Internet Engineering Task Force (IETF)
Request for Comments: 7872
Category: Informational
ISSN: 2070-1721

F. Gont
SI6 Networks / UTN-FRH
J. Linkova
Google
T. Chown
Jisc
W. Liu
Huawei Technologies

June 2016

Observations on the Dropping of Packets with IPv6 Extension Headers in the Real World

Abstract

This document presents real-world data regarding the extent to which packets with IPv6 Extension Headers (EHs) are dropped in the Internet (as originally measured in August 2014 and later in June 2015, with similar results) and where in the network such dropping occurs. The aforementioned results serve as a problem statement that is expected to trigger operational advice on the filtering of IPv6 packets carrying IPv6 EHs so that the situation improves over time. This document also explains how the results were obtained, such that the corresponding measurements can be reproduced by other members of the community and repeated over time to observe changes in the handling of packets with IPv6 EHs.

Status of This Memo

This document is not an Internet Standards Track specification; it is published for informational purposes.

This document is a product of the Internet Engineering Task Force (IETF). It represents the consensus of the IETF community. It has received public review and has been approved for publication by the Internet Engineering Steering Group (IESG). Not all documents approved by the IESG are a candidate for any level of Internet Standard; see Section 2 of RFC 7841.

Information about the current status of this document, any errata, and how to provide feedback on it may be obtained at http://www.rfc-editor.org/info/rfc7872.

Table of Contents

1. Introduction

IPv6 Extension Headers (EHs) allow for the extension of the IPv6 protocol and provide support for core functionality such as IPv6 fragmentation. While packets employing IPv6 EHs have been suspected to be dropped in some IPv6 deployments, there was not much concrete data on the topic. Some preliminary measurements have been presented in [PMTUD-Blackholes], [Gont-IEPG88], and [Gont-Chown-IEPG89], whereas [Linkova-Gont-IEPG90] presents more comprehensive results on which this document is based.

This document presents real-world data regarding the extent to which packets containing IPv6 EHs are dropped in the Internet, as measured in August 2014 and later in June 2015 with similar results (pending operational advice in this area). The results presented in this document indicate that in the scenarios where the corresponding measurements were performed, the use of IPv6 EHs can lead to packet drops. We note that, in particular, packet drops occurring at transit networks are undesirable, and it is hoped and expected that this situation will improve over time.

2. Support of IPv6 Extension Headers in the Internet

This section summarizes the results obtained when measuring the support of IPv6 EHs on the path towards different types of public IPv6 servers. Two sources of information were employed for the list of public IPv6 servers: the "World IPv6 Launch" site <http://www.worldipv6launch.org> and Alexa's list of the Top 1-Million Web Sites <http://www.alexa.com>. For each list of domain names, the following datasets were obtained:

- Web servers (AAAA records of the aforementioned list)
- Mail servers (MX -> AAAA records of the aforementioned list)
- Name servers (NS -> AAAA records of the aforementioned list)

Duplicate addresses and IPv6 addresses other than global unicast addresses were eliminated from each of those lists prior to obtaining the results included in this document. Additionally, addresses that were found to be unreachable were discarded from the dataset (please see Appendix B for further details).

For each of the aforementioned address sets, three different types of probes were employed:

- IPv6 packets with a Destination Options header of 8 bytes;

- IPv6 packets resulting in two IPv6 fragments of 512 bytes each (approximately); and
- IPv6 packets with a Hop-by-Hop Options header of 8 bytes.

In the case of packets with a Destination Options header and the case of packets with a Hop-by-Hop Options header, the desired EH size was achieved by means of PadN options [RFC2460]. The upper-layer protocol of the probe packets was, in all cases, TCP [RFC793] with the Destination Port set to the service port [IANA-PORT-NUMBERS] of the corresponding dataset. For example, the probe packets for all the measurements involving web servers were TCP segments with the Destination Port set to 80.

Besides obtaining the packet drop rate when employing the aforementioned IPv6 EHs, we tried to identify whether the Autonomous System (AS) dropping the packets was the same as the AS of the destination/target address. This is of particular interest since it essentially reveals whether the packet drops are under the control of the intended destination of the packets. Packets dropped by the destination AS are less of a concern since the device dropping the packets is under the control of the same organization as that to which the packets are destined (hence, it is probably easier to update the filtering policy if deemed necessary). On the other hand, packets dropped by transit ASes are more of a concern since they affect the deployability and usability of IPv6 EHs (including IPv6 fragmentation) by a third party (the destination AS). In any case, we note that it is impossible to tell whether, in those cases where IPv6 packets with EHs get dropped, the packet drops are the result of an explicit and intended policy or the result of improper device configuration defaults, buggy devices, etc. Thus, packet drops that occur at the destination AS might still prove to be problematic.

Since there is some ambiguity when identifying the AS to which a specific router belongs (see Appendix B.2), each of our measurements results in two different values: one corresponding to the "best-case scenario" and one corresponding to the "worst-case scenario". The "best-case scenario" is that in which, when in doubt, the packets are assumed to be dropped by the destination AS, whereas the "worst-case scenario" is that in which, when in doubt, the packets are assumed to be dropped by a transit AS (please see Appendix B.2 for details). In the following tables, the values shown within parentheses represent the possibility that, when a packet is dropped, the packet drop occurs in an AS other than the destination AS (considering both the best-case scenario and the worst-case scenario).

Dataset	DO8	HBH8	FH512
Web servers	11.88% (17.60%/20.80%)	40.70% (31.43%/40.00%)	30.51% (5.08%/6.78%)
Mail servers	17.07% (6.35%/26.98%)	48.86% (40.50%/65.42%)	39.17% (2.91%/12.73%)
Name servers	15.37% (14.29%/33.46%)	43.25% (42.49%/72.07%)	38.55% (3.90%/13.96%)

Table 1: WIPv6LD Dataset: Packet Drop Rate for Different Destination Types, and Estimated (Best-Case / Worst-Case) Percentage of Packets That Were Dropped in a Different AS

NOTE: In the tables above and below, "HBH8" stands for "packets with a Hop-By-Hop Options extension header of 8 bytes", "DO8" stands for "packets with a Destination Options extension header of 8 bytes", and "FH512" stands for "IPv6 packets with a Fragment Header of 512 bytes".

NOTE: As an example, we note that the cell describing the support of IPv6 packets with DO8 for web servers (containing the value "11.88% (17.60%/20.80%)") should be read as: "when sending IPv6 packets with DO8 to public web servers, 11.88% of such packets get dropped. Among those packets that get dropped, 17.60%/20.80% (best case / worst case) of them get dropped at an AS other than the destination AS".

Dataset	DO8	HBH8	FH512
Web servers	10.91% (46.52%/53.23%)	39.03% (36.90%/46.35%)	28.26% (53.64%/61.43%)
Mail servers	11.54% (2.41%/21.08%)	45.45% (41.27%/61.13%)	35.68% (3.15%/10.92%)
Name servers	21.33% (10.27%/56.80%)	54.12% (50.64%/81.00%)	55.23% (5.66%/32.23%)

Table 2: Alexa's Top 1M Sites Dataset: Packet Drop Rate for Different Destination Types, and Estimated (Best-Case / Worst-Case) Percentage of Packets That Were Dropped in a Different AS

There are a number of observations to be made based on the results presented above. Firstly, while it has been generally assumed that it is IPv6 fragments that are dropped by operators, our results indicate that it is IPv6 EHs in general that result in packet drops. Secondly, our results indicate that a significant percentage of such packet drops occurs in transit ASes; that is, the packet drops are not under the control of the same organization as the final destination.

3. Security Considerations

This document presents real-world data regarding the extent to which IPv6 packets employing EHs are dropped in the Internet. As such, this document does not introduce any new security issues.

4. References

4.1. Normative References

[RFC793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, DOI 10.17487/RFC0793, September 1981, <http://www.rfc-editor.org/info/rfc793>.

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

4.2. Informative References

[Gont-Chown-IEPG89] Gont, F. and T. Chown, "A Small Update on the Use of IPv6 Extension Headers", IEPG meeting before IETF 89, March 2014, <http://www.iepg.org/2014-03-02-ietf89/fgont-iepg-ietf89-eh-update.pdf>.

[Gont-IEPG88] Gont, F., "Fragmentation and Extension Header Support in the IPv6 Internet", IEPG meeting before IETF 88, November 2013, <http://www.iepg.org/2013-11-ietf88/fgont-iepg-ietf88-ipv6-frag-and-eh.pdf>.

[IANA-PORT-NUMBERS] IANA, "Service Name and Transport Protocol Port Number Registry", <http://www.iana.org/assignments/service-names-port-numbers>.

[IPv6-Toolkit]
SI6 Networks, "SI6 Networks' IPv6 Toolkit v2.0 (Guille)", <http://www.si6networks.com/tools/ipv6toolkit>.

[Linkova-Gont-IEPG90]
Linkova, J. and F. Gont, "IPv6 Extension Headers in the Real World v2.0", IEPG Meeting before IETF 90, July 2014, <http://www.iepg.org/2014-07-20-ietf90/iepg-ietf90-ipv6-ehs-in-the-real-world-v2.0.pdf>.

[PMTUD-Blackholes]
De Boer, M. and J. Bosma, "Discovering Path MTU black holes on the Internet using RIPE Atlas", July 2012, <http://www.nlnetlabs.nl/downloads/publications/pmtu-black-holes-msc-thesis.pdf>.

Appendix A. Reproducing Our Experiment

This section describes, step by step, how to reproduce the experiment with which we obtained the results presented in this document. Each subsection represents one step in the experiment. The tools employed for the experiment are traditional UNIX-like tools (such as gunzip) and the SI6 Networks' IPv6 Toolkit v2.0 (Guille) [IPv6-Toolkit].

Throughout this appendix, "#" denotes the command-line prompt for commands that require superuser privileges, whereas "$" denotes the prompt for commands that do not require superuser privileges.

A.1. Obtaining the List of Domain Names

The primary data source employed was Alexa's Top 1M web sites, available at: <http://s3.amazonaws.com/alexa-static/top-1m.csv.zip>. The file is a zipped file containing the list of the most popular web sites, in Comma-Separated Value (CSV) format. The aforementioned file can be extracted with

```
$ gunzip < top-1m.csv.zip > top-1m.csv
```

A list of domain names (i.e., with other data stripped) can be obtained with the following command [IPv6-Toolkit]:

```
$ cat top-1m.csv | script6 get-alexa-domains > top-1m.txt
```

This command will create a "top-1m.txt" file containing one domain name per line.

> NOTE: The domain names corresponding to the WIPv6LD dataset is available at <http://www.si6networks.com/datasets/wipv6day-domains.txt>. Since the corresponding file is a text file containing one domain name per line, the steps produced in this subsection need not be performed. The WIPv6LD dataset should be processed in the same way as the Alexa dataset, starting from Appendix A.2.

A.2. Obtaining AAAA Resource Records

The file obtained in the previous subsection contains a list of domain names that correspond to web sites. The AAAA records for such domain names can be obtained with:

```
$ cat top-1m.txt | script6 get-aaaa > top-1m-web-aaaa.txt
```

The AAAA records corresponding to the mail servers of each of the aforementioned domain names can be obtained with:

```
$ cat top-1m.txt | script6 get-mx | script6 get-aaaa >
top-1m-mail-aaaa.txt
```

The AAAA records corresponding to the name servers of each of the aforementioned domain names can be obtained with:

```
$ cat top-1m.txt | script6 get-ns | script6 get-aaaa >
top-1m-dns-aaaa.txt
```

A.3. Filtering the IPv6 Address Datasets

The lists of IPv6 addresses obtained in the previous step could possibly contain undesired addresses (e.g., non-global unicast addresses) and/or duplicate addresses. In order to remove both undesired and duplicate addresses, each of the three files from the previous section should be filtered accordingly:

```
$ cat top-1m-web-aaaa.txt | addr6 -i -q -B multicast -B unspec -k
global > top-1m-web-aaaa-unique.txt

$ cat top-1m-mail-aaaa.txt | addr6 -i -q -B multicast -B unspec -k
global > top-1m-mail-aaaa-unique.txt

$ cat top-1m-dns-aaaa.txt | addr6 -i -q -B multicast -B unspec -k
global > top-1m-dns-aaaa-unique.txt
```

A.4. Performing Measurements with Each IPv6 Address Dataset

A.4.1. Measurements with Web Servers

In order to measure DO8 with the list of web servers:

```
# cat top-1m-web-aaaa-unique.txt | script6 trace6 do8 tcp 80 >
top-1m-web-aaaa-do8-m.txt
```

In order to measure HBH8 with the list of web servers:

```
# cat top-1m-web-aaaa-unique.txt | script6 trace6 hbh8 tcp 80 >
top-1m-web-aaaa-hbh8-m.txt
```

In order to measure FH512 with the list of web servers:

```
# cat top-1m-web-aaaa-unique.txt | script6 trace6 fh512 tcp 80 >
top-1m-web-aaaa-fh512-m.txt
```

A.4.2. Measurements with Mail Servers

In order to measure DO8 with the list of mail servers:

```
# cat top-1m-mail-aaaa-unique.txt | script6 trace6 do8 tcp 25 >
top-1m-mail-aaaa-do8-m.txt
```

In order to measure HBH8 with the list of mail servers:

```
# cat top-1m-mail-aaaa-unique.txt | script6 trace6 hbh8 tcp 25 >
top-1m-mail-aaaa-hbh8-m.txt
```

In order to measure FH512 with the list of mail servers:

```
# cat top-1m-mail-aaaa-unique.txt | script6 trace6 fh512 tcp 25 >
top-1m-mail-aaaa-fh512-m.txt
```

A.4.3. Measurements with Name Servers

In order to measure DO8 with the list of name servers:

```
# cat top-1m-dns-aaaa-unique.txt | script6 trace6 do8 tcp 53 >
top-1m-dns-aaaa-do8-m.txt
```

In order to measure HBH8 with the list of name servers:

```
# cat top-1m-dns-aaaa-unique.txt | script6 trace6 hbh8 tcp 53 >
top-1m-dns-aaaa-hbh8-m.txt
```

In order to measure FH512 with the list of name servers:

```
# cat top-1m-dns-aaaa-unique.txt | script6 trace6 fh512 tcp 53 >
top-1m-dns-aaaa-fh512-m.txt
```

A.5. Obtaining Statistics from Our Measurements

A.5.1. Statistics for Web Servers

In order to compute the statistics corresponding to our measurements of DO8 with the list of web servers:

```
$ cat top-1m-web-aaaa-do8-m.txt | script6 get-trace6-stats >
top-1m-web-aaaa-do8-stats.txt
```

In order to compute the statistics corresponding to our measurements of HBH8 with the list of web servers:

```
$ cat top-1m-web-aaaa-hbh8-m.txt | script6 get-trace6-stats >
top-1m-web-aaaa-hbh8-stats.txt
```

In order to compute the statistics corresponding to our measurements of FH512 with the list of web servers:

```
$ cat top-1m-web-aaaa-fh512-m.txt | script6 get-trace6-stats >
top-1m-web-aaaa-fh512-stats.txt
```

A.5.2. Statistics for Mail Servers

In order to compute the statistics corresponding to our measurements of DO8 with the list of mail servers:

```
$ cat top-1m-mail-aaaa-do8-m.txt | script6 get-trace6-stats >
top-1m-mail-aaaa-do8-stats.txt
```

In order to compute the statistics corresponding to our measurements of HBH8 with the list of mail servers:

```
$ cat top-1m-mail-aaaa-hbh8-m.txt | script6 get-trace6-stats >
top-1m-mail-aaaa-hbh8-stats.txt
```

In order to compute the statistics corresponding to our measurements of FH512 with the list of mail servers:

```
$ cat top-1m-mail-aaaa-fh512-m.txt | script6 get-trace6-stats >
top-1m-mail-aaaa-fh512-stats.txt
```

A.5.3. Statistics for Name Servers

In order to compute the statistics corresponding to our measurements of DO8 with the list of name servers:

```
$ cat top-1m-dns-aaaa-do8-m.txt | script6 get-trace6-stats >
top-1m-dns-aaaa-do8-stats.txt
```

In order to compute the statistics corresponding to our measurements of HBH8 with the list of mail servers:

```
$ cat top-1m-dns-aaaa-hbh8-m.txt | script6 get-trace6-stats >
top-1m-dns-aaaa-hbh8-stats.txt
```

In order to compute the statistics corresponding to our measurements of FH512 with the list of mail servers:

```
$ cat top-1m-dns-aaaa-fh512-m.txt | script6 get-trace6-stats >
top-1m-dns-aaaa-fh512-stats.txt
```

Appendix B. Measurements Caveats

A number of issues have needed some consideration when producing the results presented in this document. These same issues should be considered when troubleshooting connectivity problems resulting from the use of IPv6 EHs.

B.1. Isolating the Dropping Node

Let us assume that we find that IPv6 packets with EHs are being dropped on their way to the destination system 2001:db8:d::1 and that the output of running traceroute towards such destination is:

1. 2001:db8:1:1000::1
2. 2001:db8:2:4000::1
3. 2001:db8:3:4000::1
4. 2001:db8:3:1000::1
5. 2001:db8:4:4000::1
6. 2001:db8:4:1000::1
7. 2001:db8:5:5000::1
8. 2001:db8:5:6000::1
9. 2001:db8:d::1

Additionally, let us assume that the output of EH-enabled traceroute to the same destination is:

1. 2001:db8:1:1000::1
2. 2001:db8:2:4000::1
3. 2001:db8:3:4000::1
4. 2001:db8:3:1000::1
5. 2001:db8:4:4000::1

For the sake of brevity, let us refer to the last-responding node in the EH-enabled traceroute ("2001:db8:4:4000::1" in this case) as "M". Assuming that packets in both traceroutes employ the same path, we'll refer to "the node following the last responding node in the EH-enabled traceroute" ("2001:db8:4:1000::1" in our case), as "M+1", etc.

Based on traceroute information above, which node is the one actually dropping the EH-enabled packets will depend on whether the dropping node filters packets before making the forwarding decision or after

making the forwarding decision. If the former, the dropping node will be M+1. If the latter, the dropping node will be "M".

Throughout this document (and our measurements), we assume that those nodes dropping packets that carry IPv6 EHs apply their filtering policy, and only then, if necessary, forward the packets. Thus, in our example above, the last responding node to the EH-enabled traceroute ("M") is "2001:db8:4:4000::1", and we assume the dropping node to be "2001:db8:4:1000::1" ("M+1").

Additionally, we note that when isolating the dropping node we assume that both the EH-enabled and the EH-free traceroutes result in the same paths. However, this might not be the case.

B.2. Obtaining the Responsible Organization for the Packet Drops

In order to identify the organization operating the dropping node, one would be tempted to lookup the Autonomous System Numbers (ASNs) corresponding to the dropping node. However, assuming that M and M+1 are two peering routers, any of these two organizations could be providing the address space employed for such peering. Or, in the case of an Internet Exchange Point (IXP), the address space could correspond to the IXP AS rather than to any of the participating ASes. Thus, the organization operating the dropping node (M+1) could be the AS for M+1, but it might as well be the AS for M+2. Only when the ASN for M+1 is the same as the ASN for M+2 do we have certainty about who the responsible organization for the packet drops is (see slides 21-23 of [Linkova-Gont-IEPG90]).

In the measurement results presented in Section 2, the aforementioned ambiguity results in a "best-case" and a "worst-case" scenario (rather than a single value): the lowest percentage value means that, when in doubt, we assume the packet drops occur in the same AS as the destination; on the other hand, the highest percentage value means that, when in doubt, we assume the packet drops occur at a different AS than the destination AS.

We note that the aforementioned ambiguity should also be considered when troubleshooting and reporting IPv6 packet drops since identifying the organization responsible for the packet drops might prove to be a non-trivial task.

Finally, we note that a specific organization might be operating more than one AS. However, our measurements assume that different ASNs imply different organizations.

Appendix C. Troubleshooting Packet Drops Due to IPv6 Extension Headers

Isolating IPv6 blackholes essentially involves performing IPv6 traceroute for a destination system with and without IPv6 EHs. The EH-free traceroute would provide the full working path towards a destination while the EH-enabled traceroute would provide the address of the last-responding node for EH-enabled packets (say, "M"). In principle, one could isolate the dropping node by looking-up "M" in the EH-free traceroute with the dropping node being "M+1" (see Appendix B.1 for caveats).

At the time of this writing, most traceroute implementations do not support IPv6 EHs. However, the path6 tool of [IPv6-Toolkit] provides such support. Additionally, the blackhole6 tool of [IPv6-Toolkit] automates the troubleshooting process and can readily provide information such as: dropping node's IPv6 address, dropping node's AS, etc.

Acknowledgements

The authors would like to thank (in alphabetical order) Mikael Abrahamsson, Mark Andrews, Fred Baker, Brian Carpenter, Gert Doering, C. M. Heard, Nick Hilliard, Joel Jaeggli, Tatuya Jinmei, Merike Kaeo, Warren Kumari, Ted Lemon, Mark Smith, Ole Troan, and Eric Vyncke for providing valuable comments on draft versions of this document. Additionally, the authors would like to thank participants of the V6OPS and OPSEC working groups for their valuable input on the topics discussed in this document.

The authors would like to thank Fred Baker for his guidance in improving this document.

Fernando Gont would like to thank Jan Zorz of Go6 Lab <http://go6lab.si/> and Jared Mauch of NTT America for providing access to systems and networks that were employed to produce some of the measurement results presented in this document. Additionally, he would like to thank SixXS <https://www.sixxs.net> for providing IPv6 connectivity.

Fernando Gont would like to thank Nelida Garcia and Guillermo Gont for their love and support.

Authors' Addresses

Fernando Gont
SI6 Networks / UTN-FRH
Evaristo Carriego 2644
Haedo, Provincia de Buenos Aires 1706
Argentina

Phone: +54 11 4650 8472
Email: fgont@si6networks.com
URI: http://www.si6networks.com

J. Linkova
Google
1600 Amphitheatre Parkway
Mountain View, CA 94043
United States

Email: furry@google.com

Tim Chown
Jisc
Lumen House, Library Avenue
Harwell Oxford, Didcot OX11 0SG
United Kingdom

Email: tim.chown@jisc.ac.uk

Will (Shucheng) Liu
Huawei Technologies
Bantian, Longgang District
Shenzhen 518129
China

Email: liushucheng@huawei.com