DNS Survival Guide

Artyom Gavrichenkov <ag@qrator.net>

A bit of a history: **DNS**

1983:

```
(int32)*host_str;
```

A bit of a history: **DNS**

1983:

```
(int32)*host_str;
```

1997-2017:

- load balancing
- geobalancing
- ASN policies

A bit of a history: **DNS**

1983:

```
(int32)*host_str;
```

1997-2017:

- load balancing
- geobalancing
- ASN policies
- failover
- EDNS0

A bit of a history: **DNS**

1983:

```
(int32)*host_str;
```

1997-2017:

- load balancing
- geobalancing
- ASN policies
- failover
- EDNS0
- AAAA
- DNSSEC
- DANE, CAA, …

Problem statement

How should an Internet company maintain its DNS infrastructure?

- In-house?
- Outsourcing?

Problem statement

How should an Internet company maintain its DNS infrastructure?

- In-house
  - How to choose a software product?
- Outsourcing
  - How to choose a service provider?

1. How to choose a software product?

Naïve approach:

a) It must be scalable
b) It should support *features*

DNS benchmarks, 2013

- Knot (1.2.0 & 1.3.0-RC5)
- Yadifa (1.0.2)
- NSD3 (3.2.15)
- NSD4 (4.0.0b4)
- PowerDNS (3.3)
- TinyDNS (1.05)
- Unbound (1.4.16)
- Pdnsd (1.2.8)

- Server:
  Dual Xeon E5-2670
  32Gb RAM DDR3 1333Mhz
  Intel X520-DA2 10Gbit
- Generator:
  Single Xeon E5-2670
  32Gb RAM DDR3 1333Mhz
  Intel X520-DA2 10Gbit
- Gentoo Linux 3.7.9

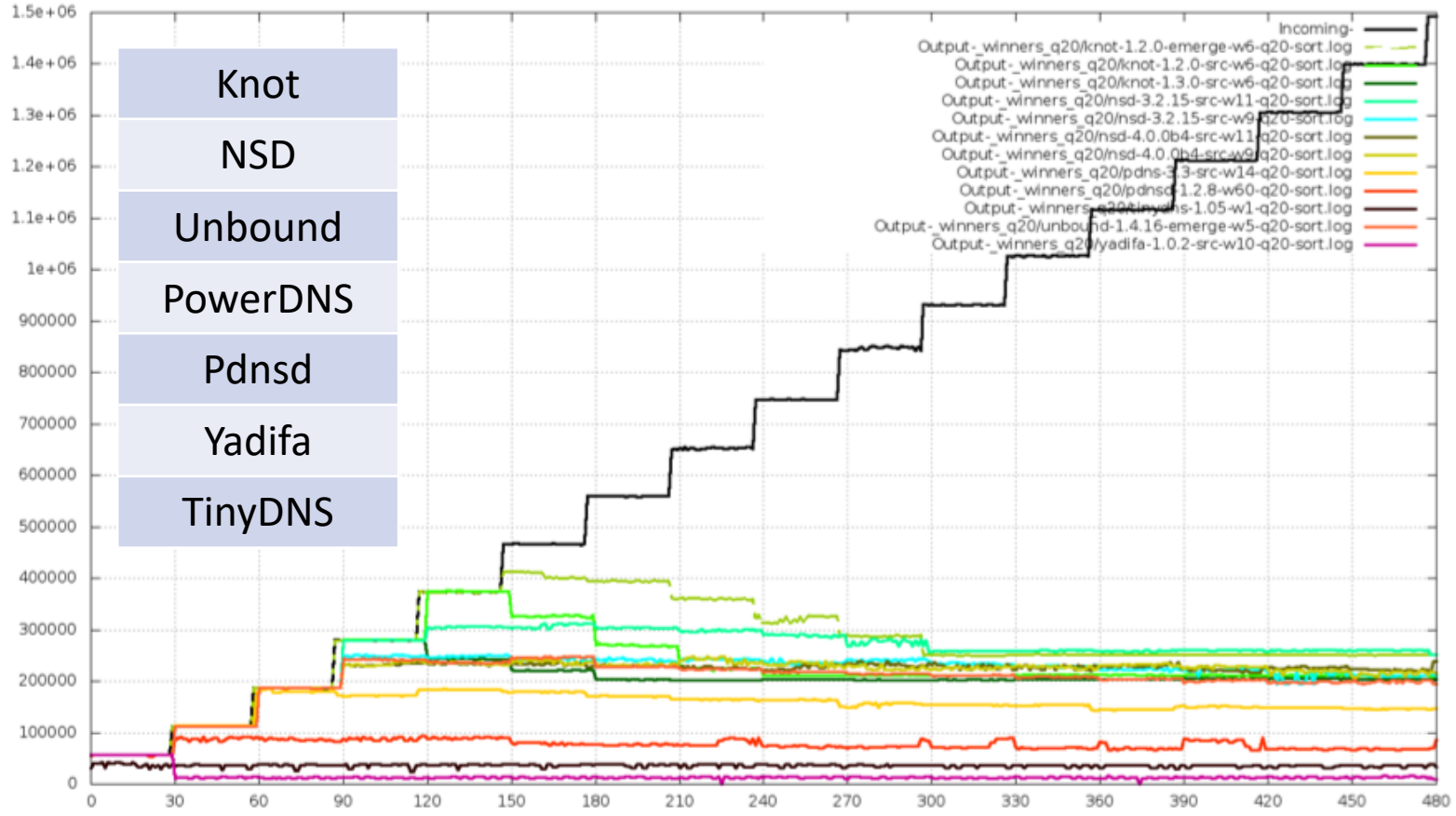
DNS benchmarks, 2013. Setup

- Vanilla DNS software!
- Purpose:
  purely academic (who runs better codebase)
- Authoritative:
  300 zones
- Caching:
  Same amount of data in cache

DNS benchmarks, 2013.

https://www.slideshare.net/ximaera/dns-server-benchmarking

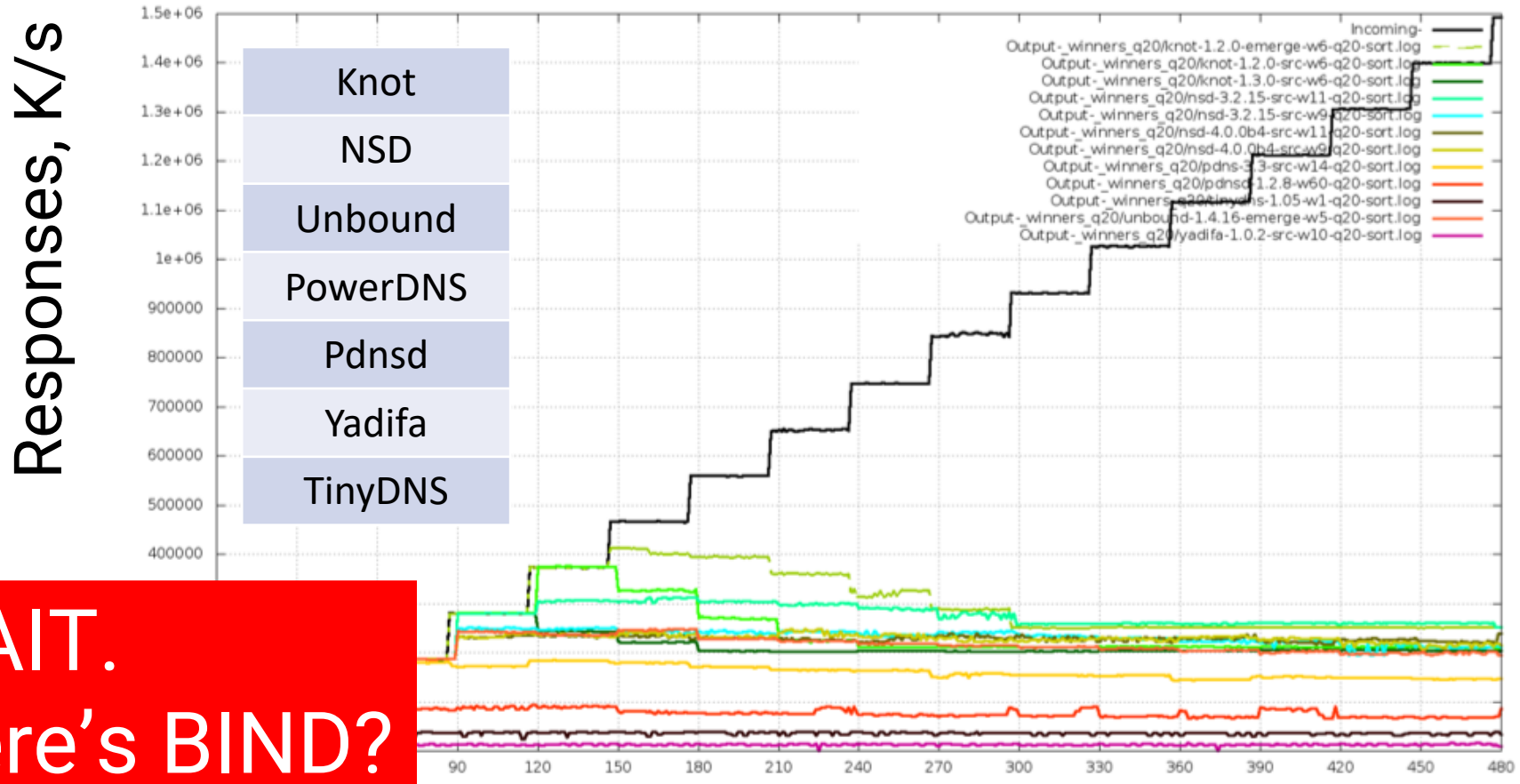
Responses, K/s



Queries, K/s

DNS benchmarks, 2013.

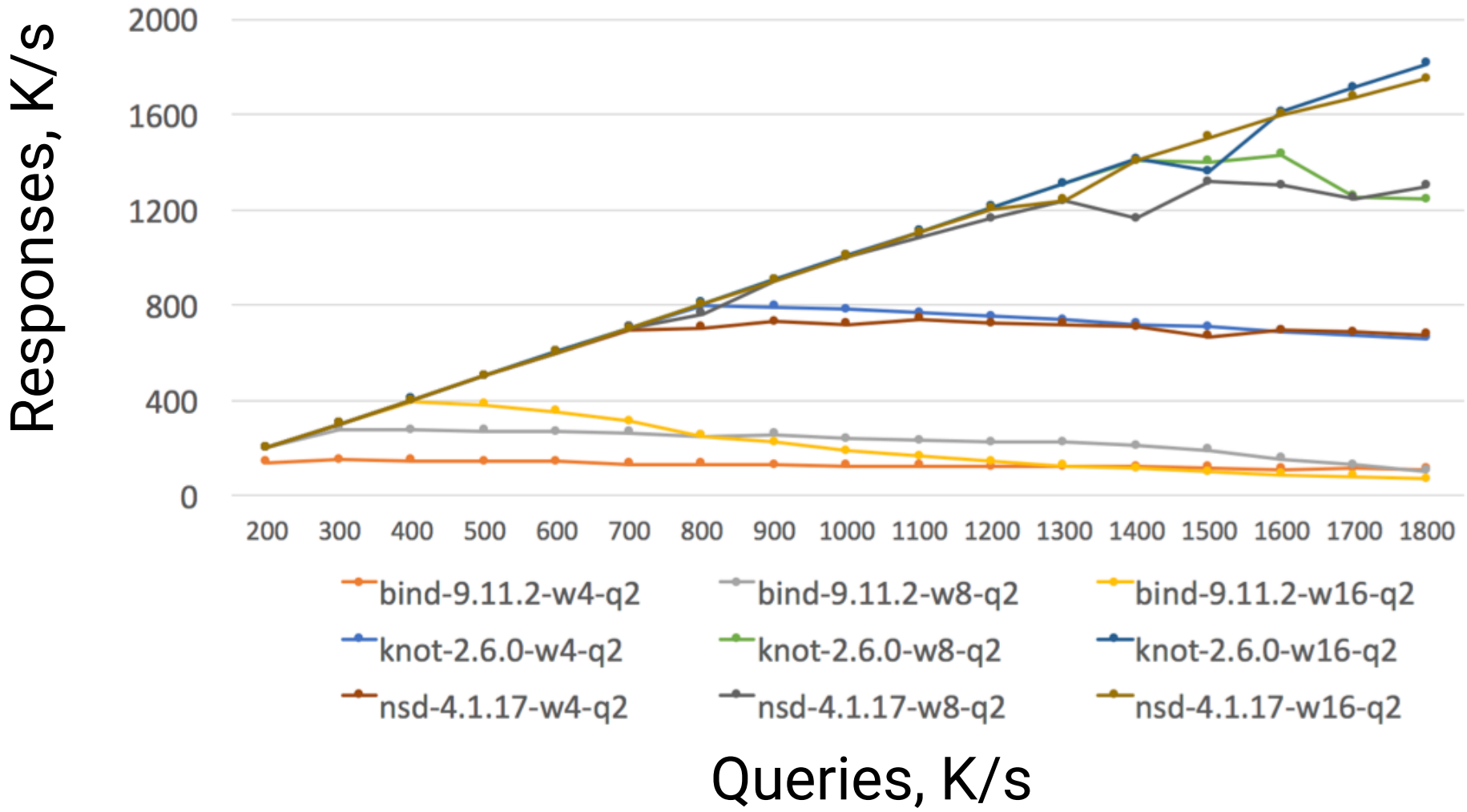
https://www.slideshare.net/ximaera/dns-server-benchmarking



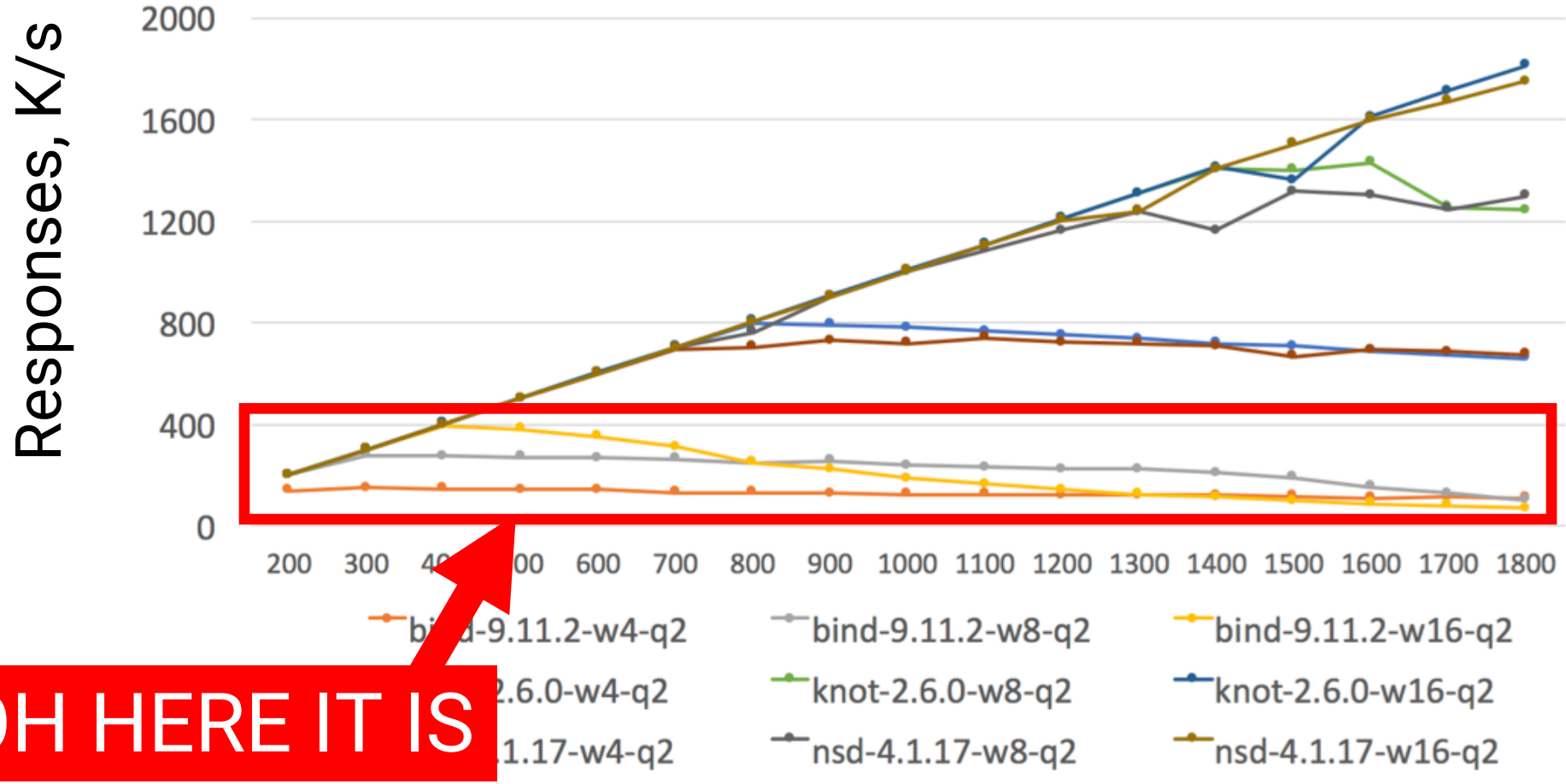
…WAIT.
Where's BIND?

Queries, K/s

DNS benchmarks, 2017

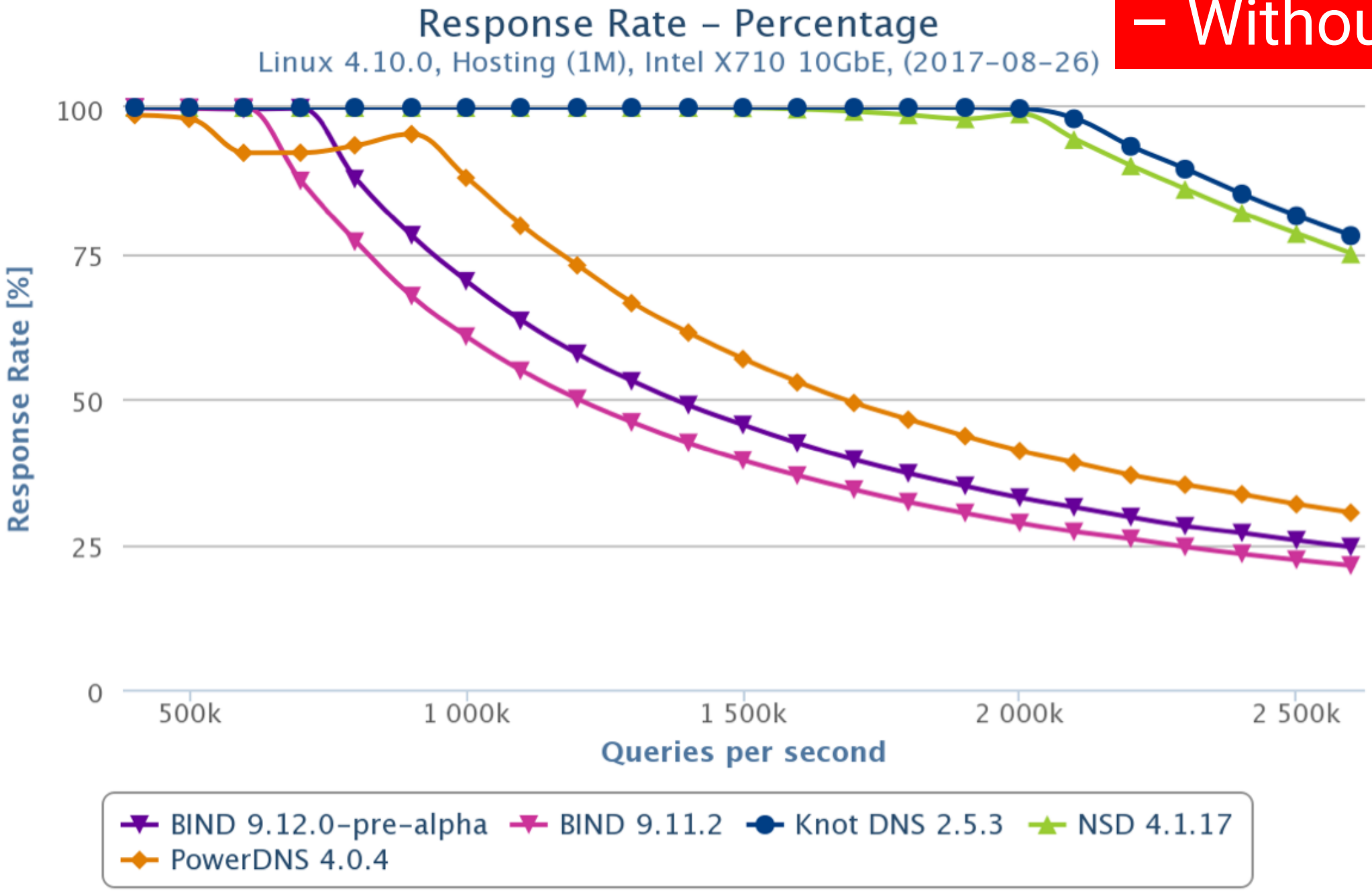


DNS benchmarks, 2017

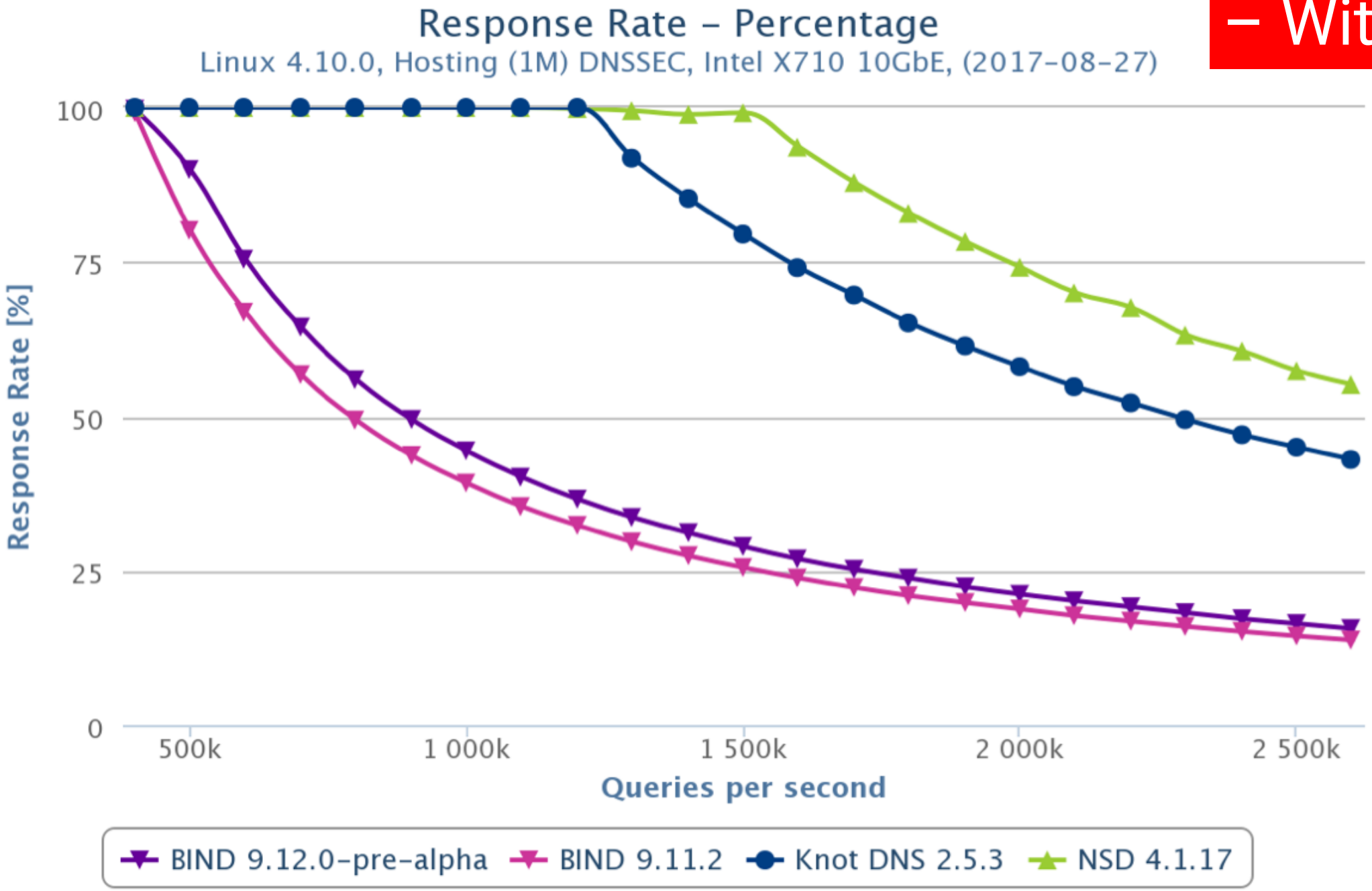


OH HERE IT IS

– Without DNSSEC



– With DNSSEC


Response Rate – Percentage
Linux 4.10.0, Hosting (1M) DNSSEC, Intel X710 10GbE, (2017-08-27)
Response Rate [%]
100
75
50
25
0
500k
1 000k
1 500k
2 000k
2 500k
Queries per second
BIND 9.12.0-pre-alpha
BIND 9.11.2
Knot DNS 2.5.3
NSD 4.1.17

This is not good.

The de-facto standard software doesn’t scale well

This is not good.

The de-facto standard software doesn’t scale well

- Yes, a balancer (Nginx) with a soccer field full of BIND servers will do.
- Definite overkill for a small task

This is not good.

The de-facto standard software doesn’t scale well

What scales well causes concern in other areas

- Maintainability?
- Reliability?
- Support?
- Backward compatibility?
- Patches and security?
- Features?

Back to the requirements.

Naïve approach:

a) It must be scalable – **how scalable?**
b) It should support *features* – **what features do we really want?**

DNS lookup


facebook
Email or Phone
Password
Log In
Forgotten account?
NANOG
72
Atlanta, Georgia
February 19 to February 21, 2018
North American Network Operators' Group (NANOG)
@NANOG.org
Home
About
Events
Photos
Community
Reviews
Posts
Create a Page
Like
Share
Learn More
Message
Photos
Charity in Bellevue, Washington
4.8
Community
See all
1,264 people like this
1,276 people follow this
About
See All
Contact North American Network

DNS lookup

```
ximaera@nostromo:~$ sudo tcpdump -qni any tcp > /dev/null
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on any, link-type LINUX_SLL (Linux cooked), capture size 65535 bytes
^C
792 packets captured
794 packets received by filter
0 packets dropped by kernel
ximaera@nostromo:~$ sudo tcpdump -qni any port 53 > /dev/null
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on any, link-type LINUX_SLL (Linux cooked), capture size 65535 bytes
^C
104 packets captured
156 packets received by filter
0 packets dropped by kernel
ximaera@nostromo:~$
```

DNS lookup

```
10:00:34.510826 IP
   (proto UDP (17), length 56)
   192.168.1.5.63097 > 8.8.8.8.53:
      9508+
      A? facebook.com.
      (30)
10:00:34.588632 IP
   (proto UDP (17), length 72)
   8.8.8.8.53 > 192.168.1.5.63097:
      9508 1/0/0
      facebook.com. A 31.13.72.36
      (45)
```

DNS lookup

- Apparently, not rocket science?

- Well, it's not – **for the `(int32)*host_str` feature.**

More to it?

- Geobalancing

MaxMind GeoIP database

GeoIP2 City Database Demo

IP Addresses

8.8.8.8

Enter up to 25 IP addresses separated by spaces or commas. You can also test your own IP address.

Submit

GeoIP2 City Results

IP Address	Country Code	Location	Postal Code	Approximate Coordinates*	Accuracy Radius	ISP	Organization	Domain	Metr Cod
8.8.8.8	US	United States, North America		37.751, -97.822	1000	Google	Google		

37.751, -97.822


Cheney
Wildlife Area
Cheney Reservoir
17
556
Google
Map data ©2018 Google
Terms
2 km

MaxMind GeoIP database

GeoIP2 City Database Demo

IP Addresses

8.8.8.8

Enter up to 25 IP addresses separated by spaces or commas. You can also test your own IP address.

Submit

Sorry, this is wrong!

GeoIP2 City Results

IP Address	Country Code	Location	Postal Code	Approximate Coordinates*	Accuracy Radius	ISP	Organization	Domain	Metr Cod
8.8.8.8	US	United States, North America		37.751, -97.822	1000	Google	Google		

MaxMind GeoIP database

Has its “owner location vs actual location” dilemma.

Generally unreliable for anything except statistics.

- https://stackoverflow.com/questions/22986794/continuously-decreasing-accuracy-of-maxmind-geolite-city
- https://www.techdirt.com/articles/20160413/12012834171/how-bad-are-geolocation-tools-really-really-bad.shtml
- https://splinternews.com/how-an-internet-mapping-glitch-turned-a-random-kansas-f-1793856052

MaxMind GeoIP database

Has its “owner location vs actual location” dilemma.

Generally unreliable for anything except statistics.

- There’s no *geography* on the Internet, just **network topology**.

- There are no countries,
  just autonomous systems and their relations.

ASN and prefix targeting: example

https://ns1.com/solutions/technical-solutions/filter-chain

- Filters are like little programs that run inline for every DNS query.
- They are attached directly to RFC-compliant DNS records

NETFENCE_ASN Restrict to answers where Autonomous System (AS) of requester IP matches AS list

NETFENCE_PREFIX Restrict to answers where requester IP matches prefix list

Contemporary DNS server requirements

- Latency reduction: ~~geobalancing~~ **prefix targeting**

Dynamic configuration

```
08:35 < j***k> and also VERY FUNNY PEOPLE
08:35 < m***k> j***k likes us \o/
08:35 < s***k> we like j***k
08:36 < d***n> DNS  DEPLOY
08:36 < j***v> What is this "DNS DEPLOY" thingy you guys keep screaming?
08:36 < d***n> j***v, when we deploy new dns content
08:36 < j***k> http://i.qkme.me/364h55.jpg
08:36 < j***v> Alearting eachother?
08:36 < d***n> yup
08:37 < j***v> d***n: Why?
08:37 < d***n> in case there's problems and I guess also as a locking mechanism
```

https://labs.spotify.com/2017/03/31/spotifys-lovehate-relationship-with-dns/

Dynamic configuration

DNS is not a static config anymore, this is essentially an API for configuration management systems and applications:

- Provisioning
- Stats
- Policy management

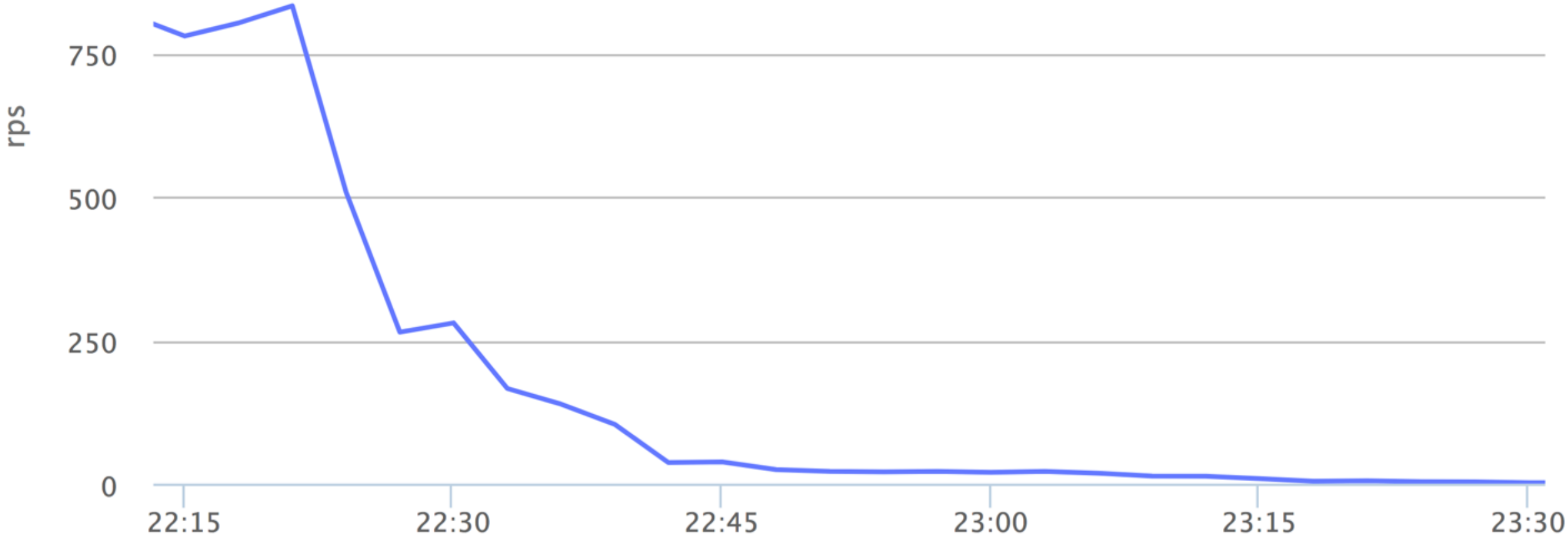
Enterprises **will** want this sooner or later.

Treating DNS **not** as an API is error-prone.

Contemporary DNS server requirements

- Latency reduction: ~~geobalancing~~ **prefix targeting**
- Dynamic configuration
- Failover

Failover, TTL 120s



Contemporary DNS server requirements

- Latency reduction: ~~geobalancing~~ **prefix targeting**
- Dynamic configuration
- Failover
- Vulnerability intelligence
- **DDoS attacks**

DNS DDoS

- Volumetric attacks:
  effective line rate challenges/handshake
- Water Torture and so on:
  query analysis, statistics and blacklists
- Anycast is **necessary**

Contemporary DNS server requirements

- Latency reduction: ~~geobalancing~~ **prefix targeting**
- Dynamic configuration
- Failover
- **DDoS attacks**
- DNSSEC, TLS, etc. More than 180 RFCs

Contemporary DNS server requirements

- Latency reduction: ~~geobalancing~~ **prefix targeting**
- Dynamic configuration
- Failover
- **DDoS attacks**
- DNSSEC, TLS, etc. More than 180 RFCs

Okay, now this is rocket science ☹

What about service providers?

Thousands out there!

- Dyn
- NS1
- Route 53
- Name.com
- Azure DNS
- Google Cloud DNS
- Cloudflare
- … (sorry for not putting *your* favorite provider in the list)

What about service providers?

Thousands out there!

- How to choose?

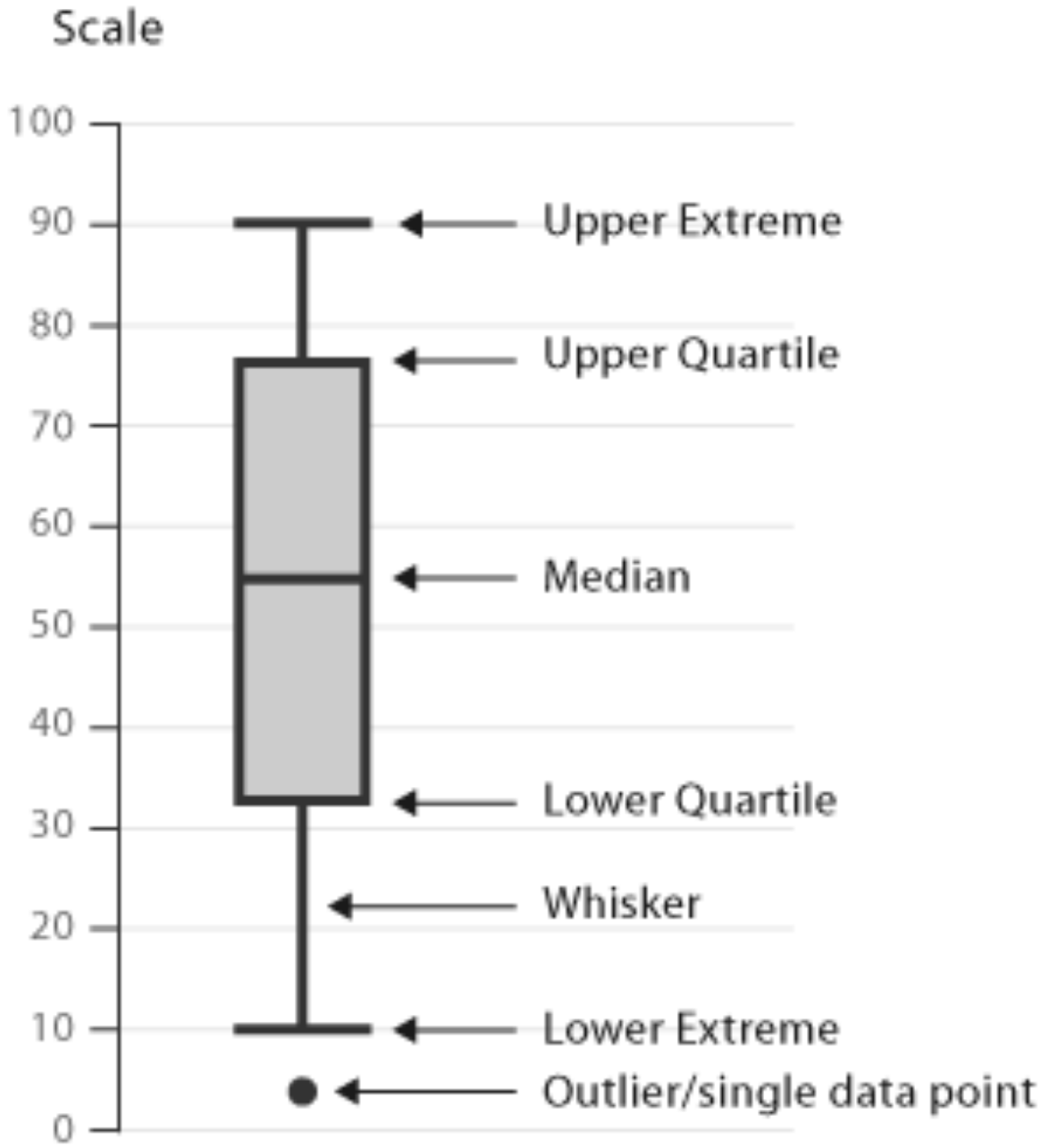
What about service providers?

Thousands out there!

- How to choose?
- Well, why?

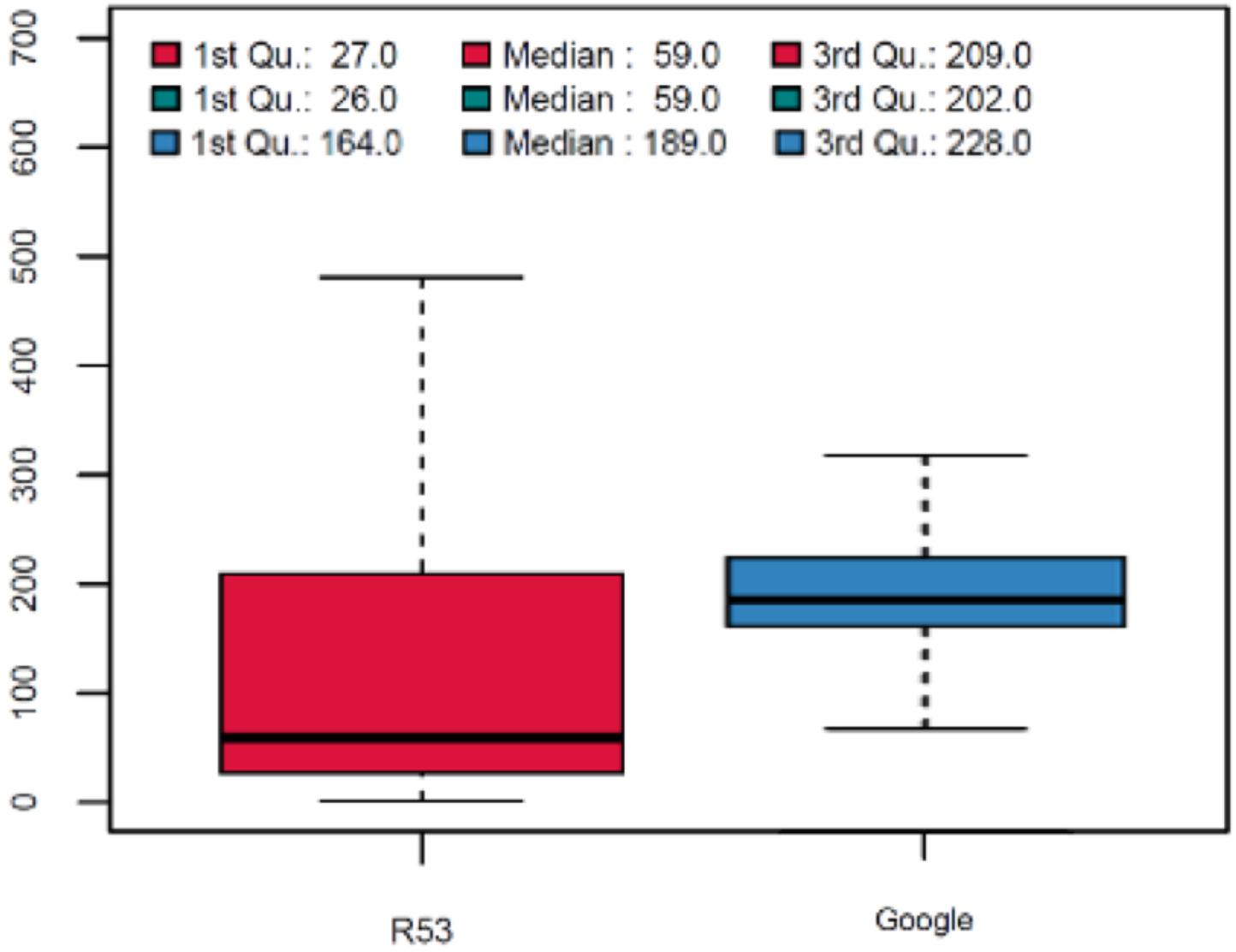
SRTT: Smoothed Round Trip Time

- A mechanism intended to help to run a lot of nameservers simultaneously for a zone
- Deployed in most SOHO and enterprise networks
- NS1 study suggests up to 90% Internet traffic being serviced by SRTT-enabled resolvers

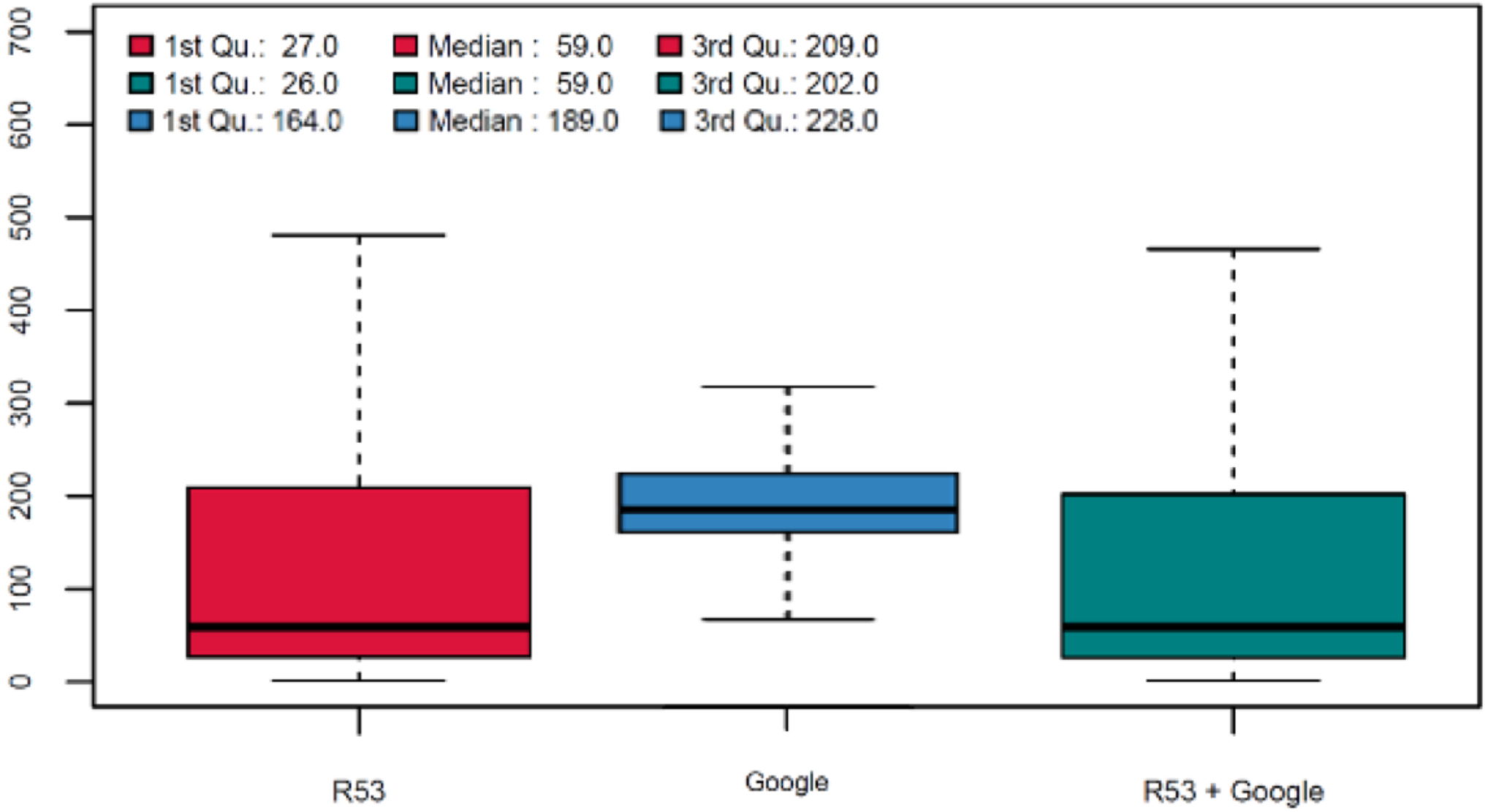

Scale
100
90
80
70
60
50
40
30
20
10
0
Upper Extreme
Upper Quartile
Median
Lower Quartile
Whisker
Lower Extreme
Outlier/single data point

“Boxplot”

SRTT

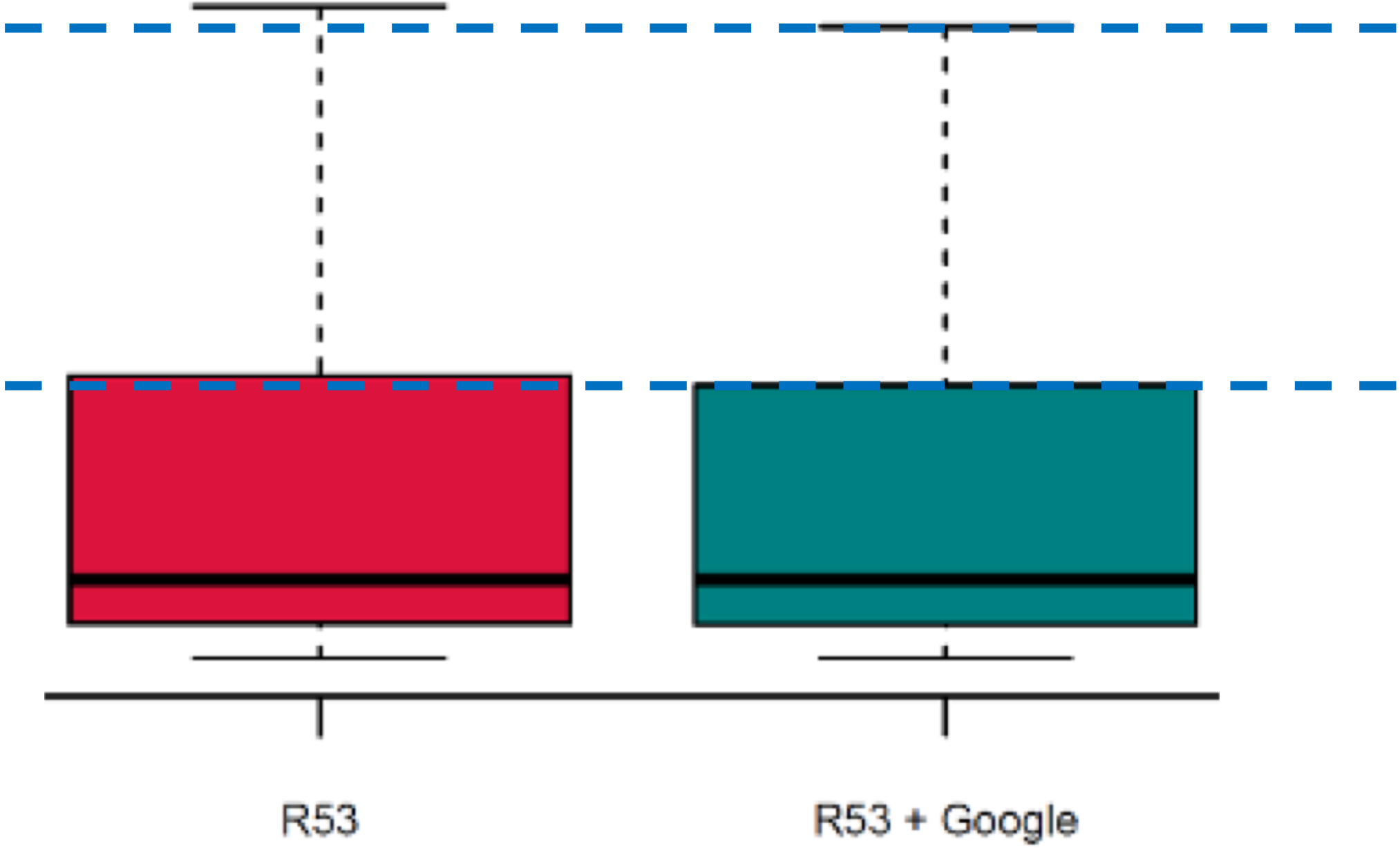

1st Qu.: 27.0
Median : 59.0
3rd Qu.: 209.0
1st Qu.: 26.0
Median : 59.0
3rd Qu.: 202.0
1st Qu.: 164.0
Median : 189.0
3rd Qu.: 228.0
0
100
200
300
400
500
600
700
R53
Google

SRTT



https://blog.serverfault.com/2017/01/09/surviving-the-next-dns-attack/

SRTT



How to choose a service provider

- The more you have, the better
  - Up to 4-6 will be fine
- Easy to compare and replace the underperforming ones
- Helps also with maintenance windows and downtime issues
- AXFR doesn’t support a lot of features
  - Prefer providers with nice API

Q&A