-due.de/~be0001/