DNS and IPv6

Geoff Huston
Joao Damas
APNIC Labs

This is not a talk about DoH and DoT

For over a year DoH is still the current hot topic in the DNS

- and there is much that can be said about DoH and DoT

But let's talk about something different but familiar at the same time

I want to get back to a much older but still not well understood issue in the DNS …

Dual Stack DNS

How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
    - Is there a "happy eyeballs" version of DNS server selection?
    - Or is there a reverse bias to use IPv4?
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you?
3. And what about DNSSEC?
    - How well does IPv6 support large UDP packets?

A word or two about "how" to talk about the DNS

We really don't understand what a "resolver" is!

- It could be a single platform running an instance of DNS resolver code
- It could be a collection of independent back-end systems with a load distributor front end facing clients
- It could be a hybrid collection where the back ends synchronise each other to emulate a common cache
- It is a stub, recursive, or forwarding resolver
- A resolver may have 1 client or millions of clients or anything in between

When we talk about "resolvers" its challenging to understand exactly what we are talking about!

Another word, this time about "how" to talk about DNS queries

We don’t understand what a query is!

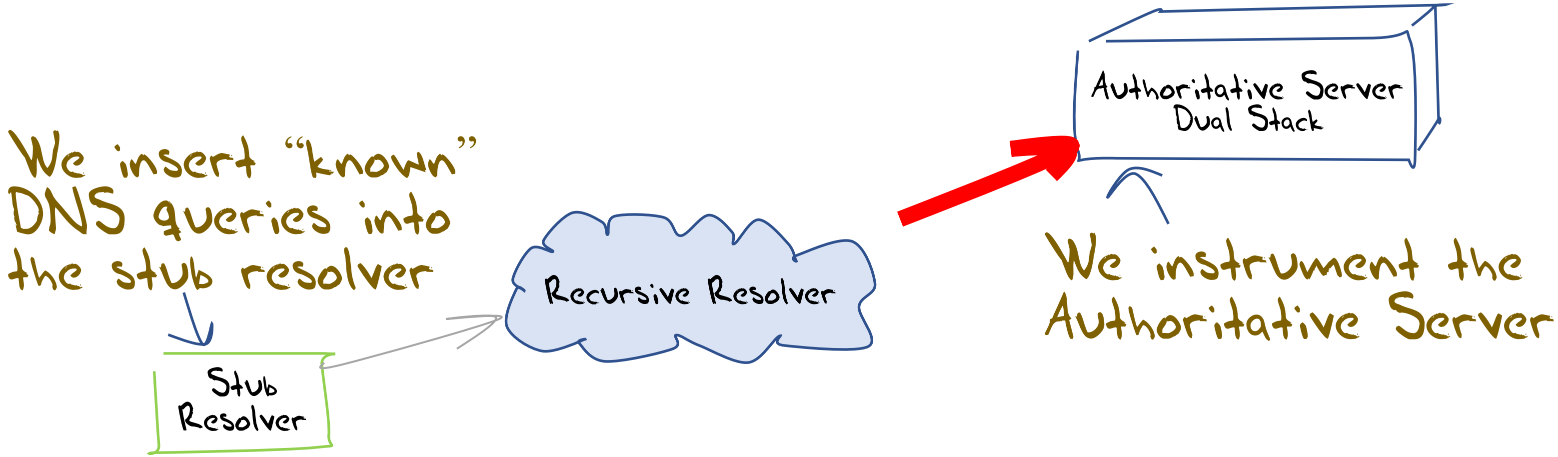
- Which sounds silly, but the distributed resolution process causes a ‘fan out’ of queries as part of the resolution process when a single query may cause a number of ‘discovery’ queries to establish the identify of the authoritative server(s) for the name
- Resolvers all use their own timers for retransmission
- Queries have no “hop count” or “resolver path” attached
  - there is no context to understand the reason for a query!
- Queries have a life of their own

APNIC's DNS Experimental Rig

We make parts of the name unique to each experiment

That way the recursive resolvers have no cached data and are forced to query the authoritative server

We "see" the recursive to authoritative query process by instrumenting the authoritative server, and match experiment placement records to the server's DNS logs



Dual Stack DNS

How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
   - Is there a “happy eyeballs” version of DNS server selection?
   - Or is there a reverse bias to use IPv4?
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you?
3. And what about DNSSEC?
   - How well does IPv6 support large UDP packets?

Dual Stack DNS

A “happy eyeballs*” DNS approach would be to prefer to use the IPv6 address of the authoritative server in preference to the IPv4 address

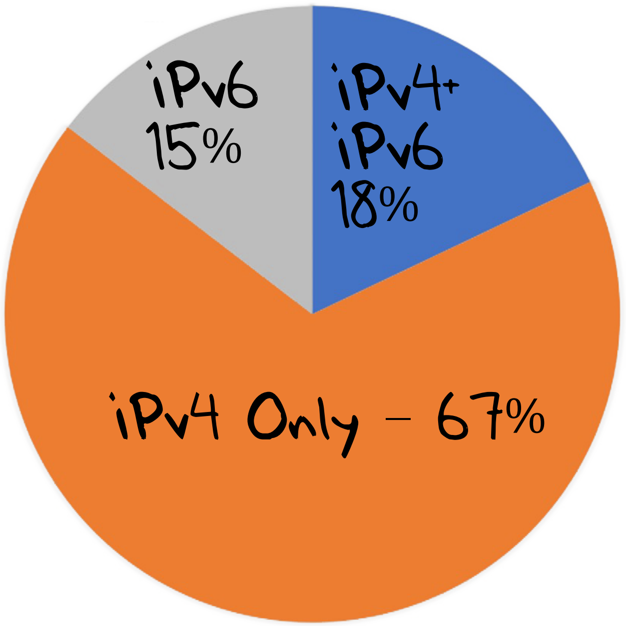
A “reverse bias” DNS approach would be to prefer to use the IPv4 address

* Why is happy eyeballs important?

Dual Stack DNS

A “happy eyeballs*” DNS approach would be to prefer to use the IPv6 address of the authoritative server in preference to the IPv4 address

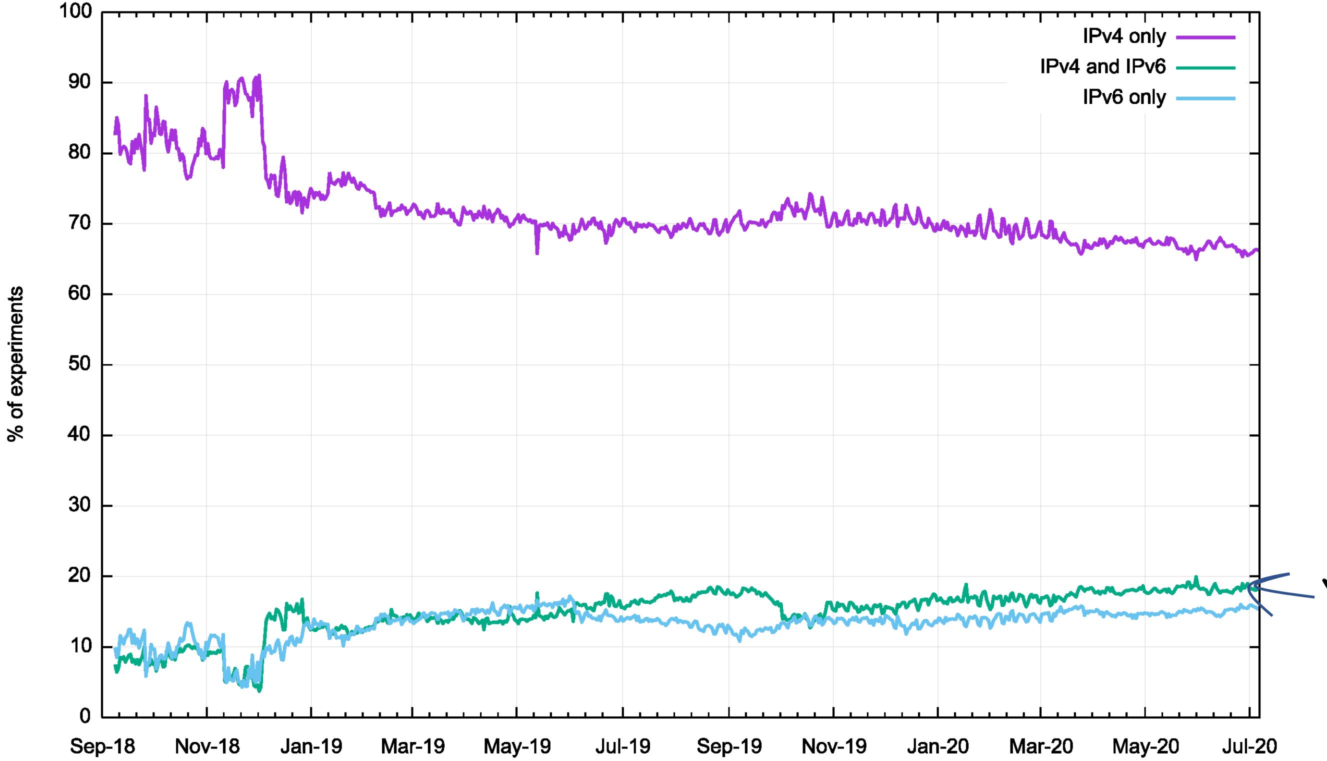
A “reverse bias” DNS approach would be to prefer to use the IPv4 address



* Why is happy eyeballs important?

Dual Stack DNS

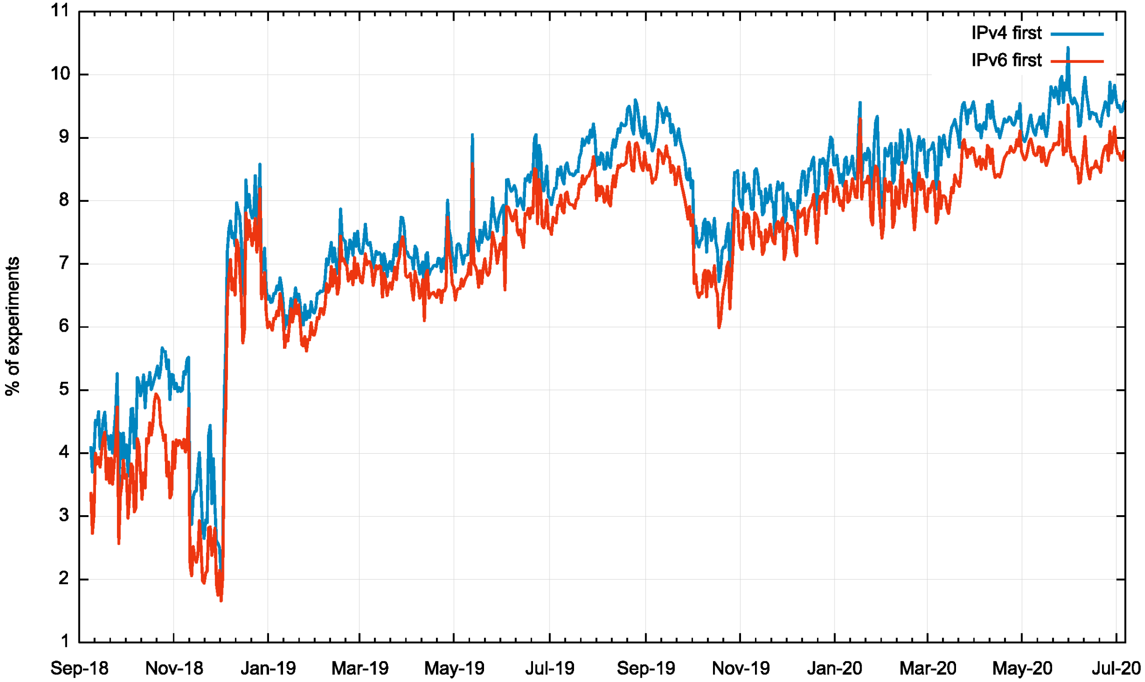
Time series of V4 / V6 query rates



What's going on here?

Dual Stack DNS

Time series of V4 / V6 query rates



Where we see both iPv4 and iPv6 used to query the authoritative server, there is a small bias to prefer to use iPv4 for the first query

Dual Stack DNS

How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
   - Is there a "happy eyeballs" version of DNS server selection? No!
   - Or is there a reverse bias to use IPv4? Yes!
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you?
3. And what about DNSSEC?
   - How well does IPv6 support large UDP packets?

Dual Stack DNS

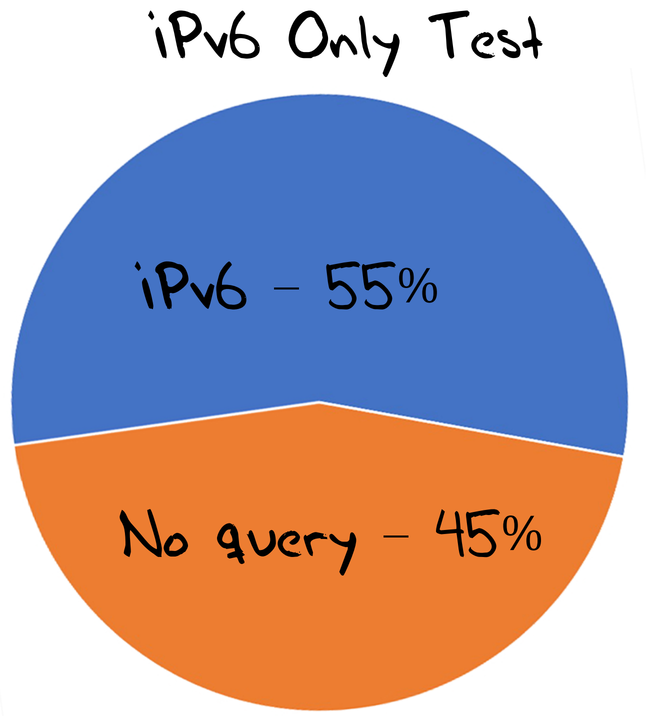
How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
   - Is there a "happy eyeballs" version of DNS server selection? No!
   - Or is there a reverse bias to use IPv4? Yes!
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you?
3. And what about DNSSEC?
   - How well does IPv6 support large UDP packets?

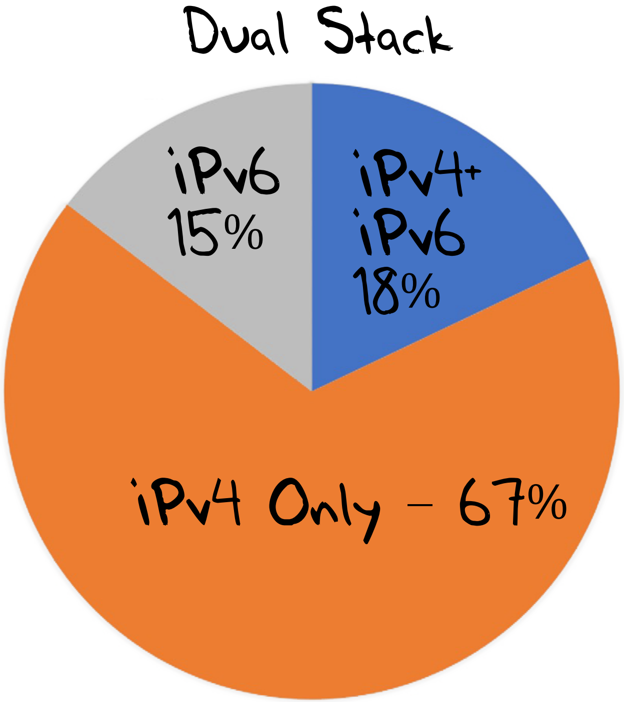
Dual Stack vs IPv6 only DNS

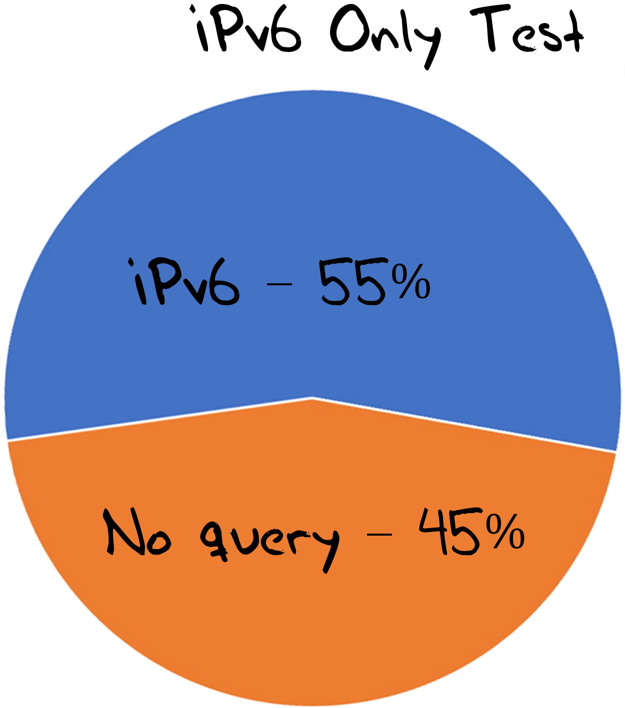
In this case the authoritative name server only has an IPv6 address

Of all the clients that are presented with an experiment (41M over 5 days) 55% of names are seen asking for the experiment name over IPv6



Dual Stack vs IPv6 only DNS





Dual Stack DNS

How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
   - Is there a "happy eyeballs" version of DNS server selection? No!
   - Or is there a reverse bias to use IPv4? Yes!
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you? 55%
3. And what about DNSSEC?
   - How well does IPv6 support large UDP packets?

Dual Stack DNS

How well is IPv6 supported in the DNS?

1. How does the DNS handle dual-stacked authoritative servers?
   - Is there a "happy eyeballs" version of DNS server selection? No!
   - Or is there a reverse bias to use IPv4? Yes!
2. If you placed authoritative servers on an IPv6-only service how many users would be able to reach you? 55%
3. And what about DNSSEC?
   - How well does IPv6 support large UDP packets?

IPv6 and Packet Fragmentation

IPv6 made two major changes to IP's handling of packet fragmentation:

- The fragmentation control header has been moved out of the IP header to become an **extension header**
  - In other words the UDP / TCP protocol header is pushed further into the packet and to find it you need to follow the header chain
- The IPv4 'Don't Fragment' bit is jammed **on** in IPv6
  - In the case of path MTU issues IPv6 routers should not perform fragmentation on the fly, but are required to pass an ICMPv6 PTB message back to the packet's sender

Who uses Fragmentation anyway?

- Well, the DNS is a good place to start looking!

Who uses large DNS packets anyway?

- Well, t

```
$ dig +dnssec DNSKEY org
; <<>> DiG 9.8.3-P1 <<>> +dnssec DNSKEY org
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 21353
;; flags: qr rd ra ad; QUERY: 1, ANSWER: 7, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 512;
; QUESTION SECTION:
;org.          IN         DNSKEY

;; ANSWER SECTION:
org.          861        IN         DNSKEY     256 3 7 AwEAAXxsMmN/JgpEE9Y4uFNRJm7Q9GBwmEYUCsCxuKlg
org.          861        IN         DNSKEY     256 3 7 AwEAAayiVbuM+ehlsKsuAL1CI3mA+5JM7ti3VeY8ysmo
org.          861        IN         DNSKEY     257 3 7 AwEAAZTjbIO5kIpxWUtyXc8avsKyHIIZ+LjC2Dv8naO+
org.          861        IN         DNSKEY     257 3 7 AwEAAcMnWBKLuvG/LwnPVykcmpvnntwxfshHlHRhlY0F

org.          861        IN         RRSIG      DNSKEY 7 1 900 20170815152632 20170725142632 3947
org.          861        IN         RRSIG      DNSKEY 7 1 900 20170815152632 20170725142632 9795
org.          861        IN         RRSIG      DNSKEY 7 1 900 20170815152632 20170725142632 17883

;; Query time: 134 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Mon Jul 31 12:07:1? 2017
;; MSG SIZE  rcvd: 1625
```

The response to a DNSKEY query for .org used a response of 1,625 octets!

Who uses lar… ts anyway?

That was back in 2017 – what about today?

```
$ dig +dnssec DNSKEY org

; <<>> DiG 9.16.2 <<>> +dnssec DNSKEY org
;; global options: +cmd
;; Got answer:

…

;; WHEN: Sat May … 22:02 AEST 2020
;; MSG SIZE  rcvd: 1058
```

```
…Q9GBwmEYUCsCxuKlg
…AL1CI3mA+5JM7ti3VeY8ysmo
…kIpxWUtyXc8avsKyHIIZ+LjC2Dv8naO+
…AAcMnWBKLuvG/LwnPVykcmpvnntwxfshHlHRhlY0F

                 DNSKEY 7 1 900 20170815152632 20170725142632 3947
        RRSIG    DNSKEY 7 1 900 20170815152632 20170725142632 9795
        RRSIG    DNSKEY 7 1 900 20170815152632 20170725142632 17883

…time: 134 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Mon Jul 31 12:07:… 2017
;; MSG SIZE  rcvd: 1625
```

The response to a DNSKEY query for .org used a response of 1,625 octets!

Who uses lar… ts anyway?

That was back in 2017 – what about today?

.org may have dropped its response size for DNSKEY queries

but there are still very large entries for other TLDs

```
;; WHEN: Sat May … 22:02 AEST 2020
;; MSG SIZE  rcvd: 1058
```

```
…SuAL1CI3mA+5JM7ti3VeY8ysmo
…IpxWUtyXc8avsKyHIIZ+LjC2Dv8naO+
…AAcMnWBKLuvG/LwnPVykcmpvnntwxfshHlHRhlY0F

…           DNSKEY 7 1 900 20170815152632 20170725142632 3947
RRSIG       DNSKEY 7 1 900 20170815152632 20170725142632 9795
RRSIG       DNSKEY 7 1 900 20170815152632 20170725142632 17883

;; …time: 134 msec
;; SERVER: 8.8.8.8#53(8.8.8.8)
;; WHEN: Mon Jul 31 12:07:… 2017
;; MSG SIZE  rcvd: 1625
```

The response to a DNSKEY query for .org used a response of 1,625 octets!

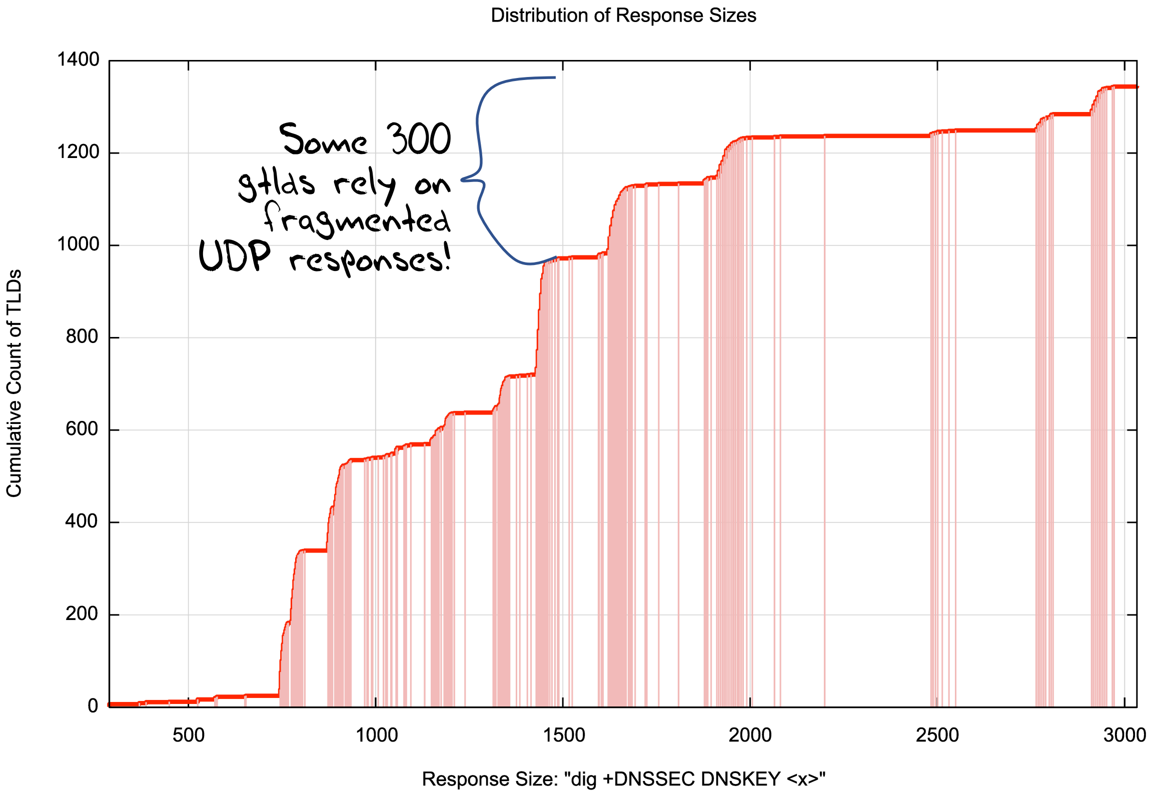
Who uses large DNS packets anyway?

These folk do! →

```
.booking 2932
.winners 2932
.watches 2932
.ferrero 2932
.lincoln 2932
.chintai 2932
.citadel 2932
.oldnavy 2932
.banamex 2932
.farmers 2932
.athleta 2932
.jpmorgan 2937
.discover 2937
.homegoods 2942
.marshalls 2942
.analytics 2942
.homesense 2942
.statefarm 2942
.swiftcover 2947
.xn--kpu716f 2952
.weatherchannel 2967
.bananarepublic 2967
.americanexpress 2972
.gdn 3033
```

Size of dnssec-signed DNSKEY response for some gtlds

Who uses large DNS packets anyway?



However…

UDP Fragmentation has its problems

- UDP trailing fragments in IPv4 and IPv6 may encounter fragment filtering rules on firewalls in front of resolvers

- Large UDP packets in IPv6 may encounter path MTU mismatch problems, and the ICMP6 Packet Too Big diagnostic message may be filtered.

  Even if it is delivered, the host may not process the message due to the lack of verification of the authenticity of the ICMP6 message.

  Because the protocol is UDP, receipt of an ICMP6 message will not cause retransmission of a re-framed packet.

- UDP fragments in IPv6 are implemented by Extension Headers. There is some evidence of deployment of IPv6 switching equipment that unilaterally discards IPv6 packets with extension headers

Is this a problem for today's IPv6 Internet?

- Can we measure the extent to which users might be affected with this scenario of large DNS responses, DNS resolvers and IPv6?

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

V6, the DNS and Fragmented UDP

Total number of tests (DNS over UDP over IPv6): 27,619,047

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

IPv6 Fragmentation Failure Rate: **41%**

Data gathered 26 April – 1 May 2020

V6, the DNS and Fragmented UDP

Total number of tests (DNS over UDP over IPv6): 27,619,047

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

IPv6 Fragmentation Failure Rate: **41%**

That's awesomely bad!

Data gathered 26 April – 1 May 2020

What to do?

Accepting a future IPv6-only Internet means we are going to have to take the problem of IPv6 Fragmentation seriously

- Because relying on IPv4 as a backup is a hack with an indeterminate future!

Which means that we need to figure out how to change the appalling drop rate for fragmented IPv6 packets both in the DNS and in end-to-end paths in the net

Should we try and fix the network problem or try to work around it?

What do the RFC's say?

What do the RFC's say?

Internet Engineering Task Force (IETF)
Request for Comments: 8085
BCP: 145
Obsoletes: 5405
Category: Best Current Practice
ISSN: 2070-1721

L. Eggert
NetApp
G. Fairhurst
University of Aberdeen
G. Shepherd
Cisco Systems
March 2017

UDP Usage Guidelines

Abstract

The User Datagram Protocol (UDP) provides a minimal message-passing transport that has no inherent congestion control mechanisms. This document provides guidelines on the use of UDP for the designers of applications, tunnels, and other protocols that use UDP. Congestion control guidelines are a primary focus, but the document also provides guidance on other topics, including message sizes, reliability, checksums, middlebox traversal, the use of Explicit Congestion Notification (ECN), Differentiated Services Code Points

What do the RFC's say?

Internet Engineering Task Force (IETF) L. Eggert
Request for Comments: 8085 NetApp
BCP: 145
Obsoletes
Category:
ISSN: 207

Abstract

The Us
transp
docume
applic
contro
provid
reliab
Conges

Due to these issues, an application SHOULD NOT send UDP datagrams that result in IP packets that exceed the Maximum Transmission Unit (MTU) along the path to the destination. Consequently, an application SHOULD either use the path MTU information provided by the IP layer or implement Path MTU Discovery (PMTUD) itself [RFC1191]

endpoint.

Applications that do not follow the recommendation to do PMTU/PLPMTUD discovery SHOULD still avoid sending UDP datagrams that would result in IP packets that exceed the path MTU. Because the actual path MTU is unknown, such applications SHOULD fall back to sending messages that are shorter than the default effective MTU for sending (EMTU_S in [RFC1122]). For IPv4, EMTU_S is the smaller of 576 bytes and the first-hop MTU [RFC1122]. For IPv6, EMTU_S is 1280 bytes [RFC2460].

What do the RFC's say?

Internet Engineering Task Force (IETF) L. Eggert
Request for Comments: 8085
BCP: 145
Obsoletes
Category:
ISSN: 207

Due to these issues, an application SHOULD NOT send UDP datagrams that result in IP packets that exceed the Maximum Transmission Unit (MTU) along the path to the destination. Consequently, an application SHOULD either use the path

Applications that ... PMTUD discovery ... result in ... MTU is ... s tha ... S in ... bytes and the fir... 1280 bytes [RFC2460].

DON'T FRAGMENT!

This BCP is saying that using EDNS(0) in the DNS to signal the capability of accepting large fragmented DNS responses was unwise, and if a host/application does not know the path MTU, it should truncate at UDP at 1280 octets (and IPv4 should truncate at 512 octets!)

What can we do about it?

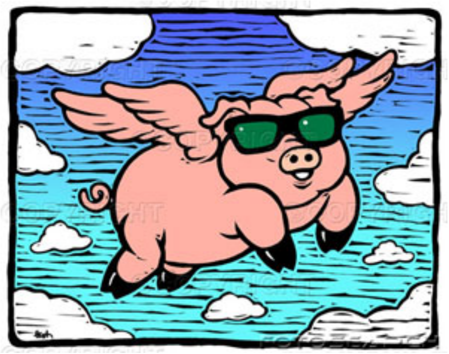
Fix it!

Get all the deployed routers, switches and firewalls and related network middleware to accept packets with IPv6 Fragmentation Headers

What can we do about it?

Fix it!

Get all the deployed routers, switches and firewalls and related network middleware to accept packets with IPv6 Fragmentation Headers



What can we do about it?

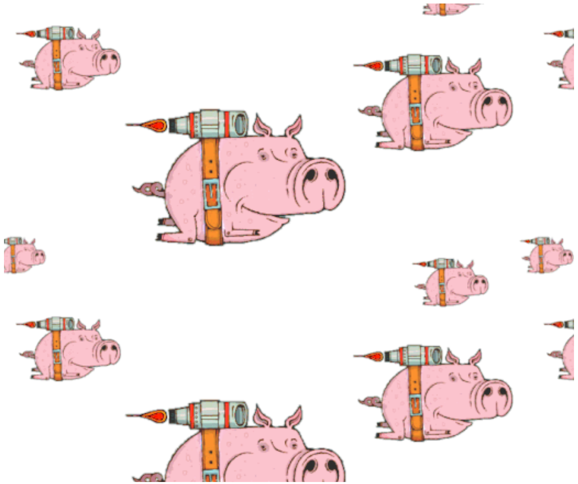
Change it!

Change the way in which IPv6 manages IP fragmentation and the use of Extension Headers as Fragmentation Control fields

What can we do about it?

Change it!

Change the way in which IPv6 manages IP fragmentation and the use of Extension Headers as Fragmentation Control fields



What can we do about it?

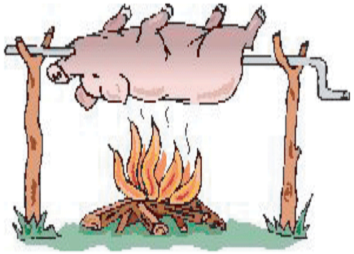
Avoid it!

Change application behaviour to avoid the use of packet fragmentation completely

What can we do about it?

Avoid it!

Change application behaviour to avoid the use of packet fragmentation completely



Which?

All of these options have a certain level of pain, cost and potential inconvenience

Its hard to work out what is the best course of action, but it seems like a lot of extra effort if we take on all three at once!

Large DNS Responses and IPv6

Change the transport protocol?

- DNS over TCP by default
- DNS over TLS over TCP by default
- DNS over HTTP over TLS over TCP
- DNS over QUIC
- Devise some new DNS framing protocol that uses multiple packets with firewall-friendly packet and protocol headers instead of IP fragmentation

Large DNS Responses and IPv6

Change the application protocol behaviour?

- Perform UDP MTU discovery using EDNS(0) UDP Buffer Size variations as a probe
- Shift Additional Records into additional explicit UDP query/response transactions rather than bloating the original DNS response
- Add a truncated minimal UDP response to trail a fragmented response (ATR)

Where now?

- We have a decent idea of the problem space we need to resolve
- We'd prefer a plan that allows each of us to work independently rather than a large scale orchestrated common change
- We're not sure we can clean up all the ICMPv6 filters and EH packet droppers in the IPv6 network
- And it sure seems a bit late in the day to contemplate IPv6 protocol changes
- Which means that we are probably looking at working around the problem by changing the behaviour of DNS and use an upper payload size of 1232 octets

See https://dnsflagday.net/2020/

Thanks!

Additional Material

Top 20 Tables

Which Resolvers?

- 73,354 IPv6 resolver /48 subnets seen
- 12,287 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 can't

Resolver	Hits	ASN	AS Name
2001:12e0:0800::	138,036	10429	TELEFONICA BRASIL S.A, BR
2800:0680:0012::	71,239	3816	COLOMBIA TELECOMUNICACIONES S.A. ESP, CO
2804:0d40:0080::	51,254	7738	Telemar Norte Leste S.A., BR
2800:0190:0005::	46,883	19037	AMX Argentina S.A., AR
2800:0190:0004::	42,801	19037	AMX Argentina S.A., AR
2803:7180:2220::	17,408	21575	ENTEL PERU S.A., PE
2803:7180:2222::	16,667	21575	ENTEL PERU S.A., PE
2402:8780:0000::	15,400	63859	MYREPUBLIC-AS-ID PT. Eka Mas Republik, ID
2806:0260:1004::	15,294	13999	Mega Cable, S.A. de C.V., MX
2001:12e0:0802::	15,196	10429	TELEFONICA BRASIL S.A, BR
2806:0260:1008::	13,893	13999	Mega Cable, S.A. de C.V., MX
2a00:1028:0001::	10,735	5610	O2-CZECH-REPUBLIC, CZ
2804:0d40:0081::	9,962	7738	Telemar Norte Leste S.A., BR
2a00:1028:0004::	6,295	5610	O2-CZECH-REPUBLIC, CZ
2001:1284:ff02::	5,831	14868	COPEL Telecomunicacoes S.A., BR
2804:0d50:0080::	5,821	8167	Brasil Telecom S/A - Filial Distrito Federal, BR
2001:4860:4801::	4,767	15169	GOOGLE, US
2001:0df4:5f00::	3,190	134033	HIREACH-BROADBAND-AS HIREACH BROADBAND PRIVATE LTD, IN
2402:3a80:c029::	3,038	38266	VODAFONE-IN Vodafone India Ltd., IN
2402:3a80:c035::	2,722	38266	VODAFONE-IN Vodafone India Ltd., IN

All these resolvers appear to be unable to receive fragmented UDP

This is the top 20, as measured by the number of tests that used this resolver subnet

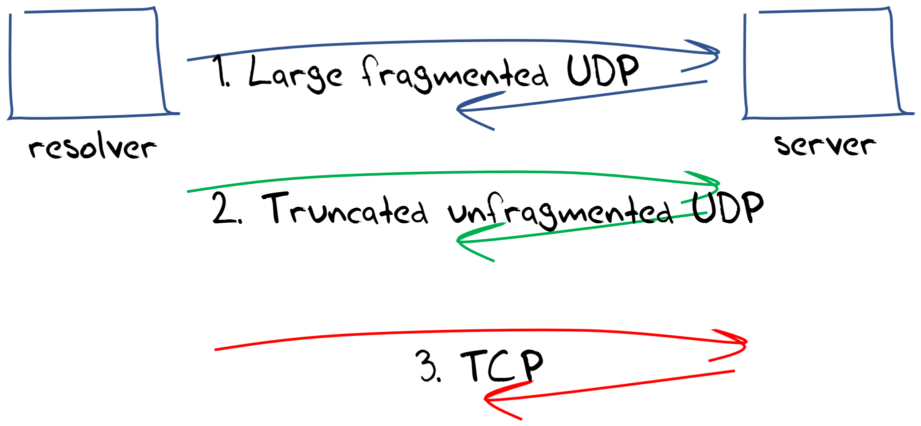
Which Network's Resolvers can't

AS	Tests	AS Tests		AS Name
10429	153,235	165,213		TELEFONICA BRASIL S.A, BR
19037	89,684	89,684		AMX Argentina S.A., AR
3816	71,239	71,239		COLOMBIA TELECOMUNICACIONES S.A. ESP, CO
7738	61,216	578,551		Telemar Norte Leste S.A., BR
21575	34,075	34,075		ENTEL PERU S.A., PE
30873	31,908	38,641		PTC-YEMENNET, YE
4837	30,739	181,862		CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
13999	29,187	128,896		Mega Cable, S.A. de C.V., MX
5610	17,036	17,080		O2-CZECH-REPUBLIC, CZ
63859	15,400	15,400		MYREPUBLIC-AS-ID PT. Eka Mas Republik, ID
38091	13,128	108,205		HELLONET-AS-KR LG HelloVision Corp., KR
38266	9,586	38,872		VODAFONE-IN Vodafone India Ltd., IN
263656	9,542	9,777		BRSULNET TELECOM LTDA, BR
14868	5,831	6,116		COPEL Telecomunicacoes S.A., BR
8167	5,821	87,254		Brasil Telecom S/A - Filial Distrito Federal, BR
15169	4,767	7,389,028		GOOGLE, US
134033	4,601	4,601		HIREACH-BROADBAND-AS HIREACH BROADBAND PRIVATE LTD, IN
6621	3,768	10,730		HNS-DIRECPC, US
6939	2,669	14,147		HURRICANE, US
262459	1,549	1,550		Osirnet Info Telecom Ltda., BR

All these networks host resolvers that appear to be unable to receive fragmented UDP
This is the Top 20, as measured by the query count per origin AS of the resolver address

Its not quite so simple as "can't"

When a resolver fails to receive a UDP response it _may_ re-query with a smaller UDP buffer size parameter
This will cause our server to respond with a truncated DNS response over UDP
This _should_ cause the resolver to re-query using TCP



It's not quite so simple as "can't"

When a resolver fails to receive a UDP response it _may_ re-query with a smaller UDP buffer size parameter
This will cause our server to respond with a truncated DNS response over UDP
This _should_ cause the resolver to re-query using TCP

There are four cases of fragmentation failure :

1. can’t receive fragmented UDP and won’t use a smaller UDP Buffer Size
2. can’t receive fragmented UDP, will use a smaller UDP Buffer Size but cannot make a TCP connection
3. can’t receive fragmented UDP, will use a smaller UDP Buffer Size but cannot complete a TCP connection
4. can’t receive fragmented UDP, will use a smaller UDP Buffer Size and make a TCP connection (fail then recovery)

Case 1 - Which Resolvers really can't

Resolver		AS	Tests	AS Tests		AS Name
2001:12e0:0800::	3	10429	138,036	165,213		TELEFONICA BRASIL S.A, BR
2800:0680:0012::	3	3816	71,239	71,239		COLOMBIA TELECOMUNICACIONES S.A. ESP, CO
2804:0d40:0080::	3	7738	51,254	578,551		Telemar Norte Leste S.A., BR
2800:0190:0005::	3	19037	46,883	89,684		AMX Argentina S.A., AR
2800:0190:0004::	3	19037	42,801	89,684		AMX Argentina S.A., AR
2803:7180:2220::	3	21575	17,408	34,075		ENTEL PERU S.A., PE
2803:7180:2222::	3	21575	16,667	34,075		ENTEL PERU S.A., PE
2402:8780:0000::	3	63859	15,400	15,400		MYREPUBLIC-AS-ID PT. Eka Mas Republik, ID
2806:0260:1004::	3	13999	15,294	128,896		Mega Cable, S.A. de C.V., MX
2001:12e0:0802::	3	10429	15,196	165,213		TELEFONICA BRASIL S.A, BR
2806:0260:1008::	3	13999	13,893	128,896		Mega Cable, S.A. de C.V., MX
2a00:1028:0001::	3	5610	10,735	17,080		O2-CZECH-REPUBLIC, CZ
2804:0d40:0081::	3	7738	9,962	578,551		Telemar Norte Leste S.A., BR
2a00:1028:0004::	3	5610	6,295	17,080		O2-CZECH-REPUBLIC, CZ
2001:1284:ff02::	3	14868	5,831	6,116		COPEL Telecomunicacoes S.A., BR
2804:0d50:0080::	3	8167	5,821	87,254		Brasil Telecom S/A - Filial Distrito Federal, BR
2001:4860:4801::	3	15169	4,767	7,389,028		GOOGLE, US
2001:0df4:5f00::	3	134033	3,190	4,601		HIREACH-BROADBAND-AS HIREACH BROADBAND PRIVATE LTD, IN
2001:0df4:5f01::	3	134033	1,411	4,601		HIREACH-BROADBAND-AS HIREACH BROADBAND PRIVATE LTD, IN
2a02:79e0:0020::	3	60032	673	1,317		CORIOLIS-AS, FR

These resolvers appear to be unable to receive fragmented UDP and will not re-query with a smaller UDP buffer size
This is the Top 20, as measured by the query count per resolver subnet address

Case 1 – Which Networks' Resolvers really can't

AS	Tests	AS Tests		AS Name
10429	153,235	165,213		TELEFONICA BRASIL S.A, BR
19037	89,684	89,684		AMX Argentina S.A., AR
3816	71,239	71,239		COLOMBIA TELECOMUNICACIONES S.A. ESP, CO
7738	61,216	578,551		Telemar Norte Leste S.A., BR
21575	34,075	34,075		ENTEL PERU S.A., PE
13999	29,187	128,896		Mega Cable, S.A. de C.V., MX
5610	17,036	17,080		O2-CZECH-REPUBLIC, CZ
63859	15,400	15,400		MYREPUBLIC-AS-ID PT. Eka Mas Republik, ID
14868	5,831	6,116		COPEL Telecomunicacoes S.A., BR
8167	5,821	87,254		Brasil Telecom S/A - Filial Distrito Federal, BR
15169	4,767	7,389,028		GOOGLE, US
134033	4,601	4,601		HIREACH-BROADBAND-AS HIREACH BROADBAND PRIVATE LTD, IN
6621	3,768	10,730		HNS-DIRECPC, US
6939	2,666	14,147		HURRICANE, US
60032	1,317	1,317		CORIOLIS-AS, FR
0	1,229	151,916		-Reserved AS-, ZZ
34602	1,170	1,218		STARLINK-AS Moscow, Russia, RU
63949	869	31,213		LINODE-AP Linode, LLC, US
263047	766	766		Speednet Provedor de Acesso a Internet Ltda, BR
34524	615	665		DIGICOM-AS, BG

These networks host resolvers that appear to be unable to receive fragmented UDP and will not re-query with a smaller UDP buffer size

This is the Top 20, as measured by the query count per ASN

Case 2 - Which Resolvers can't TCP

2402:3a80:c029::	7	38266	3,038	38,872	VODAFONE-IN Vodafone India Ltd., IN
2402:3a80:c035::	7	38266	2,722	38,872	VODAFONE-IN Vodafone India Ltd., IN
2402:3a80:c026::	7	38266	2,040	38,872	VODAFONE-IN Vodafone India Ltd., IN
2402:3a80:c020::	7	38266	1,694	38,872	VODAFONE-IN Vodafone India Ltd., IN
2607:fb90:c12c::	7	21928	1,443	296,729	T-MOBILE-AS21928, US
2804:108c:cab0::	7	28146	1,396	1,396	MHNET TELECOM, BR
2001:4478:4010::	7	4802	988	4,396	ASN-IINET iiNet Limited, AU
2804:04b0:0101::	7	262459	818	1,550	Osirnet Info Telecom Ltda., BR
2804:04b0:0201::	7	262459	731	1,550	Osirnet Info Telecom Ltda., BR
2a04:d380:cc00::	7	41997	613	613	CONNECT-AS-1, AZ
2408:8663:1614::	7	4837	304	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8662:8001::	7	4837	286	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8662:1607::	7	4837	284	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8663:8001::	7	4837	282	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8663:1001::	7	4837	261	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8662:1639::	7	4837	254	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8663:1615::	7	4837	244	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2405:7b00:3100::	7	38091	241	108,205	HELLONET-AS-KR LG HelloVision Corp., KR
2408:8662:1615::	7	4837	215	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2408:8663:f000::	7	4837	214	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN

These resolvers appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size but not open a TCP session

This is the top 20, as measured by the query count per resolver subnet address

Case 2 – Which Networks' Resolvers can't TCP

AS	Tests	AS Tests	AS Name
4837	9,935	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
38266	9,586	38,872	VODAFONE-IN Vodafone India Ltd., IN
38091	2,377	108,205	HELLONET-AS-KR LG HelloVision Corp., KR
262459	1,549	1,550	Osirnet Info Telecom Ltda., BR
21928	1,443	296,729	T-MOBILE-AS21928, US
28146	1,396	1,396	MHNET TELECOM, BR
4802	988	4,396	ASN-IINET iiNet Limited, AU
41997	613	613	CONNECT-AS-1, AZ
2527	511	108,369	SO-NET So-net Entertainment Corporation, JP
262704	456	4,066	FD Informatica Ltda EPP, BR
28213	197	200	LCI Equipamentos de Informatica LTDA - LCI Telecom, BR
57096	192	192	GAT-AS, KG
263656	151	9,777	BRSULNET TELECOM LTDA, BR
262320	109	880	BANDA TURBO PROVEDORES DE INTERNET LTDA, BR
0	86	151,916	-Reserved AS-, ZZ
265187	71	6,729	STEEL WEB PROVEDORES DE ACESSO LTDA, BR
23910	57	2,191	CNGI-CERNET2-AS-AP China Next Generation Internet CERNET2, CN
1659	49	1,865	ERX-TANET-ASN1 Taiwan Academic Network (TANet) Information Center, TW
52951	44	2,561	Catanduva sistemas a cabo ltda., BR
262759	41	1,809	Vilson Giazzoni, BR

These networks host resolvers that appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size but not open a TCP session

This is the top 20, as measured by the query count per resolver subnet address

Case 3 – Which Resolvers can't complete TCP

Resolver		AS	Tests	AS Tests	AS Name
2405:7b00:2401::	15	38091	2,102	108,205	HELLONET-AS-KR LG HelloVision Corp., KR
2a02:2718:400a::	15	30873	2,056	38,641	PTC-YEMENNET, YE
2a02:2718:0000::	15	30873	1,858	38,641	PTC-YEMENNET, YE
2a02:2718:4001::	15	30873	1,834	38,641	PTC-YEMENNET, YE
2a02:2718:0004::	15	30873	1,580	38,641	PTC-YEMENNET, YE
2a02:2718:0002::	15	30873	1,560	38,641	PTC-YEMENNET, YE
2a02:2718:0001::	15	30873	1,493	38,641	PTC-YEMENNET, YE
2a02:2718:ffff::	15	30873	1,381	38,641	PTC-YEMENNET, YE
2a02:2718:400c::	15	30873	1,343	38,641	PTC-YEMENNET, YE
2408:8663:0180::	15	4837	1,163	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2a02:2718:8000::	15	30873	1,130	38,641	PTC-YEMENNET, YE
2408:8662:0180::	15	4837	1,126	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
2a02:2718:4010::	15	30873	1,089	38,641	PTC-YEMENNET, YE
2a02:2718:50da::	15	30873	992	38,641	PTC-YEMENNET, YE
2a02:2718:8003::	15	30873	807	38,641	PTC-YEMENNET, YE
2405:7b00:1607::	15	38091	805	108,205	HELLONET-AS-KR LG HelloVision Corp., KR
2a02:2718:0300::	15	30873	735	38,641	PTC-YEMENNET, YE
2a02:2718:d000::	15	30873	727	38,641	PTC-YEMENNET, YE
2a02:2718:0800::	15	30873	702	38,641	PTC-YEMENNET, YE
2408:8000:d501::	15	4837	700	181,862	CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN

These resolvers appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size, will open a TCP session, but appear to have a TCP MTU black hole issue
This is the top 20, as measured by the query count per resolver subnet address

Case 3 – Which Networks' Resolvers can't complete TCP

AS	Tests	AS Tests		AS Name
30873	31,638	38,641		PTC-YEMENNET, YE
4837	20,638	181,862		CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
38091	10,296	108,205		HELLONET-AS-KR LG HelloVision Corp., KR
263656	9,198	9,777		BRSULNET TELECOM LTDA, BR
45365	304	17,405		JBTV-AS-KR LG HelloVision Corp., KR
2527	135	108,369		SO-NET So-net Entertainment Corporation, JP
262494	75	14,786		Virtex Ltda, BR
17816	73	12,479		CHINA169-GZ China Unicom IP network China169 Guangdong province, CN
264586	70	1,423		SPEED TELECOMUNICACOES, BR
38019	61	61		CMNET-V4TIANJIN-AS-AP tianjin Mobile Communication Company Limited, CN
52846	50	2,595		INNOVANET Telecom LTDA., BR
265187	46	6,729		STEEL WEB PROVEDORES DE ACESSO LTDA, BR
0	44	151,916		-Reserved AS-, ZZ
23910	42	2,191		CNGI-CERNET2-AS-AP China Next Generation Internet CERNET2, CN
52739	37	1,441		PROVEDOR INTERSOUSA LTDA, BR
39280	34	1,393		ULTELNET-AS, AZ
47562	32	296		FASTLINK, RU
9808	29	48,553		CMNET-GD Guangdong Mobile Communication Co.Ltd., CN
47975	27	3,666		KT-AS-47975, AM
267114	26	680		LINK NET BANDA LARGA EIRELI - ME, BR

These networks host resolvers that appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size, will open a TCP session, but appear to have a TCP MTU black hole issue
This is the top 20, as measured by the query count per resolver subnet address

Case 4 – Which Resolvers can't Frag but can TCP

Resolver		AS	Tests	AS Tests		AS Name
2a00:1450:400a::	31	15169	1,087,935	7,389,028		GOOGLE, US
2404:6800:4003::	31	15169	510,013	7,389,028		GOOGLE, US
2800:03f0:4001::	31	15169	494,635	7,389,028		GOOGLE, US
2404:6800:4005::	31	15169	493,653	7,389,028		GOOGLE, US
2a00:1450:4001::	31	15169	485,880	7,389,028		GOOGLE, US
2a00:1450:400c::	31	15169	448,655	7,389,028		GOOGLE, US
2404:6800:4008::	31	15169	378,751	7,389,028		GOOGLE, US
2a00:1450:4025::	31	15169	326,480	7,389,028		GOOGLE, US
2a00:1450:4010::	31	15169	315,765	7,389,028		GOOGLE, US
2402:0800:20ff::	31	7552	298,418	298,419		VIETEL-AS-AP Viettel Group, VN
2607:f8b0:400c::	31	15169	286,540	7,389,028		GOOGLE, US
2001:1890:01ff::	31	7018	276,859	276,905		ATT-INTERNET4, US
2405:0200:1613::	31	55836	177,707	2,770,023		RELIANCEJIO-IN Reliance Jio Infocomm Limited, IN
2607:f8b0:4003::	31	15169	165,860	7,389,028		GOOGLE, US
2407:0000:0000::	31	4761	160,579	160,579		INDOSAT-INP-AP INDOSAT Internet Network Provider, ID
2607:f8b0:4004::	31	15169	145,827	7,389,028		GOOGLE, US
2405:0200:1638::	31	55836	134,246	2,770,023		RELIANCEJIO-IN Reliance Jio Infocomm Limited, IN
2405:0200:160a::	31	55836	123,872	2,770,023		RELIANCEJIO-IN Reliance Jio Infocomm Limited, IN
2001:4288:0201::	31	6713	117,358	154,970		IAM-AS, MA
2405:0200:1637::	31	55836	115,742	2,770,023		RELIANCEJIO-IN Reliance Jio Infocomm Limited, IN

These resolvers appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size, get the truncated UDP response and successfully complete the TCP session

This is the top 20, as measured by the query count per resolver subnet address

Case 4 – Which Networks' Resolvers can't Frag but can TCP

AS	Tests	AS Tests		AS Name
15169	5,285,016	7,389,028		GOOGLE, US
55836	2,755,672	2,770,023		RELIANCEJIO-IN Reliance Jio Infocomm Limited, IN
7922	725,983	733,871		COMCAST-7922, US
7552	298,418	298,419		VIETEL-AS-AP Viettel Group, VN
7018	276,863	276,905		ATT-INTERNET4, US
28573	260,478	319,671		CLARO S.A., BR
22394	257,440	257,442		CELLCO, US
20057	222,069	222,070		ATT-MOBILITY-LLC-AS20057, US
21928	214,833	296,729		T-MOBILE-AS21928, US
3462	204,971	205,022		HINET Data Communication Business Group, TW
4134	173,831	233,976		CHINANET-BACKBONE No.31,Jin-rong Street, CN
4761	160,579	160,579		INDOSAT-INP-AP INDOSAT Internet Network Provider, ID
3352	157,695	157,695		TELEFONICA_DE_ESPANA, ES
6713	154,970	154,970		IAM-AS, MA
14080	133,968	133,968		Telmex Colombia S.A., CO
8708	132,146	132,147		RCS-RDS 73-75 Dr. Staicovici, RO
4837	131,977	181,862		CHINA169-BACKBONE CHINA UNICOM China169 Backbone, CN
10507	118,654	118,662		SPCS, US
15557	109,150	109,169		LDCOMNET, FR
2527	100,682	108,369		SO-NET So-net Entertainment Corporation, JP

These networks host resolvers that appear to be unable to receive fragmented UDP, will re-query with a smaller UDP buffer size, get the truncated UDP response and successfully complete the TCP session
This is the top 20, as measured by the query count per resolver subnet address