IPv6 Extension Header (Performance and Diagnostic Metrics (PDM) Destination Option) Testing Across the Internet

IEPG: IETF114

N. Elkins: Industry Network Technology Council

M. Ackermann: Industry Network Technology Council

A. Deshpande: NITK Surathkal

Can IPv6 Extension Headers Be Used on the Internet?

- Controversy for many years
- A number of studies showing that IPv6 extension headers “don’t work”
- Studies (by and large) sent “fake” IPv6 extension headers to Alexa top n sites
- If this is true, our work on our IPv6 Extension Header Destination Option Performance and Diagnostic Metrics (PDM) is really for naught

Brief explanation of PDM

- RFC8250: IPv6 Performance and Diagnostic Metrics (PDM) Destination Option

- To assess performance problems, this document describes optional headers embedded in each packet that provide sequence numbers and timing information as a basis for measurements. Such measurements may be interpreted in real time or after the fact. This document specifies the Performance and Diagnostic Metrics (PDM) Destination Options header.

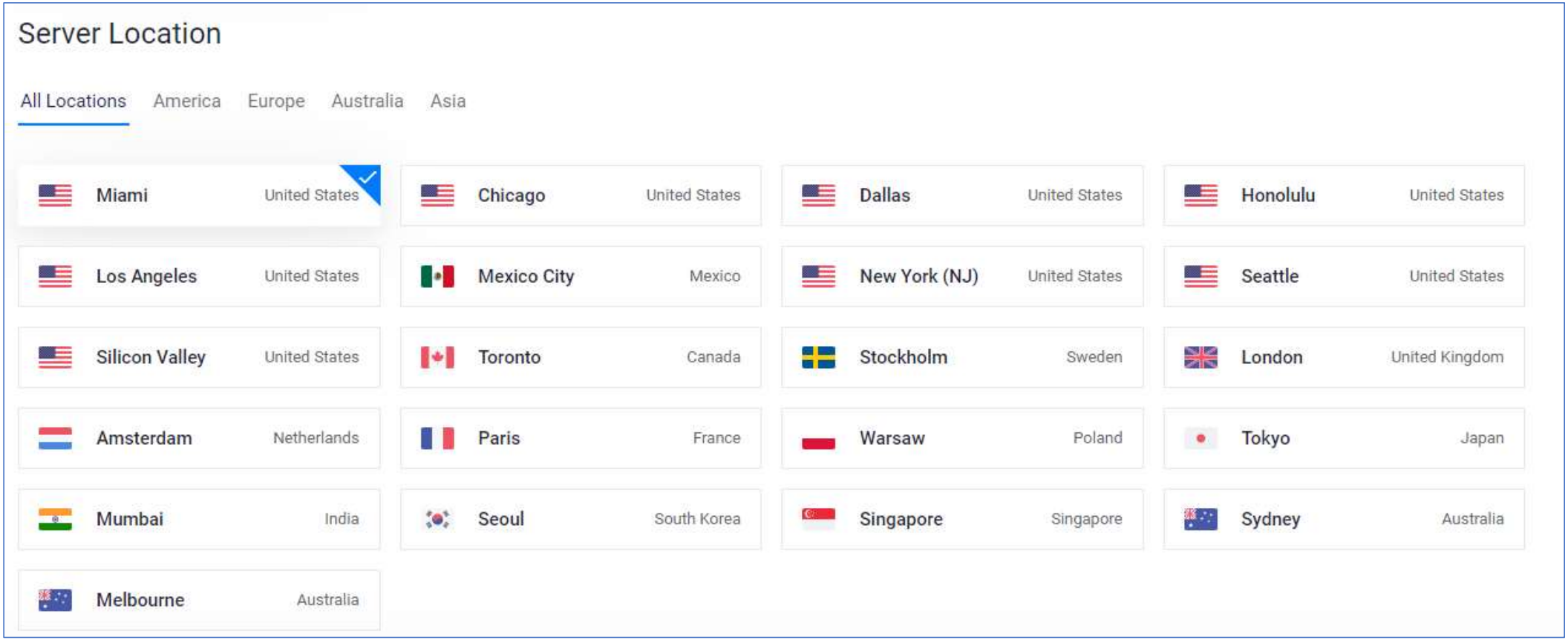
What we did

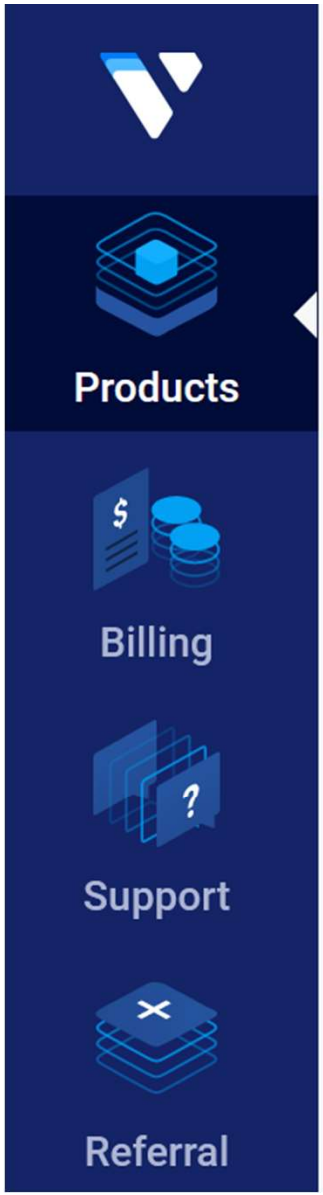
- Used a small hosting service (not one of the “brand-name” ones)

- Locations throughout the world

1. PDM-Warsaw
2. PDM-Toronto
3. PDM-Seattle
4. PDM-Mumbai
5. PDM-Melbourne
6. PDM-Frankfurt

All machines are FreeBSD with a modification to the kernel to send PDM IPv6 Destination option with every packet


Server Location
All Locations
America
Europe
Australia
Asia
Miami
United States
Chicago
United States
Dallas
United States
Honolulu
United States
Los Angeles
United States
Mexico City
Mexico
New York (NJ)
United States
Seattle
United States
Silicon Valley
United States
Toronto
Canada
Stockholm
Sweden
London
United Kingdom
Amsterdam
Netherlands
Paris
France
Warsaw
Poland
Tokyo
Japan
Mumbai
India
Seoul
South Korea
Singapore
Singapore
Sydney
Australia
Melbourne
Australia



Server	OS	Location
PDM-Frankfurt 2048 MB AMD High Performance - 45.63.117.59		Frankfurt
PDM-Melbourne 2048 MB AMD High Performance - 67.219.99.226		Melbourne
PDM-Mumbai 2048 MB AMD High Performance - 65.20.80.111		Mumbai
PDM-Seattle 2048 MB AMD High Performance - 66.42.67.223		Seattle
PDM-Toronto 2048 MB AMD High Performance - 155.138.139.133		Toronto
PDM-Warsaw 2048 MB AMD High Performance - 70.34.248.166		Warsaw

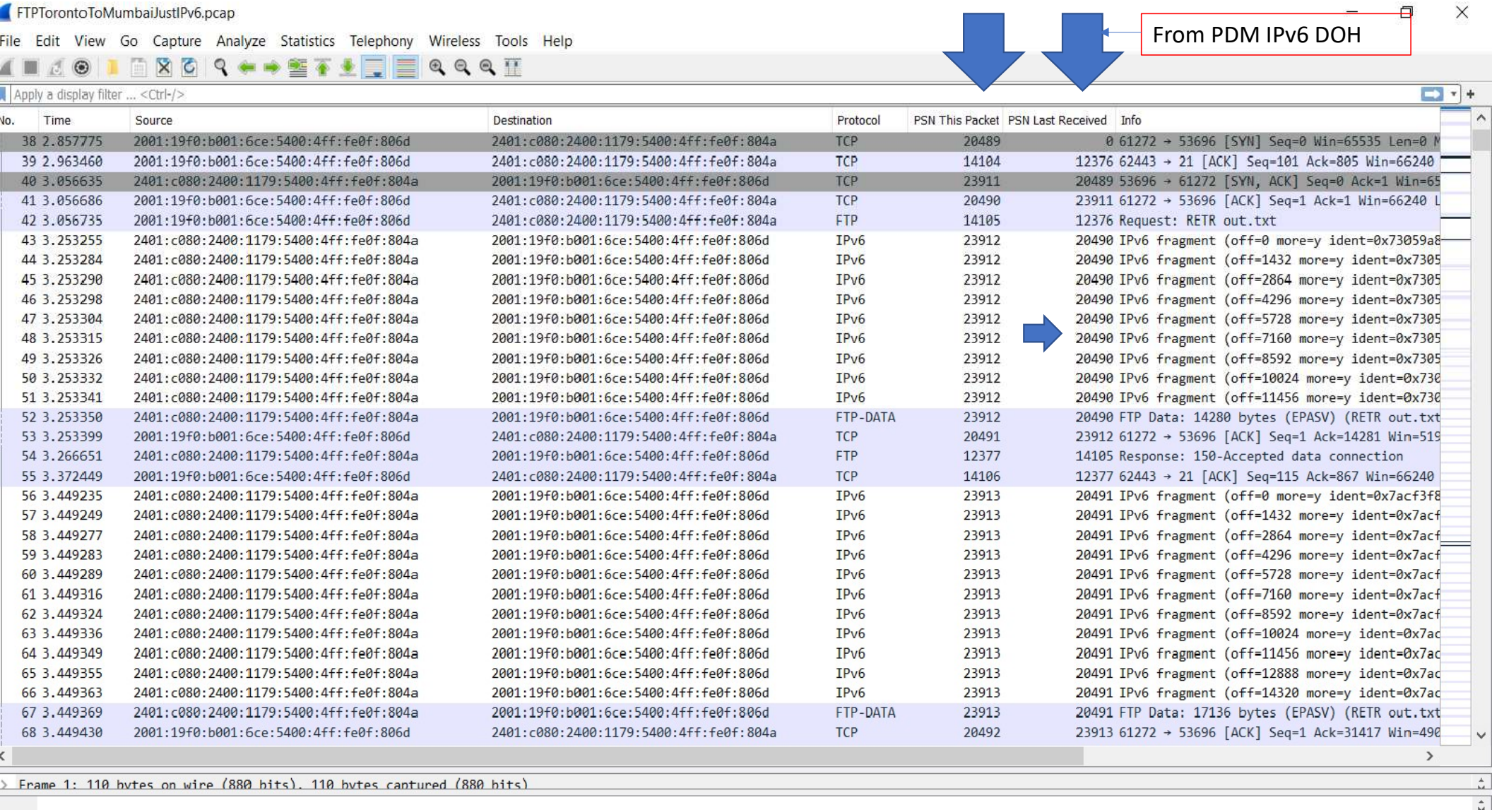
Thanks to…

In particular, Dr. Mohit Tahiliani

Tested large FTP: Toronto to Mumbai (with PDM)

- Connected to **2401:c080:2400:1179:5400:04ff:fe0f:804a.**
- 220---------- Welcome to Pure-FTPd [privsep] [TLS] ----------
- 220-You are user number 1 of 50 allowed.
- 220-Local time is now 15:12. Server port: 21.
- 220 You will be disconnected after 15 minutes of inactivity.
- 331 User PDMuser OK. Password required
- 230 OK. Current directory is /
- Remote system type is UNIX.
- Using binary mode to transfer files.

- 229 Extended Passive mode OK (|||3353|)
- 150-Accepted data connection
- 150 **27872.0 kbytes to download**
- 100% |**********| 27872 KiB 222.31 KiB/s 00:00 ETA
- 226-**File successfully transferred**
- 226 125.107 seconds (measured here), 222.78 Kbytes per second
- 28540928 bytes received in 02:05 (222.31 KiB/s)
- 221-Goodbye. You uploaded 0 and downloaded 27872 kbytes.
- 221 Logout.


FTPTorontoToMumbaiJustIPv6.pcap
File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help
From PDM IPv6 DOH
Apply a display filter ... <Ctrl-/>
No. Time Source Destination Protocol PSN This Packet PSN Last Received Info
38 2.857775 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 20489 0 61272 → 53696 [SYN] Seq=0 Win=65535 Len=0
39 2.963460 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 14104 12376 62443 → 21 [ACK] Seq=101 Ack=805 Win=66240
40 3.056635 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d TCP 23911 20489 53696 → 61272 [SYN, ACK] Seq=0 Ack=1 Win=65
41 3.056686 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 20490 23911 61272 → 53696 [ACK] Seq=1 Ack=1 Win=66240
42 3.056735 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a FTP 14105 12376 Request: RETR out.txt
43 3.253255 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=0 more=y ident=0x73059a8
44 3.253284 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=1432 more=y ident=0x7305
45 3.253290 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=2864 more=y ident=0x7305
46 3.253298 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=4296 more=y ident=0x7305
47 3.253304 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=5728 more=y ident=0x7305
48 3.253315 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=7160 more=y ident=0x7305
49 3.253326 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=8592 more=y ident=0x7305
50 3.253332 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=10024 more=y ident=0x730
51 3.253341 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23912 20490 IPv6 fragment (off=11456 more=y ident=0x730
52 3.253350 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d FTP-DATA 23912 20490 FTP Data: 14280 bytes (EPASV) (RETR out.txt
53 3.253399 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 20491 23912 61272 → 53696 [ACK] Seq=1 Ack=14281 Win=519
54 3.266651 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d FTP 12377 14105 Response: 150-Accepted data connection
55 3.372449 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 14106 12377 62443 → 21 [ACK] Seq=115 Ack=867 Win=66240
56 3.449235 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=0 more=y ident=0x7acf3f8
57 3.449249 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=1432 more=y ident=0x7acf
58 3.449277 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=2864 more=y ident=0x7acf
59 3.449283 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=4296 more=y ident=0x7acf
60 3.449289 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=5728 more=y ident=0x7acf
61 3.449316 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=7160 more=y ident=0x7acf
62 3.449324 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=8592 more=y ident=0x7acf
63 3.449336 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=10024 more=y ident=0x7ac
64 3.449349 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=11456 more=y ident=0x7ac
65 3.449355 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=12888 more=y ident=0x7ac
66 3.449363 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d IPv6 23913 20491 IPv6 fragment (off=14320 more=y ident=0x7ac
67 3.449369 2401:c080:2400:1179:5400:4ff:fe0f:804a 2001:19f0:b001:6ce:5400:4ff:fe0f:806d FTP-DATA 23913 20491 FTP Data: 17136 bytes (EPASV) (RETR out.txt
68 3.449430 2001:19f0:b001:6ce:5400:4ff:fe0f:806d 2401:c080:2400:1179:5400:4ff:fe0f:804a TCP 20492 23913 61272 → 53696 [ACK] Seq=1 Ack=31417 Win=490
Frame 1: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)

Notice in PCAP

- Packet Sequence Number This Packet and Packet Sequence Number Last Received are from PDM IPv6 Destination Option Header
- Surprisingly, packets were broken up. IPv6 Fragment Extension Header also sent
- All successfully

FTPTorontoToMumbaiJustIPv6.pcap

File Edit View Go Capture Analyze Statistics Telephony Wireless Tools Help

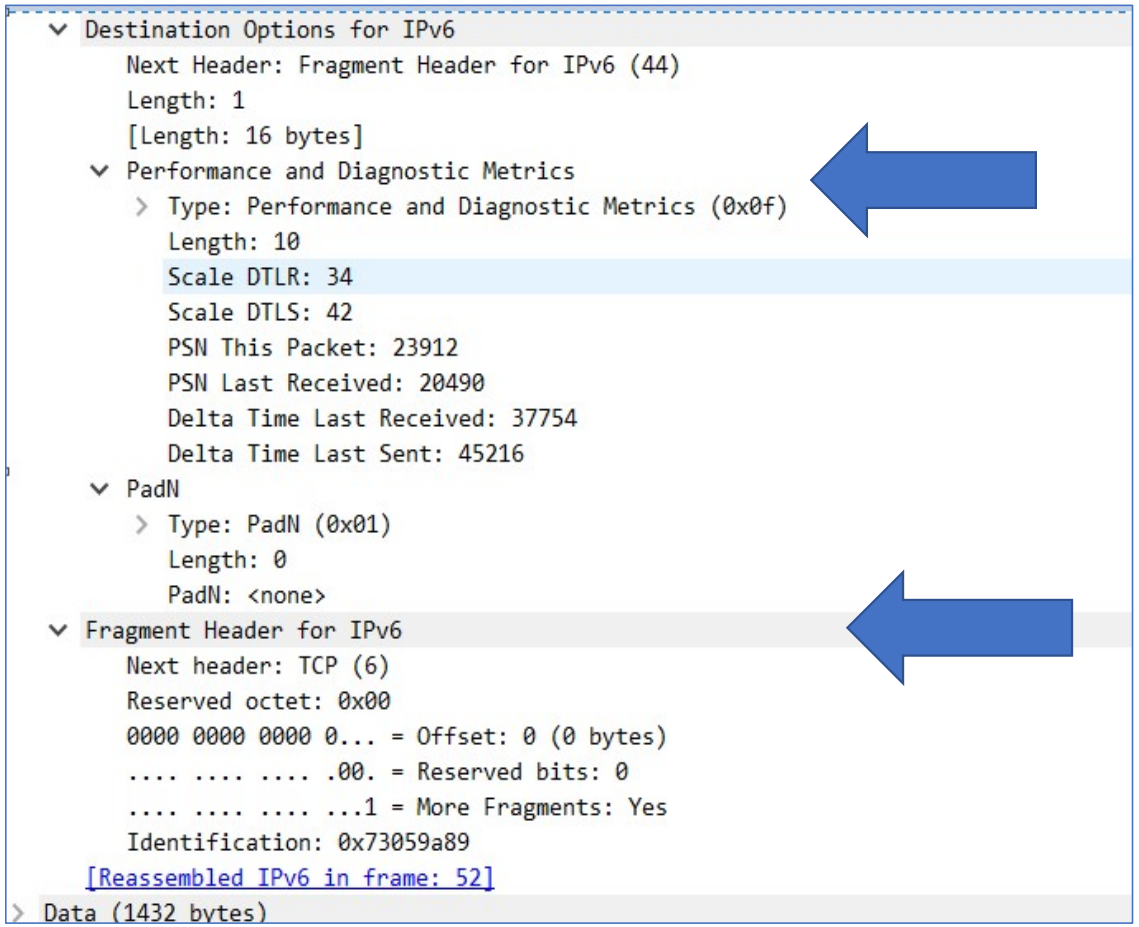
Apply a display filter ... <Ctrl-/>

No.	Time	Source	Destination	Protocol	PSN This Packet	PSN Last Received	Info
41	3.056686	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	TCP	20490	23911	61272 → 53696 [ACK] Seq=1 Ack=1 Win=66240 L
42	3.056735	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	FTP	14105	12376	Request: RETR out.txt
43	3.253255	2401:c080:2400:1179:5400:4ff:fe0f:804a	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	IPv6	23912	20490	IPv6 fragment (off=0 more=y ident=0x73059a8

```
Frame 41: 102 bytes on wire (816 bits), 102 bytes captured (816 bits)
Ethernet II, Src: 56:00:04:0f:80:6d (56:00:04:0f:80:6d), Dst: 86:1f:85:c1:55:77 (86:1f:85:c1:55:77)
Internet Protocol Version 6, Src: 2001:19f0:b001:6ce:5400:4ff:fe0f:806d, Dst: 2401:c080:2400:1179:5400:4ff:fe0f:804a
    0110 .... = Version: 6
  > .... 0000 0000 .... .... .... .... .... = Traffic Class: 0x00 (DSCP: CS0, ECN: Not-ECT)
    .... 1100 1001 0100 0100 1110 = Flow Label: 0xc944e
    Payload Length: 48
    Next Header: Destination Options for IPv6 (60)
    Hop Limit: 64
    Source Address: 2001:19f0:b001:6ce:5400:4ff:fe0f:806d
    Destination Address: 2401:c080:2400:1179:5400:4ff:fe0f:804a
  v Destination Options for IPv6
      Next Header: TCP (6)
      Length: 1
      [Length: 16 bytes]
    v Performance and Diagnostic Metrics
      > Type: Performance and Diagnostic Metrics (0x0f)
        Length: 10
        Scale DTLR: 29
        Scale DTLS: 42
        PSN This Packet: 20490
        PSN Last Received: 23911
        Delta Time Last Received: 50924
        Delta Time Last Sent: 45220
    > PadN
Transmission Control Protocol, Src Port: 61272, Dst Port: 53696, Seq: 1, Ack: 1, Len: 0
```

PDM IPv6 Extension Header Destination Option

Showing both Extension Headers



Bottom line

1. PDM-FTP Toronto to Warsaw - worked
2. PDM-FTP Toronto to Seattle - worked
3. PDM-FTP Toronto to Mumbai - worked
4. PDM-FTP Toronto to Melbourne - worked
5. PDM-FTP Toronto to Frankfurt - worked

Traces available for all to look at.

Come to the Hackathon (or HackDemo) if you want to see for yourself.

Is hosting service using an overlay network?

- Email sent by me
  - I have a question about the connection between various [hostingcompany] instances. For example, if I have an instance in Mumbai and another one in Atlanta, then do you have an overlay network? That is, do you have special connectivity between [hostingcompany] instances or is it going over the Internet?

- Response from hosting company
  - Communication between [hostingcompany] VPS residing in different datacenters will always travel on public internet exchanges. [hostingcompany] Private Cloud can create a private network, however this is only for communication between instances in the same datacenter. [hostingcompany] utilizes multiple transit providers.

```
# traceroute6 2401:c080:2400:1179:5400:04ff:fe0f:804a
traceroute6 to 2401:c080:2400:1179:5400:04ff:fe0f:804a (2401:c080:2400:1179:5400:4ff:fe0f:804a) from
2001:19f0:b001:6ce:5400:4ff:fe0f:806d, 64 hops max, 12 byte packets
 1  * * *
 2  vl199-ds1-n1-r103-b.sao1.constant.com  1.262 ms  0.400 ms  0.334 ms
 3  vl518-ds1-q8.tor1.constant.com  1.192 ms  2.403 ms
    vl818-ds2-q8.tor1.constant.com  1.124 ms
 4  vl25-er1-q2.tor1.constant.com  0.893 ms
    vl25-er2-q2.tor1.constant.com  7.116 ms  1.321 ms
 5  et-1-0-19.cr3-tor1.ip6.gtt.net  0.640 ms *  1.267 ms
 6  2001:668:0:2:ffff:0:5995:8cfd  97.981 ms
    chi-b23-v6.ip.twelve99.net  11.284 ms
    2001:668:0:2:ffff:0:5995:8cfd  98.632 ms
 7  kanc-b2-v6.ip.twelve99.net  22.989 ms  23.535 ms
    2001:668:0:3:ffff:0:4d43:50be  90.619 ms
 8  2404:a800::42  209.484 ms
    kanc-bb2-v6.ip.twelve99.net  23.407 ms  23.405 ms
 9  * dls-bb2-v6.ip.twelve99.net  34.095 ms *
10  vl22-ds2-q8.bom1.constant.com  196.490 ms
    dls-b24-v6.ip.twelve99.net  69.405 ms
    vl22-ds2-q8.bom1.constant.com  192.592 ms
11  dls-bb1-v6.ip.twelve99.net  33.698 ms  33.900 ms
    vl810-ds1-m1-c3r0407-a.bom1.constant.com  209.905 ms
12  * * *
13  * * 2401:c080:2400:1179:5400:4ff:fe0f:804a  192.645 ms
```

Traceroute showing multiple transit providers

Why are our results so different from others?

- We are using real data and a real application (e.g. PDM and FTP)
- We are NOT going to the Alexa top n
- But, we also tried to replicate the results of others
- Indeed, if you use the large hosting companies and go to the Alexa top n, there are issues
- But why?

Looking at Traceroutes with and w/out PDM

- We did PING and UDP TraceRoutes from our PDM enabled machine to many well-known sites (and some not so well-known)

- **What we are looking for is if last ICMPv6 with PDM was at hop 8, then what was hop 9 according to the non-PDM traceroute?**

- **That is, did it already get to the destination network?**

Traceroute Packet Capture

No.	Time	Source	Destination	Protocol	PSN This Pacl	Hop Limit	Info
1	0.000000	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	37868	1	53842 → 33435 Len=12
4	9.292686	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	51235	1	53842 → 33436 Len=12
7	14.698438	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	49557	1	53842 → 33437 Len=12
9	20.316429	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	48949	2	53842 → 33438 Len=12
12	20.317605	2001:19f0:fc01:b::6464:c801	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	48949	63,1	Time Exceeded (hop limit exceeded in transit)
13	20.327161	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	54702	2	53842 → 33439 Len=12
14	20.327495	2001:19f0:fc01:b::6464:c801	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	54702	63,1	Time Exceeded (hop limit exceeded in transit)
15	20.327563	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	10092	2	53842 → 33440 Len=12
16	20.327846	2001:19f0:fc01:b::6464:c801	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	10092	63,1	Time Exceeded (hop limit exceeded in transit)
17	20.327905	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	47810	3	53842 → 33441 Len=12
18	20.329037	2001:19f0:fc00::a53:ad	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	47810	62,1	Time Exceeded (hop limit exceeded in transit)
19	20.337834	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	44722	3	53842 → 33442 Len=12
20	20.340170	2001:19f0:fc00::a53:ad	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	44722	62,1	Time Exceeded (hop limit exceeded in transit)
21	20.340250	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	36784	3	53842 → 33443 Len=12
22	20.341318	2001:19f0:fc00::a53:b1	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	36784	62,1	Time Exceeded (hop limit exceeded in transit)
23	20.343616	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	59954	4	53842 → 33444 Len=12
24	20.344459	2001:19f0:fc00::a53:41	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	59954	61,1	Time Exceeded (hop limit exceeded in transit)
25	20.346746	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	4520	4	53842 → 33445 Len=12
26	20.353806	2001:19f0:fc00::a53:109	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	4520	61,1	Time Exceeded (hop limit exceeded in transit)
27	20.356364	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	13709	4	53842 → 33446 Len=12
28	20.357629	2001:19f0:fc00::a53:109	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	13709	61,1	Time Exceeded (hop limit exceeded in transit)
29	20.357689	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	46433	5	53842 → 33447 Len=12
30	20.358279	2001:668:0:3:ffff:2:0:939	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	46433	59,1	Time Exceeded (hop limit exceeded in transit)
31	21.274626	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	18819	5	53842 → 33448 Len=12
32	26.358606	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	2401:c080:2400:1179:5400:4ff:fe0f:804a	UDP	36308	5	53842 → 33449 Len=12
33	26.359776	2001:668:0:3:ffff:2:0:939	2001:19f0:b001:6ce:5400:4ff:fe0f:806d	ICMPv6	36308	59,1	Time Exceeded (hop limit exceeded in transit)

Summarizing Information being collected per traceroute

- Company
- Starting IP address and DNS name
- Destination IP address and DNS name (it may resolve to CDN, for example)
- Hop number where last PDM ICMPv6 came from (ex. 8). Let this be "n".
- Destination DNS name where last PDM ICMPv6 came from (ex. Tor)
- Hop number where last non-PDM ICMPv6 came from (ex. 10)
- Destination DNS name where last non-PDM ICMPv6 came from (ex. Cloudflare)
- Destination DNS name of hop n+1 (the next hop that the PDM ICMPv6 would have come to) (ex. Akamai)

- **What we are looking for is if last ICMPv6 with PDM was at hop 8, then what was hop 9 according to the non-PDM traceroute?**

Results

- Will present traceroute results next time!
- We welcome collaborators! Preparing VM image of FreeBSD with PDM so you can test for yourself.
- Come talk to us at Hackathon and/or HackDemo