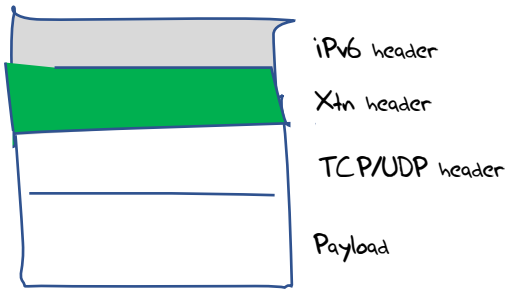
Measurement of IPv6 Extension Header Support

Geoff Huston
APNIC Labs

IPv6 Extension Header



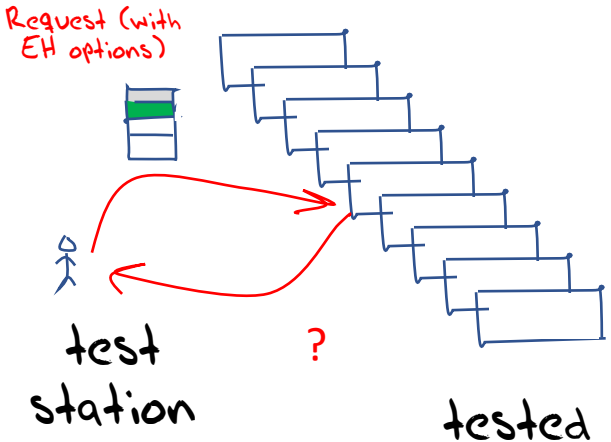
The extension header sits between the IPv6 packet header and the upper level protocol header for the leading fragged packet, and sits between the header and the trailing payload frags for the trailing packets

Practically, this means that transport-protocol aware packet processors/switches need to decode the extension header chain, if its present, which can consume additional cycles to process/switch a packet – and the additional time is not predictable. For trailing frags there is no transport header!

Or the unit can simply discard all IPv6 packets that contain extension headers - which is what a lot of transport protocol sensitive IPv6 deployed switching equipment appears to do!

RFC 7872

June 2016



One-to-many test sending sets of well-known servers requests where EH options are added to the outbound packets

The test is whether or not the server sends a response

Tested Destination Options, Hop-by-hop and Fragments

tested servers

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%)

IPv6 EH Fragmentation Handling

There is a lot of “drop” behaviour in the IPv6 Internet for Fragmentation Extension headers

RFC7872 – recorded EH packet drop rates of 30% - 55%

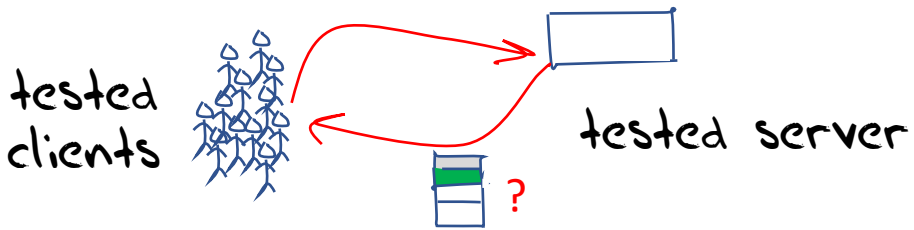
But sending fragmented queries to servers is not all that common – the reverse situation of big responses is more common

So what about sending fragmented packets BACK from servers – what’s the drop rate of the reverse case?

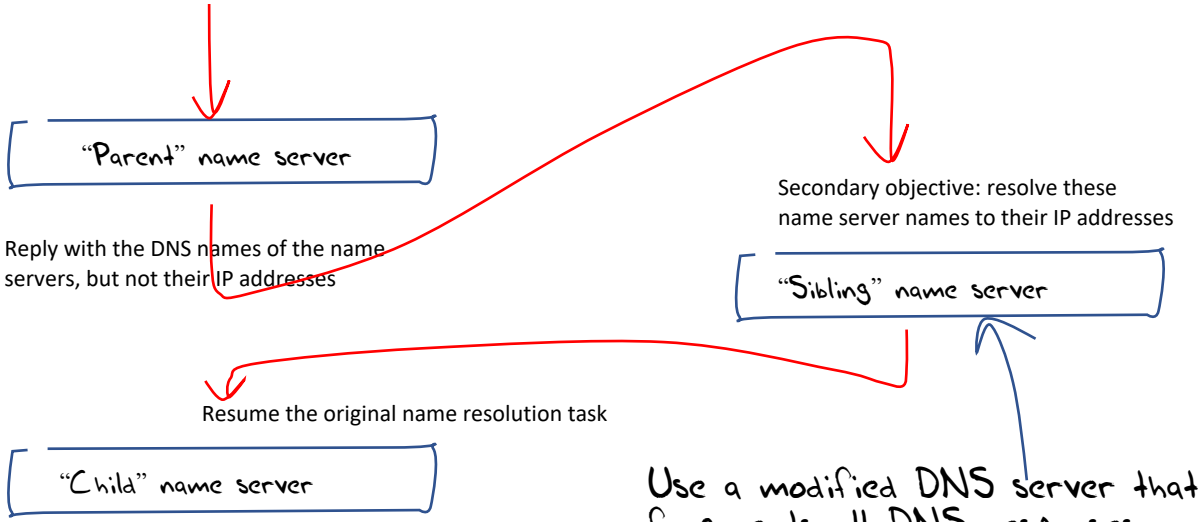
Our Measurement Approach

We use an Online Ad platform to enroll endpoints to attempt to resolve a set of DNS names:

- Each endpoint is provided with a unique name string (to eliminate the effects of DNS caching)
- The DNS name is served from our authoritative servers
- Resolving the DNS name requires the user's DNS resolvers to receive a fragmented IPv6 packet



"Glueless" Delegation to detect IPv6 Fragmentation Handling



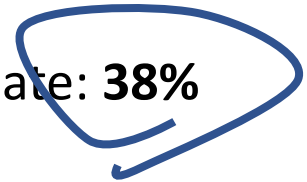
The "child" name server will only be queried if the resolver could receive the response from the sibling name server

V6, the DNS and Fragmented UDP

Total number of tests: 10,851,323

Failure Rate in receiving a large response: 4,064,356

IPv6 Fragmentation Failure Rate: **38%**



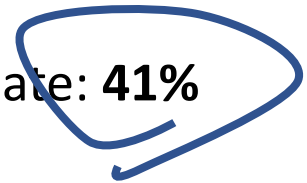
2017 data

V6, the DNS and Fragmented UDP

Total number of tests: 27,619,047

Failure Rate in receiving a large response: 11,232,833

IPv6 Fragmentation Failure Rate: **41%**



2020 data

Which Resolvers?

- 10,115 IPv6 seen resolvers
- 3,592 resolvers were consistently unable to resolve the target name (likely due to failure to receive the fragmented response)
- Which is too large a list to display here
- But we can show the top 20…

Which Resolvers?

Resolver	Hits	AS	AS Name	CC
2405:200:1606:672::5	4,178,119	55836	RELIANCEJIO-IN Reliance Jio Infocomm Limited	IN
2402:8100:c::8	1,352,024	55644	IDEANET1-IN Idea Cellular Limited	IN
2402:8100:c::7	1,238,764	55644	IDEANET1-IN Idea Cellular Limited	IN
2407:0:0:2b::5	938,584	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2a::3	936,883	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2a::6	885,322	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2b::6	882,687	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2b::2	882,305	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2a::4	881,604	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2a::5	880,870	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2a::2	877,329	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2b::4	876,723	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:2b::3	876,150	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2402:8100:d::8	616,037	55644	IDEANET1-IN Idea Cellular Limited	IN
2402:8100:d::7	426,648	55644	IDEANET1-IN Idea Cellular Limited	IN
2407:0:0:9::2	417,184	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:8::2	415,375	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:8::4	414,410	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:9::4	414,226	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
2407:0:0:9::6	411,993	4761	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID

All these resolvers appears to be unable to receive fragmented UDP DNS responses – This is the Top 20, as measured by the query count per resolver address

Resolvers in Which Networks?

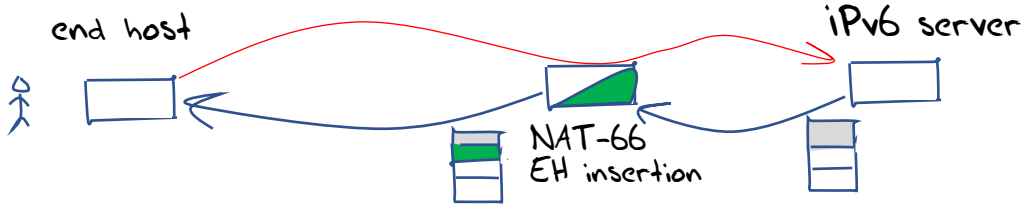
AS	Hits	% of Total	AS Name	CC
15169	7,952,272	17.3%	GOOGLE - Google Inc.	US
4761	6,521,674	14.2%	INDOSAT-INP-AP INDOSAT Internet Network Provider	ID
55644	4,313,225	9.4%	IDEANET1-IN Idea Cellular Limited	IN
22394	4,217,285	9.2%	CELLCO - Cellco Partnership DBA Verizon Wireless	US
55836	4,179,921	9.1%	RELIANCEJIO-IN Reliance Jio Infocomm Limited	IN
10507	2,939,364	6.4%	SPCS - Sprint Personal Communications Systems	US
5650	2,005,583	4.4%	FRONTIER-FRTR - Frontier Communications of America	US
2516	1,322,228	2.9%	KDDI KDDI CORPORATION	JP
6128	1,275,278	2.8%	CABLE-NET-1 - Cablevision Systems Corp.	US
32934	1,128,751	2.5%	FACEBOOK - Facebook	US
20115	984,165	2.1%	CHARTER-NET-HKY-NC - Charter Communications	US
9498	779,603	1.7%	BBIL-AP BHARTI Airtel Ltd.	IN
20057	438,137	1.0%	ATT-MOBILITY-LLC-AS20057 - AT&T Mobility LLC	US
17813	398,404	0.9%	MTNL-AP Mahanagar Telephone Nigam Ltd.	IN
2527	397,832	0.9%	SO-NET So-net Entertainment Corporation	JP
45458	276,963	0.6%	SBN-AWN-AS-02-AP SBN-ISP/AWN-ISP and SBN-NIX/AWN-NIX	TH
6167	263,583	0.6%	CELLCO-PART - Cellco Partnership DBA Verizon Wireless	US
8708	255,958	0.6%	RCS-RDS 73-75 Dr. Staicovici	RO
38091	255,930	0.6%	HELLONET-AS-KR CJ-HELLOVISION	KR
18101	168,164	0.4%	Reliance Communications DAKC MUMBAI	IN

This is the total per origin AS of those resolvers that appear to be unable to receive fragmented UDP DNS responses. This is the Top 20, as measured by the query count per origin AS

What about TCP and the IPv6 Fragmentation Header?

Let's try the same approach:

- Set up an ad-based measurement using a customised IPv6 packet handler
- Pass all TCP responses through a packet fragmenter
  - Use an IPv6 NAT-66 implementation that takes a server's IPv6 packets and wrangles them to include an EH header before passing them back to the client
  - In this case any packet with size > 512 octets was mangled to fragment at a 512 octets
- Use a packet capture to see if the fragmented TCP segment was ACKed or not



What about TCP and IPv6 Fragmentation?

1,961,561 distinct IPv6 end point addresses

434,971 failed to receive Fragmented IPv6 packets

22% failure rate

Where are TCP e-2-e drops?

AS	Samples	Failure Rate	AS Name	CC
3598	4,762	99.4%	MICROSOFT-CORP-AS - Microsoft Corporation	US
15169	6,426	98.9%	GOOGLE - Google Inc.	US
24961	252	98.4%	MYLOC-AS	DE
6621	4,431	92.8%	HNS-DIRECPC - Hughes Network Systems	US
131222	595	89.1%	MTS-INDIA-IN 334, Udyog, Vihar	IN
38229	260	86.5%	LEARN-LK Lanka Education & Research Network	LK
6939	106,057	85.2%	HURRICANE - Hurricane Electric	US
852	4,552	84.1%	ASN852 - TELUS Communications Inc.	CA
32934	359	79.7%	FACEBOOK - Facebook	US
54115	128	78.9%	FACEBOOK-CORP - Facebook Inc	US
1312	122	76.2%	Virginia Polytechnic Institute and State Univ.	US
22394	109,333	73.2%	CELLCO - Cellco Partnership DBA Verizon Wireless	US
5603	1,938	69.3%	SIOL-NET	SI
4134	171	69.0%	CHINANET-BACKBONE No.31	CN
20845	272	68.4%	DIGICABLE	HU

Top 15 networks with highest Fragmented IPv6 Drop Rates

Why do we see these high packet drop rates?

Two major factors appear to lie behind this failure rate:

- Network equipment dropping IPv6 packets with Extension Headers
- Firewalls dropping Fragmented packets

Next Measurement Steps?

Test other Extension Headers

- Hop-by-Hop Extension headers
- Destination Extension Headers

Compare TCP and UDP drop performance

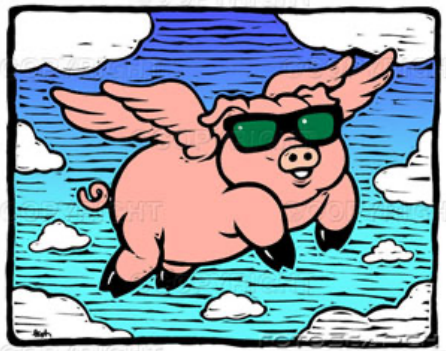
Locate Drop Point

- at end point?
- in flight?

But

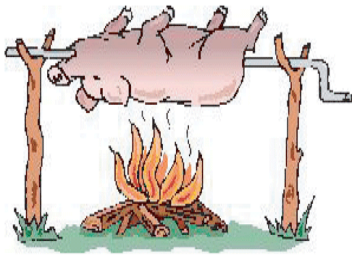
Can we fix the network anyway?

- Or is this just an exercise in trying to make the pig fly?



Or

- Like the fate of IPv4 options, just forget about using them, and declare IPv6 EH headers a bad idea!



Or, but

- If we forget about IPv6 EH then IPv6 fragmentation is no longer possible
- And that’s puts a huge strain on IPv6 UDP applications
- Like the DNS!
  - And we really don’t have a good answer for that so far!

What's the real question here?

What else do we need to understand about networks and end stack behaviours in IPv6 in order to figure out whether to abandon EH completely or try to salvage bits of it and make those bits work everywhere?

Thanks!