IPv4 Extension Headers and Flow Label
draft-herbert-ipv4-eh-01

Abstract

This specification defines extension headers for IPv4 and a definition of an IPv4 flow label. The goal is to provide a uniform and feasible method of extensibility that is shared between IPv4 and IPv6.

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/1id-abstracts.html

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

Copyright and License Notice

Copyright (c) 2019 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 Simplified BSD License.

Table of Contents

1 Introduction ............................................. 3
  1.1 Motivation ........................................... 3
  1.2 IPv4 extension headers .............................. 4
  1.3 The IPv4 flow label .................................. 5
2 IPv4 extension headers ................................. 5
  2.1 Requirements ......................................... 6
    2.1.1 General requirements ............................ 6
    2.1.2 Fragmentation and reassembly requirements ....... 7
  2.2 Interaction with standard IPv4 mechanisms .......... 7
    2.2.1 IPv4 options and IPv4 extension headers ........ 8
    2.2.2 IPv4 fragmentation and IPv4 extension headers ... 8
    2.2.3 Atomic datagram recommendation ................. 8
  2.3 ICMP errors .......................................... 9
    2.3.1 Parameter Problem codes ........................ 9
    2.3.2 Destination Unreachable codes .................. 10
  2.4 Processing limits .................................... 11
  2.5 No Next Header ....................................... 12
3 The IPv4 flow label ...................................... 12
  3.1 Sender requirements .................................. 12
  3.2 Receiver requirements ............................... 13
4 Deployability ............................................ 13
5 Security Considerations .................................. 14
6 IANA Considerations ..................................... 14
  6.1 Protocol descriptions ............................... 14
  6.2 Parameter Problem codes ............................ 15
  6.2 Destination Unreachable codes ...................... 15
7 References ................................................ 16
  7.1 Normative References ............................... 16
  7.2 Informative References .............................. 16
Author’s Address .......................................... 18
1 Introduction

This specification defines extension headers for IPv4 as well as an IPv4 flow label. The motivation is to provide an extensible mechanism in IPv4 that is unified with IPv6 and thus leverages common protocol and implementation for extensibility between the two versions of the Internet Protocol.

The extension headers defined for IPv6 in [RFC8200], specifically Hop-by-Hop Options, Destination Options, Routing Header, and Fragment Header, are permitted for use with IPv4 (note that Authentication Header and Encapsulating Security Payload are already usable with IPv4). Additionally, No Next Header (protocol number 59) is defined to be usable in IPv4 packets.

The IPv4 flow label is similarly derived from the definition of the IPv6 flow label. There is no flow label defined in the IPv4 header [RFC791], however under specific circumstances the sixteen bit Identification field may safely be used as a flow label.

1.1 Motivation

IPv6 is intended to become the standard protocol of the Internet, however it is clear that there is a large segment of users that will be using IPv4 for the foreseeable future. This is particularly true in many enterprises where a business case for transitioning to IPv6 hasn’t yet emerged [V6STATE].

In lieu of sun-setting IPv4 and expecting all users to move to IPv6 in some time frame that is unlikely to be met, this specification suggests an alternative which is to improve IPv4. However the nature of these improvements is very specific, the idea is to "backport" useful features of IPv6 into IPv4. Essentially, this makes IPv4 look more like IPv6. The rationale for this is two fold:

1) Users benefit from forward looking features being actively defined and developed for IPv6 without requiring them to transition to IPv6.

2) In making IPv4 look more like IPv6, the work required to complete a future transition to IPv6 may be reduced or simplified.

Various proposals that would use IPv6 extensions are currently being discussed in IETF. These include Segment Routing [SRHV6], Compressed Routing Header [CRH], Path MTU Option [MTUOPT], In-situ OAM [IOAM], Service-aware IPv6 Network [SAIN], and Firewall and Service Tickets [FAST]. These proposals leverage the extensibility mechanism of
extension headers defined for IPv6. All of these proposals, in some form, could be of value for use with IPv4. Unfortunately, IPv4 does not have an extensibility mechanism that meets the requirements for supporting them. IP options are quite limited and have long been considered obsolete. There have been proposals for encoding host to network signaling in UDP (e.g. [SPUD], IOAM over encapsulation like Geneve [IOAMGEN]), however these are shown to neither be generic nor robust especially in the case that encapsulated data must be modified in flight.

The proposal contained in this document is to enable IPv4 packets to carry the extension headers in the same manner that IPv6 packets can carry extension headers. In doing so, the various extensions for IPv6 can be used with IPv4 to the benefit of the user. In many cases (such as IOAM and Path MTU option), the extension being defined is protocol agnostic and would be applicable and usable with IPv4 with little or no change. In other cases, such as segment routing, the extension might be IPv6 specific, for example the segment routing header contains a list of IPv6 addresses. With some modification to the extension definition, it is also conceivable that these may work with IPv4. For instance, in the case of segment routing, the extension can be adapted for use with IPv4 by defining a routing header format that contains IPv4 addresses instead of IPv6 addresses.

1.2 IPv4 extension headers

IPv4 options were defined in [RFC0791] as the means of extending the IP protocol. IPv4 options have not been successful. Early router implementations, and even those today, either don’t process IPv4 options or relegate them to a slow path effectively making them unusable for serious applications. IPv4 options are limited to forty bytes length and, unlike TCP options, no IP options have been defined that are critical to communications. The upshot is that IPv4 options have long not been considered an option for deployment [IPNOOP].

IPv6 took a different approach. Extensibility of IPv6 is provided by extension headers. Optional internet-layer information is encoded in separate headers that may be placed between the IPv6 header and the upper-layer header in a packet [RFC8200]. IPv6 extension headers have had mixed success in deployment in that some intermediate devices have trouble processing them [RFC7872], however there are several active proposals in IETF that would make use of them (e.g. [FAST], [MTUOPT], [IOAM], [SRV6EH]).

Using extension headers with IPv4 is logically straightforward. The IPv4 Protocol field is effectively re-designated to be a Next Header field with the same meaning and semantics as the IPv6 Next Header field. In this manner, an IPv4 packet can contain IPv6 extension
headers that are recast as IPv4 extension headers. These include Hop-by-Hop Options, Routing Header, Fragment, Destination Options, Authentication, and Encapsulating Security Payload. In cases where an extension header contains IPv6 specific information, the extension header can be adapted for use with IPv4. For instance, a Routing Header carrying IPv6 addresses to visit could be adapted to carry IPv4 addresses.

A number of ICMP errors may be sent when an error condition is encountered in processing extension headers. The relevant ICMPv6 errors are defined in [RFC4443] and [ICMPLIM]. This specification adapts these ICMPv6 errors for use in IPv4.

1.3 The IPv4 flow label

IPv6 [RFC8200] introduced the concept of a flow label that has proven quite convenient to perform flow classification, such as that needed by Equal-Cost Multipath (ECMP). The base IPv4 header does not have reserved bits that could be allocated as a flow label, however the sixteen bit Identification field can be used as a flow label in atomic datagrams [RFC6864].

2 IPv4 extension headers

IPv4 extension headers are optional internet-layer information encoded in separate headers that may be placed between the IPv4 header and the upper-layer header in a packet. IPv4 extension headers are based on IPv6 extension headers and share the same basic properties and semantics [RFC8200].

Extension headers are numbered from IANA IP Protocol Numbers [IANA-PN], the same values are used for IPv4 and IPv6. When processing a sequence of Next Header values in a packet, the first one that is not an extension header [IANA-EH] indicates that the next item in the packet is the corresponding upper-layer header. A special "No Next Header" value is used if there is no upper-layer header.

As illustrated in these examples, an IPv4 packet MAY carry zero, one, or more extension headers, each identified by Protocol field of the IPv4 header or the Next Header field of a preceding extension header:
2.1 Requirements

2.1.1 General requirements

IPv4 extension headers normatively assume the requirements of IPv6 extension headers as defined in [RFC8200] section 4, with the following modifications:

* References to the IPv6 header are replaced by references to the IPv4 header.

* ICMP errors sent in the course of processing extension headers use ICMPv4 instead of ICMPv6. Applicable ICMPv4 errors for extension header processing are specified in section 2.3.

* The IPv4 header Protocol field assumes the same role and semantics with respect to extension headers as the IPv6 Next Header field.

* The Hop-by-Hop Options header is used to carry optional information that MAY be examined and processed by any node along a packet’s delivery path.

* If a legacy IPv4 destination node, one that does not support IPv4 extension headers, receives a packet with extension headers then the packet will be processed as having an unknown protocol. It is expected that the packet will be discarded and an ICMP
error may be generated.

* Extension headers or options that carry IPv6 specific data or are otherwise specific to IPv6 MUST NOT be used with IPv4 (Segment Routing [SRV6EH] for example). IPv4 variants of these may be defined if achieving the same functionality in IPv4 is desirable.

* References to the Payload Length, for instance in reassembly procedures, are reinterpreted as being the computed IPv4 payload length (i.e. IPv4 Total Length minus the length of the IPv4 header).

2.1.2 Fragmentation and reassembly requirements

The following are modifications to fragmentation and reassembly requirements:

* References to setting the Payload Length field in the IPv6 header are interpreted to be setting the Total Length in the IPv4 header taking into account the IPv4 header length.

* When creating or modifying IPv4 headers in packets, the IPv4 header checksum MUST be set correctly.

* Different fragment packets MAY contain different IPv4 options. In the reassembled packet, the IP options are taken from the first fragment packet (the one with offset of zero).

* Different fragment packets MAY contain different extension headers preceding the fragment header. In the reassembled packet, the extension headers preceding the fragment header are taken from the first fragment packet (the one with offset of zero).

* If the length and offset of a fragment are such that the Total Length of the packet reassembled from that fragment would exceed 65,535 octets, then that fragment must be discarded and an ICMP Parameter Problem, Code 0, message should be sent to the source of the fragment, pointing to the Fragment Offset field of the fragment packet.

2.2 Interaction with standard IPv4 mechanisms

IPv4 extension headers may be used concurrently with IPv4 mechanisms such as IPv4 options and IPv4 fragmentation. This section discusses the interactions.
2.2.1 IPv4 options and IPv4 extension headers

An IPv4 packet MAY contain both IPv4 options and extension headers. IPv4 options are completely independent of IPv4 extension headers. IPv4 options MUST be processed before processing any extension headers per normal requirements of processing the IP header before the IP payload.

2.2.2 IPv4 fragmentation and IPv4 extension headers

An IPv4 packet MAY be fragmented both by using a Fragment extension header as well as by standard IPv4 fragmentation. The Fragment header can only be set at the source, however intermediate devices can fragment packets using standard IPv4 fragmentation. Standard IPv4 fragmentation at a source node MUST be done only after any extension headers are set in a packet or the packet was fragmented using the Fragment header. Specifically, fragmentation using the extension header MUST NOT be done on packet fragments created by standard IPv4 fragmentation. However, a packet fragment that contains a Fragment header MAY itself be fragmented by standard IPv4 fragmentation. There is no correlation between standard IPv4 fragmentation and the IPv4 Fragment header, the identifier space for each are unrelated and reassembly procedures are independent.

At a destination, if a received packet was fragmented by standard IPv4 fragmentation, it MUST be reassembled before processing any IPv4 extension headers. This requirement ensures that standard IPv4 reassembly is done before reassembly for the Fragment header.

If an IPv4 packet containing Hop-by-Hop options is fragmented using standard IPv4 fragmentation, the Hop-by-Hop Options are not set in each of the packet fragments. An intermediate node MAY process the Hop-by-Hop options in the first fragment if the complete Hop-by-Hop extension header is contained within the fragment. If the Fragment header is used with IPv4 then the DF bit (Don’t Fragment) bit SHOULD be set and Path MTU discovery mechanisms SHOULD be used.

2.2.3 Atomic datagram recommendation

It is RECOMMENDED to only use IPv4 extensions in atomic datagrams. Atomic datagrams [RFC6864] are IPv4 packets for which the Don’t Fragment bit set, More Fragment bit is not set, and Fragment Offset is zero. In this case the packet will not be subject to IPv4 fragmentation, the Fragment header can alternatively be used for fragmentation.
2.3 ICMP errors

ICMP errors are defined to be sent in response to errors that occur in processing extension headers. Six ICMPv4 Parameter Problem codes are defined and one Destination Unreachable code is defined. These codes are adaptations of ICMPv6 codes defined in [RFC4443] and [ICMPLIM].

2.3.1 Parameter Problem codes

The format for ICMP Parameter Problem errors related to extension header employs a multi-part ICMPv4 message format as defined in [RFC4884]. The extended structure contains a pointer to the octet beyond the limit.

The format of the ICMPv4 Parameter Problem for extension headers is:

```
| 0  | 1  | 2  | 3  |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Type      |     Code      |          Checksum             |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     unused    |    Length     |             unused            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  Internet Header + leading octets of original datagram    |
|                                                               |
|                                                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Pointer                               |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

IPv4 Fields:

- **Destination Address**
  - Copied from the Source Address field of the invoking packet.

ICMPv4 Fields:

- **Type**
  - 12 (Parameter Problem Message)

- **Code** (pertinent to this specification)
  - 3 - Erroneous header field encountered
  - 4 - Unrecognized Next Header type encountered
  - 5 - Unrecognized IPv4 option encountered
6 - Extension header too big
7 - Extension header chain too long
8 - Too many options in extension header

Length
Length of the padded "original datagram" field, measured in 32-bit words.

Pointer
Identifies the octet offset within the invoking packet where the error was detected.

Codes 3, 4, and 5 are analogues of Parameter Problem codes 0, 1, and 2 defined for IPv6 in [RFC4443]. Operation and semantics of these codes are the same as their counterparts in [RFC4443] with the following differences:

* The multi-part ICMP format of [RFC4884] is used and its fields are set appropriately.

* The pointer to the offending byte in the invoking packet is contained in the 32 bit Pointer field in the extended format.

Codes 6, 7, and 8 are analogues of Parameter Problem codes 4, 5, and 6 defined for IPv6 in [ICMPLIM]. Operation and semantics of these codes are the same as their counterparts in [RFC4443] with the following differences:

* The multi-part ICMP format of [RFC4884] is used and its fields are set appropriately.

* The pointer to the offending byte in the invoking packet is contained in the 32 bit Pointer field in the extended format.

2.3.2 Destination Unreachable codes

This specification defines one IPv4 Destination Unreachable code for aggregate header limits being exceeded as described in [ICMPLIM]. The error for aggregate header limits employs a multi-part ICMPv4 message format as defined in [RFC4884]. The extended structure contains a pointer to the octet beyond the limit.

The format of the ICMPv4 message for an aggregate header limit exceeded is:
IPv4 Fields:

Destination Address
  Copied from the Source Address field of the invoking packet.

ICMPv4 Fields:

  Type
    3 (Destination Unreachable Message)

  Code (pertinent to this specification)
    16 - Headers too long

  Length
    Length of the padded "original datagram" field, measured in 32-bit words.

  Pointer
    Identifies the octet offset within the invoking packet where the error was detected.

Code 16 is an analogue of Destination Unreachable code 8 defined in [ICMP6LIM]. Operation and semantics of the code should be the same as the counterpart in [ICMP6LIM].

2.4 Processing limits

Section 5.3 of [RFC8504] describes the use of limits in processing extension headers in order to protect a node from excessive extension header options. These limits are adapted for use with IPv4 extension headers.
A node MAY impose limits on processing IPv4 extension headers and Destination and Hop-by-Hop options. This includes limits on length or number of consecutive padding options, disallowing unknown options, and limits on the number of options or length of options. If a limit is exceeded in processing a received packet, the packet is discarded and an appropriate ICMP error defined in Section 2.3 SHOULD be sent.

A node MAY allow configuration of different limits for processing IPv4 and IPv6 options. The default limits for IPv4 are assumed to be those defined in [RFC8504].

2.5 No Next Header

The value 59 may be set in the Protocol field of the IPv4 header or in Next Header field of an IPv4 extension header. This indicates that there is nothing following that header. The semantics of setting this value are the same as that described in [RFC8200] with adaptation for use with IPv4. If the Total Length field of the IPv4 header indicates the presence of octets past the end of a header whose Next Header field contains 59, those octets must be ignored and passed on unchanged if the packet is forwarded.

3 The IPv4 flow label

The Identification field of the IPv4 header is re-purposed to be the IPv4 flow label in atomic datagrams. As stated in [RFC6864]:

">> Originating sources MAY set the IPv4 ID field of atomic datagrams to any value."

This specification allows the IPv4 ID to be used as a flow label in atomic datagrams where \((DF==1)\&\&(MF==0)\&\&(frag_offset==0)\).

3.1 Sender requirements

An origin host MAY set the IPv4 Identification field as a flow label in atomic datagram packets. The IPv4 flow label is set following the same procedures for setting the IPv6 flow label as described in [RFC6437] with the following modifications:

* The Identification field MUST only be used as a flow label in atomic datagrams. That is Don’t Fragment (DF) bit MUST be set, More Fragment (MF) bit MUST NOT be set, and Fragment Offset MUST be zero.

* If the IPv4 Identification field is not used as a flow label in atomic fragments, the Identification field MUST be set to zero.
* Only stateless flow labels can be set.

* The value to set, e.g. from a hash computation over packet headers, is truncated to sixteen bits (the size of the Identification field).

* Intermediate nodes MUST NOT set the Identification field in atomic datagrams.

3.2 Receiver requirements

Receivers, including intermediate hosts, MAY process a non-zero Identification field in the IPv4 header of atomic datagrams as being a flow label. The IPv4 flow label for instance can be used as input to ECMP as described in [RFC6438].

If the Identification field is zero or the packet is not an atomic datagram (either the More Fragment bit is set, the Don’t Fragment bit is not set, or Fragment Offset is non-zero) then the Identification field MUST NOT be considered as a flow label.

4 Deployability

If a legacy host device receives an IPv4 packet with IPv4 extension headers, the packet will be treated as having an unknown protocol and should dropped. Intermediate devices might also see packets with a protocol unknown to them and will forward the packet inasmuch as they would forward any packet with an unknown protocol.

In the Internet, it is well known that there are some intermediate nodes that will drop packets with protocols that are unknown to them (firewalls would commonly to this for instance). Therefore, it is unlikely that packets with IPv4 extension headers can be ubiquitously deployed over the Internet. A workaround to this might be to encapsulate extension headers in UDP [EHUDPENCAP].

In a limited domain [LIMDOM], an operator would have control over intermediate nodes and could ensure that at a minimum they properly forward packets with IPv4 extension headers. Routers in a limited domain can be updated to process IPv4 Hop-by-Hop Options or Routing headers to provide the functionality of features like IOAM and Segment Routing in IPv4. Similarly, they could be updated to support the IPv4 flow label to provide flow based ECMP in the same manner that the IPv6 flow label is used for ECMP [RFC6438].
5 Security Considerations

This specification enables use of IPv6 extension headers in IPv4. Related security mechanisms of IPv6 extension headers can be applied for use with IPv4 extension headers.

The IPv4 flow label has similar security properties as the IPv6 flow label.

6 IANA Considerations

6.1 Protocol descriptions

IANA is requested to change the descriptions of IPv6 extension headers and No Next Header protocol numbers to reflect that they are not IPv6 specific.

In the Assigned Internet Protocol Numbers Registry, the modified protocols descriptions are:

<table>
<thead>
<tr>
<th>Decimal</th>
<th>Keyword</th>
<th>Protocol</th>
<th>IPv6 Extension header</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>HOPOPT</td>
<td>Hop-by-Hop Option</td>
<td>[RFC8200][RFCXXXX]</td>
<td>[This document]</td>
</tr>
<tr>
<td>43</td>
<td>Route</td>
<td>Routing Header</td>
<td>[Steve_Deering]</td>
<td>[This document]</td>
</tr>
<tr>
<td>44</td>
<td>Frag</td>
<td>Fragment Header</td>
<td>[Steve_Deering]</td>
<td>[This document]</td>
</tr>
<tr>
<td>59</td>
<td>NoNxt</td>
<td>No Next Header</td>
<td>[RFC8200][RFCXXXX]</td>
<td>[This document]</td>
</tr>
<tr>
<td>60</td>
<td>Opts</td>
<td>Destination Options</td>
<td>[RFC8200][RFCXXXX]</td>
<td>[This document]</td>
</tr>
</tbody>
</table>

IANA is requested to update "Internet Protocol Version 4 (IPv4) Parameters" to include sections for "IPv6 Extension Header Types", "Destination Options and "Hop-by-Hop Options", and "Routing Types". These are based on the similarly named sections in "Internet Protocol Version 6 (IPv6) Parameters" with appropriate modifications for IPv4.
6.2 Parameter Problem codes

IANA is requested to assign the following codes in "ICMP Type Numbers" for type 12 "Parameter Problem":

3 - Erroneous header field encountered
4 - Unrecognized Next Header type encountered
5 - Unrecognized IPv4 option encountered
6 - Extension header too big
7 - Extension header chain too long
8 - Too many options in extension header

6.2 Destination Unreachable codes

IANA is requested to assign the following codes in "ICMP Type Numbers" for type 3 "Destination Unreachable":

16 - Headers too long
7 References

7.1 Normative References


7.2 Informative References


[RFC6437] Amante, S., Carpenter, B., Jiang, S., and J. Rajahalme,


[IOAM] F. Brockners, S. Bhandari, V. Govindan, C. Pignataro, H. Gredler, J. Leddy, S. Youell, T. Mizrahi, D. Mozes, P. Lapukhov, R. Chang, "Encapsulations for In-situ OAM Data" draft-brockners-inband-oam-transport-05


[FAST] Herbert, T., "Firewall and Service Tickets", draft-herbert-fast-03


[IOAMGEN] Brockners, F. et al., "Geneve encapsulation for In-situ OAM Data", draft-brockners-ippm-ioam-geneve-01

[IPNOOP] Rodrigo Fonseca, George Manning Porter, Randy H. Katz, Scott Shenker and Ion Stoica, "IP Options are not an option", <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2005/EECS-2005-24.html>
[ICMPLIM] Herbert, T., "ICMPv6 errors for discarding packets due to processing limits", draft-ietf-6man-icmp-limits-01


Author’s Address

Tom Herbert
Quantonium
Santa Clara, CA
USA

Email: tom@quantonium.net