Intarea Working Group
Internet-Draft
Intended status: Standards Track
Expires: October 27, 2012

C. Bormann
Universitaet Bremen TZI

April 25, 2012

Adaptation Layer Fragmentation Indication
draft-bormann-intarea-alfi-00

Abstract

IPv6 defines a minimum MTU of 1280 bytes. Many link layers are more limited in their choice of packet size. Typically, IP adaptation layers for these link layers define a segmentation or fragmentation scheme to transport larger IP packets in multiple link layer packets.

Often, adaption layer fragmentation schemes reduce some performance metric, such as the packet delivery rate. Where application or transport protocols have a choice, it would therefore be desirable for them to know about any adaptation layer fragmentation that is going on, so they can choose packet sizes that minimize adaptation layer fragmentation.

At the IP layer, fragmentation can be detected using a number of mechanisms used in Packetization Layer Path MTU Discovery [RFC4821]. However, adaptation later fragmentation schemes are often designed to be "transparent", i.e. there is no way at higher layers to find out they had to be employed (except maybe by elaborate measurement schemes targeting one of the impacted performance metrics; this approach does not appear to be viable) [WEI].

The present specification defines a mechanism for IPv6 adaptation layers to indicate the presence of adaptation layer fragmentation, as well as an indication of preferred packet sizes.

The main objective of this version of the draft is to present a complete design in order to be able to gauge the complexity of the approach against the gains to be expected from implementing it.

Comments are appreciated and should go to the intarea@ietf.org mailing list.

Table of Contents

1. Introduction 4
  1.1. Terminology 4
2. Objectives and Considerations 5
3. The ALFI option 6
4. IANA Considerations 8
5. Security Considerations 9
6. Acknowledgements 10
7. References 11
  7.1. Normative References 11
  7.2. Informative References 11
Author's Address 13

1. Introduction

(To be written - for now please read the Abstract.)

1.1. Terminology

The following terms are used in this specification:

ALF Adaptation Layer Fragmentation

MUALTU Maximum Unfragmented Adaptation Layer Transmission Unit, i.e. the largest piece of IPV6 packet (measured in bytes) that can be transferred by the adaptation layer without invoking ALF

IFMUALTU Initial-Fragment MUALTU, the MUALTU for the initial adaptation layer fragment of an IP packet

FFMUALTU Following-Fragment MUALTU, the estimated minimum MUALTU for all but the initial adaptation layer fragments of an IP packet

ALFI Adaptation Layer Fragmentation Indication, i.e. indication that ALF was performed on a packet

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 RFC 2119. They indicate requirement levels for compliant implementations [RFC2119].

The term "byte" is used in its now customary sense as a synonym for "octet".

2. Objectives and Considerations

This draft is shaped by the requirements of 6LoWPAN networks [I-D.bormann-6lowpan-roadmap], including variants such as Bluetooth/ Low Energy [I-D.ietf-6lowpan-btle] or DECT/ULE [I-D.mariager-6lowpan-v6over-dect-ule]. However, it should be beneficial with any adaptation layer that requires the use of ALF.

One important consideration for ALFI is that the ALF scheme may not be able to provide a consistent MUALTU. E.g., header compression may cause variable overheads, and initial and following fragments are likely to cause different MUALTUs. Header compression may be dependent on the specific characteristics of the packets employed, so indications will be most accurate if they can be made on the basis of actual packets as they are intended to be transferred.

Therefore, ALFI provides the ability to equip packets with a probe that collects for any adaptation layer fragmentation that may be available on the path.

Note that probing for MUALTUs changes the MUALTU. Implementations SHOULD attempt to indicate a MUALTU for a non-probe packet, i.e. the packet under consideration with the ALFI option (and its hop-by-hop header, if applicable) removed. If that is not possible, implementations SHOULD err towards indicating smaller MUALTUs, within reason.

Note that not all nodes will immediately implement ALFI. ALFI just "fails ignorant" (but see below).

Note that an adaptation layer instance may want to manipulate ALFI for other reasons than to indicate ALF (cf. "MSS clamping"). (In particular, while other nodes don't have ALFI yet, a border router such as a 6LBR [I-D.ietf-6lowpan-nd] may want to provide some ALFI guessing.)

Generally speaking, ALFI can be used to indicate any significant, step function degradation of some performance metric based on packet size.

ALFI SHOULD NOT be set for segmentation implementations with limited performance impact. E.g., AAL5 implementations SHOULD NOT set ALFI.

3. The ALFI option

The ALFI option is an IPv6 option in the sense of section 4.2 of [RFC2460]. It is only used in the hop-by-hop header.

The option type identifier is chosen to select the following behavior as detailed in section 4.2 of [RFC2460]:

- 00 - skip over this option and continue processing the header (this enables the "fail-ignorant" backwards compatibility behavior)

- 1 - Option Data may change en-route (the option is used to record information en-route)

```
    .                               +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    .                               |0 0 1 x x x x x|       4       |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
    |    Initial-Fragment MUALTU    |   Following-Fragment MUALTU   |
    +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1

In IFMUALTU and FFMUALTU, the value zero represents infinity. All other values are unsigned integers in network byte order, representing a MUALTU in bytes.

The originator of a packet MAY, for occasional probing, insert an ALFI option into packets where it can choose the packet size and the performance metrics of which are important to the application.

When generating the packet, it sets Initial-Fragment MUALTU (IFMUALTU) and Following-Fragment MUALTU (FFMUALTU) to zero. (Its adaptation layer can then update them as described in the following paragraphs.)

Each instance of an adaptation layer that employs ALF computes its own estimate of IFMUALTU and FFMUALTU for the type of packet that has this option, ignoring the option itself and, if the option was the only option in the hop-by-hop header, the hop-by-hop header. For each estimate, if it is below the existing value (where zero is infinity) of the respective field, the instance updates the field to the estimate.

The receiver of the packet relays the information in the ALFI option to the transport layer and/or application.

(TBD: How to ship this information through the IPv6 socket interface [RFC3493]/[RFC3542]. Constrained implementations won't have this specific problem.)

The transport layer and/or application can then make this information available to the peer instance, which enables it to choose IPv6 packet sizes of IFMUALTU or lower, or, if this cannot be achieved, at least below IFMUALTU+n*FFMUALTU for a small n. For instance, in CoAP [I-D.ietf-core-coap], the Block2 option [I-D.ietf-core-block] can be used to negotiate a block size accordingly.

4. IANA Considerations

IANA needs to allocate an IPv6 option number for the ALFI option, "Destination Options and Hop-by-Hop Options" registry in "Internet Protocol Version 6 (IPv6) Parameters", with act=00 and chg=1 (i.e., similar to the Quick-Start option [RFC4782]).

5. Security Considerations

It is hard to like hop-by-hop options from a security point of view.

(This section will certainly grow as additional security considerations beyond those listed in the base specifications become known.)

6. Acknowledgements

Peter van der Stok prompted the author to finally write up this protocol, a couple of years after the need for it had been shown in [WEI].

7. References

7.1. 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.

7.2. Informative References

[I-D.bormann-6lowpan-roadmap]
Bormann, C., "6LoWPAN Roadmap and Implementation Guide", draft-bormann-6lowpan-roadmap-01 (work in progress), March 2012.

[I-D.ietf-6lowpan-btle]
Nieminen, J., Patil, B., Savolainen, T., Isomaki, M., Shelby, Z., and C. Gomez, "Transmission of IPv6 Packets over Bluetooth Low Energy", draft-ietf-6lowpan-btle-06 (work in progress), March 2012.

[I-D.ietf-6lowpan-nd]
Shelby, Z., Chakrabarti, S., and E. Nordmark, "Neighbor Discovery Optimization for Low Power and Lossy Networks (6LoWPAN)", draft-ietf-6lowpan-nd-18 (work in progress), October 2011.

[I-D.ietf-core-block]
Bormann, C. and Z. Shelby, "Blockwise transfers in CoAP", draft-ietf-core-block-08 (work in progress), February 2012.

[I-D.ietf-core-coap]
Shelby, Z., Hartke, K., Bormann, C., and B. Frank, "Constrained Application Protocol (CoAP)", draft-ietf-core-coap-09 (work in progress), March 2012.

[I-D.mariager-6lowpan-v6over-dect-ule]
Mariager, P. and J. Petersen, "Transmission of IPv6 Packets over DECT Ultra Low Energy", draft-mariager-6lowpan-v6over-dect-ule-01 (work in progress), October 2011.

[RFC3493] Gilligan, R., Thomson, S., Bound, J., McCann, J., and W. Stevens, "Basic Socket Interface Extensions for IPv6", RFC 3493, February 2003.

[RFC3542] Stevens, W., Thomas, M., Nordmark, E., and T. Jinmei, "Advanced Sockets Application Program Interface (API) for IPv6", RFC 3542, May 2003.

[RFC4782] Floyd, S., Allman, M., Jain, A., and P. Sarolahti, "Quick-Start for TCP and IP", RFC 4782, January 2007.

[RFC4821] Mathis, M. and J. Heffner, "Packetization Layer Path MTU Discovery", RFC 4821, March 2007.

[WEI]   Shelby, Z. and C. Bormann, "6LoWPAN: the Wireless Embedded Internet", ISBN 9780470747995, 2009.

Author's Address

Carsten Bormann
Universitaet Bremen TZI
Postfach 330440
Bremen D-28359
Germany

Phone: +49-421-218-63921
Fax: +49-421-218-7000
Email: cabo@tzi.org