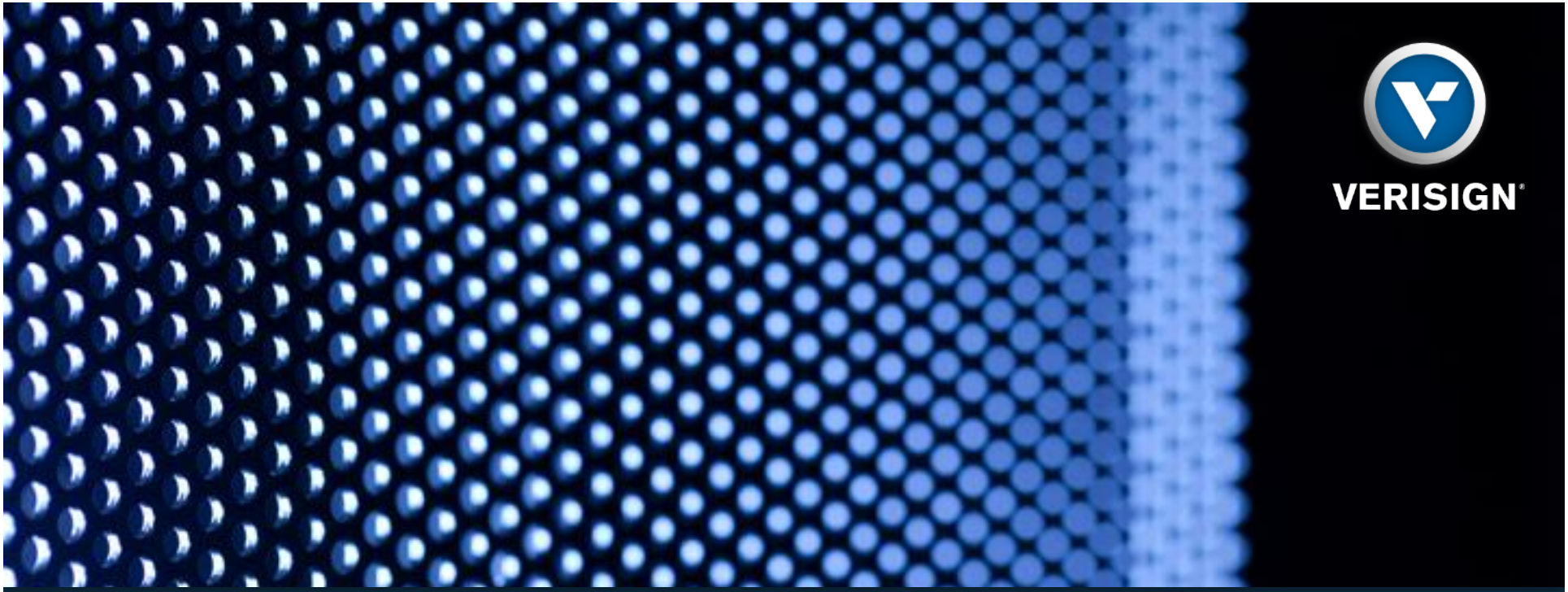


VERISIGN

An Analysis of TCP Traffic in Root Server DITL Data

Matt Thomas & Duane Wessels, Verisign Labs

DNS-OARC 2014 Fall Workshop, Los Angeles

October 12, 2014

Agenda

- TCP Refresher
  - Taxonomy of TCP Session
  - Longitudinal Measure of TCP Sessions
- A Synopsis of TCP-based DNS Queries
  - IP Version, rCode, qType, Message size, Port randomization, qNames
- Measuring RTTs
  - Number of distinct IPs, /24s, and ASNs and their request distribution
  - By year and root
  - Geographically

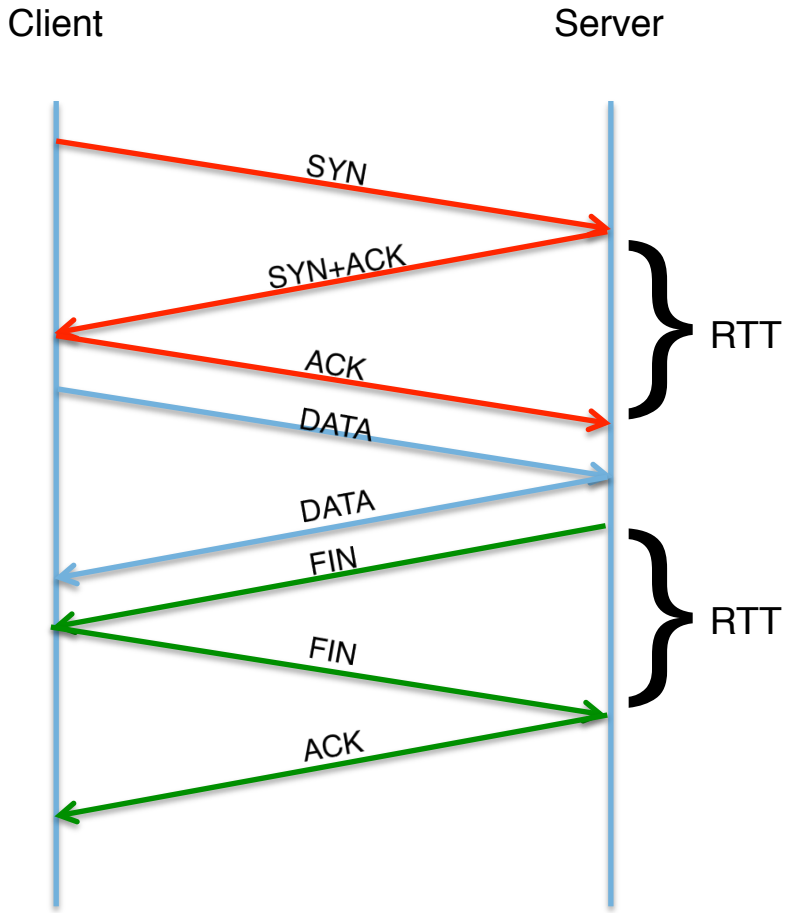
DNS & TCP

- Truncation
- Spoof prevention
- Privacy

- TCP provides opportunities to calculate round trip time (RTT) latency.
  - Setup
  - Teardown

- *Assuming it was all captured*

… oh wait – it wasn't.



Caveats on DITL Data

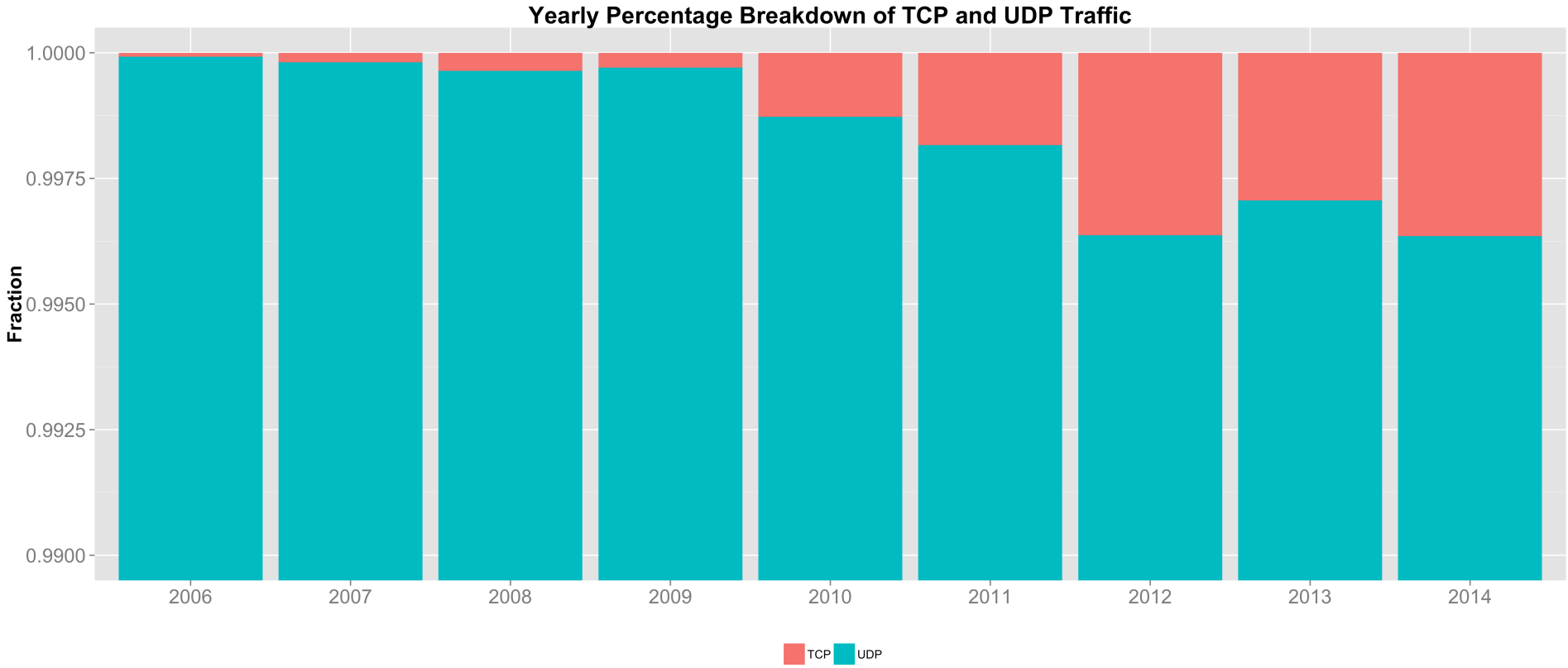
- DITL data is … messy and confusing!
- Not all operators provide data every year
- Sometimes we get UDP but not TCP
- Sometimes we get TCP in one direction only
- Sometimes we get TCP from a subset of the operator's sites
- Sometimes we capture attack traffic which skews the results (i.e., 2012)

- DNS responses are either not collected or filtered out for the "clean" data sets.

 = UDP ↓=TCP recv ↑=TCP xmit

	2009	2010	2011	2012	2013	2014
A	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
B		↓↑				
C	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
D		↓↑	↓↑		↓↑	
E	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
F	↓	↓	↓	↓	↓	↓
G		↓↑				
H	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑
I		↓	↓	↓	↓	↓
J		↓↑	↓↑	↓↑	↓↑	↓↑
K	↓	↓↑	↓↑	↓↑	↓↑	↓↑
L		↓↑	↓↑	↓↑	↓	
M	↓↑	↓↑	↓↑	↓↑	↓↑	↓↑

UDP vs. TCP Connections Over Time



TCP has grown from 0.002% in 2009 to 0.36% in 2014.

Year	Fraction TCP
2009	0.00029
2010	0.0022
2011	0.0018
2012*	0.0036
2013	0.0029
2014	0.0036

* 2012 data skewed by attack traffic

TCP Sessions Taxonomy

DNS Query :: Complete “standard” DNS query

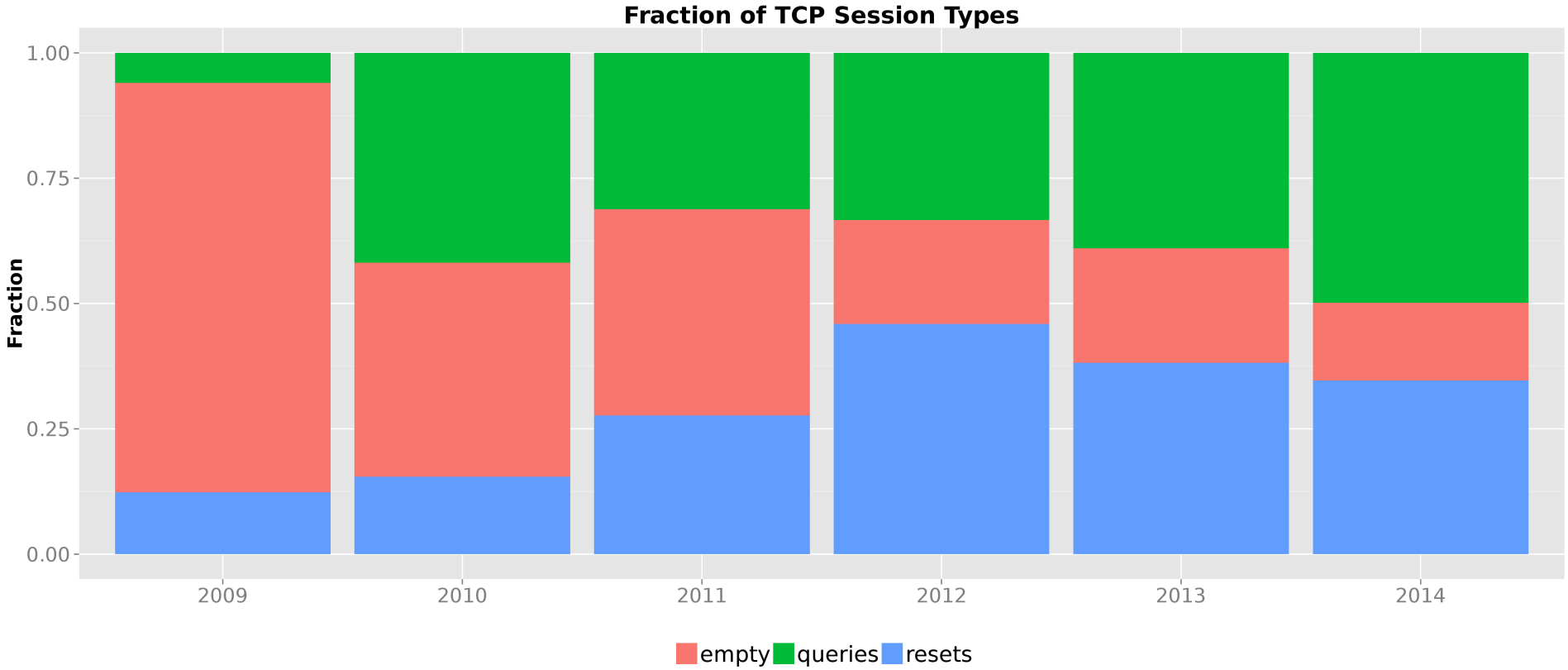
EMPTY :: No payload is transmitted.

- Session teardown usually incomplete

RESET :: A reset is sent to immediately kill session.

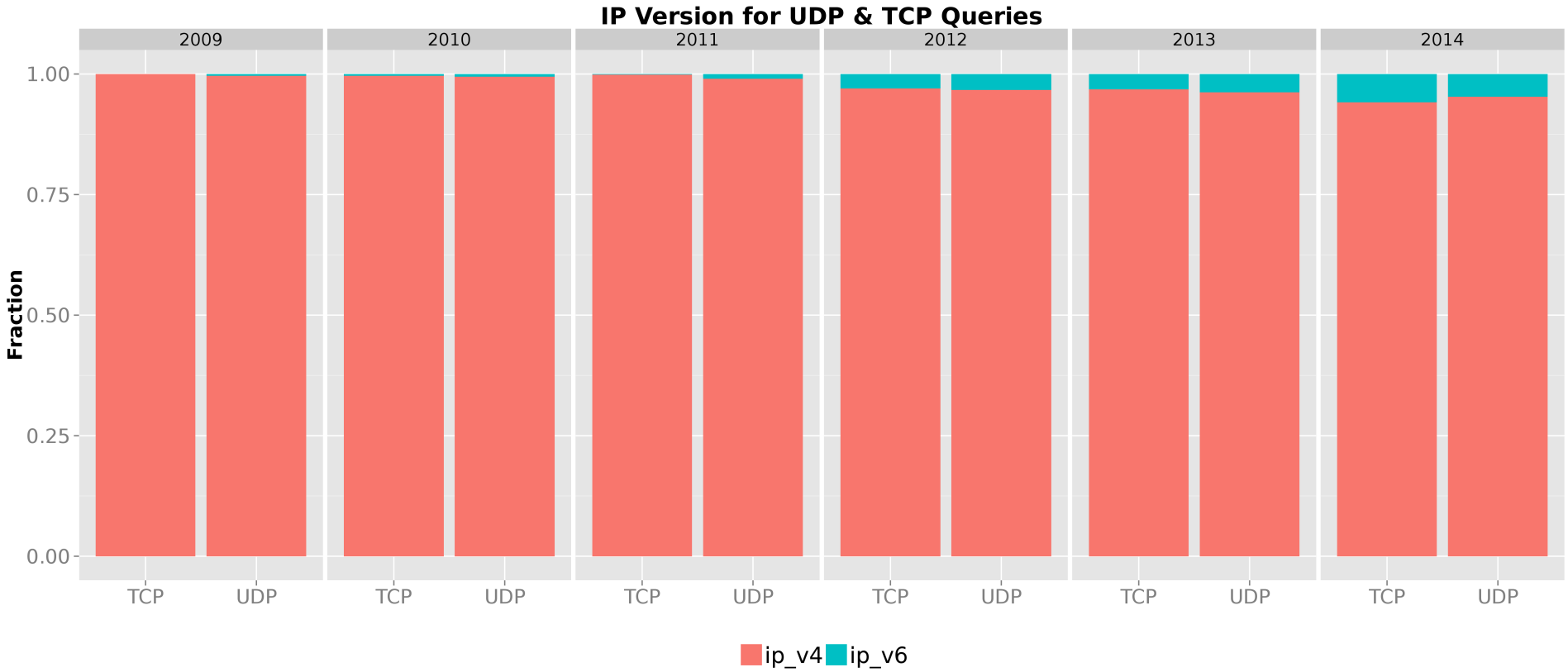
- Typically follows initial SYN/ACK from server
- Can be sent at any time though.

TCP Session Types Over The DITL Years



- We observe more valid TCP queries (as a percentage) over time.

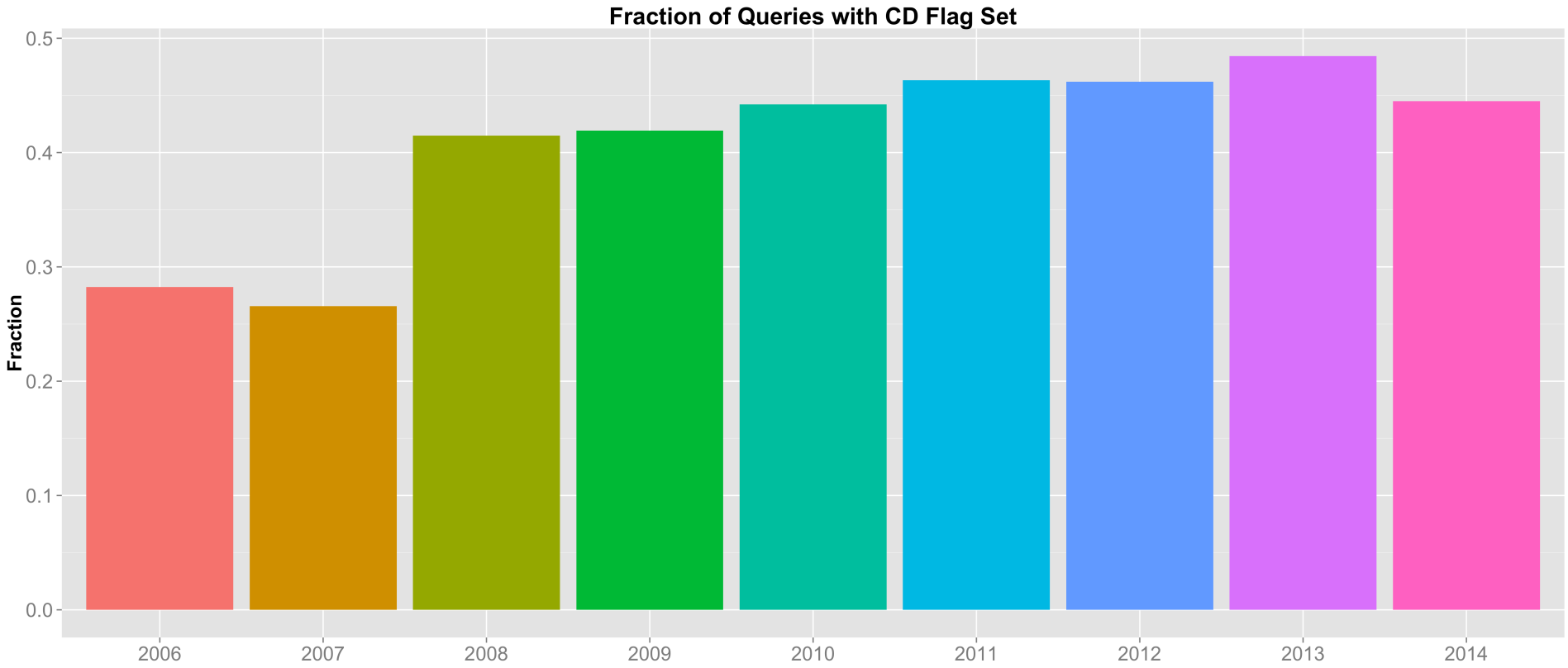
Growth in IPv6



- IPv6 growth rates in UDP and TCP are similar over time.
- In 2014, 4.7% of UDP traffic was IPv6, and 5.8% of TCP was IPv6

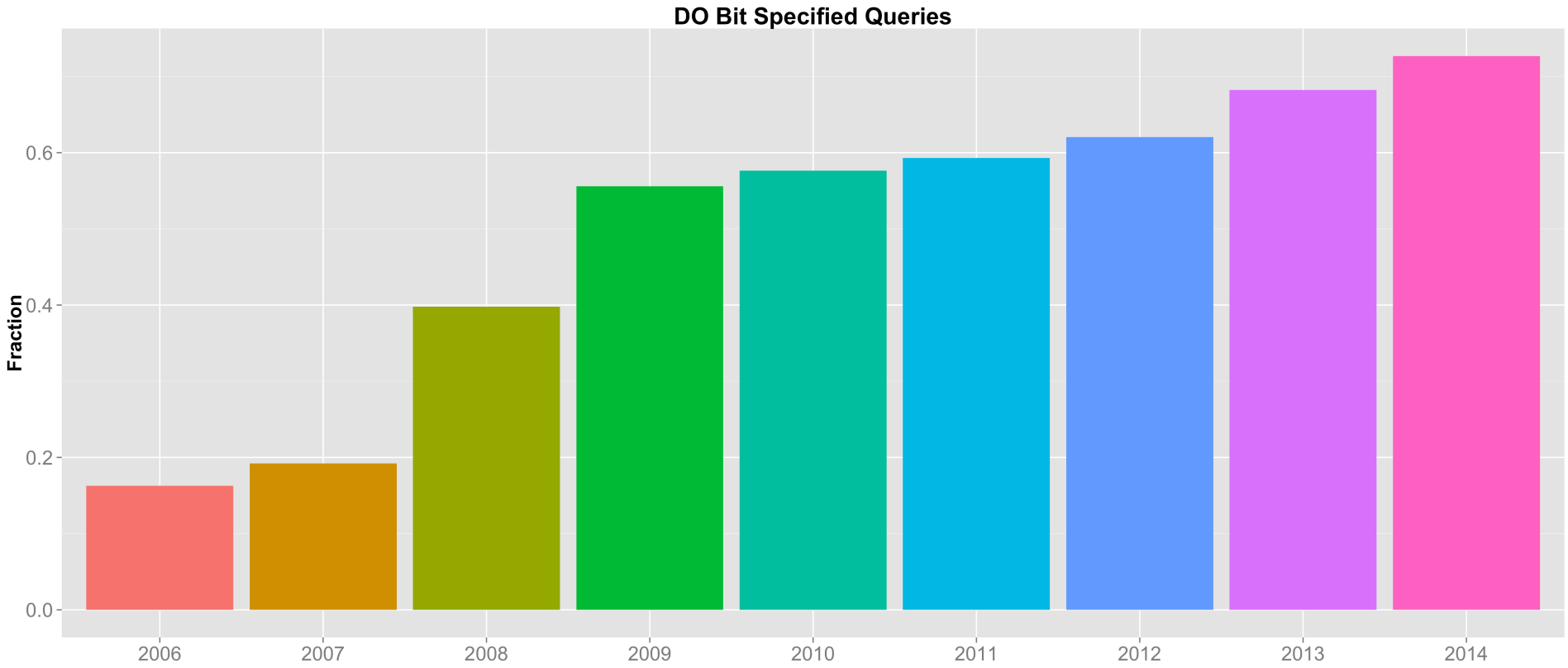
EDNS0, Bufsize, Flags and Friends

CD Bit Specified



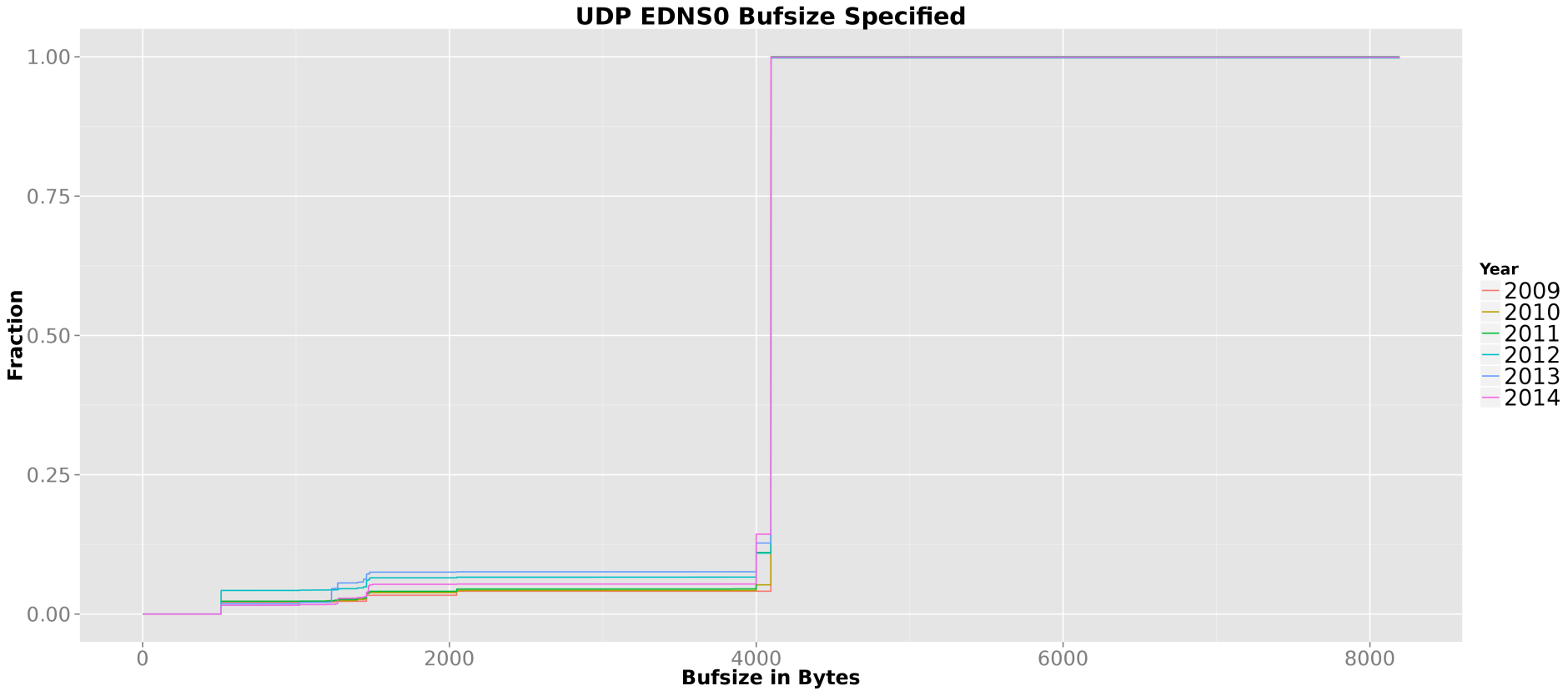
- Upward trend in CD Flags
- Not proportional to D0=1 query growth rate though…
- Why is the percentage so high?

DO Bit Specified



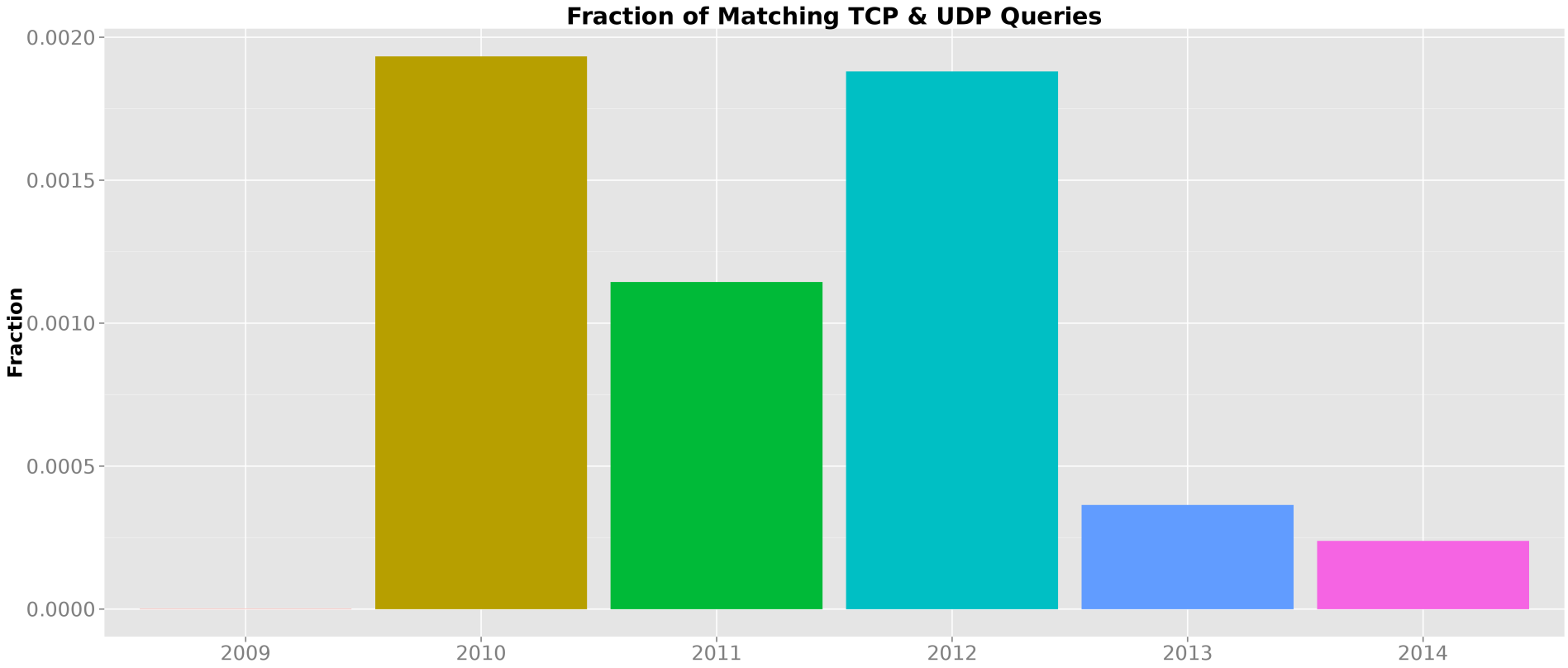
- Steady increase in queries from DNSSEC-enabled clients.

EDNS0 Bufsize



- 4096 remains the lion's share
- Previous smaller Bufsizes in 2012/2013 have shifted to higher

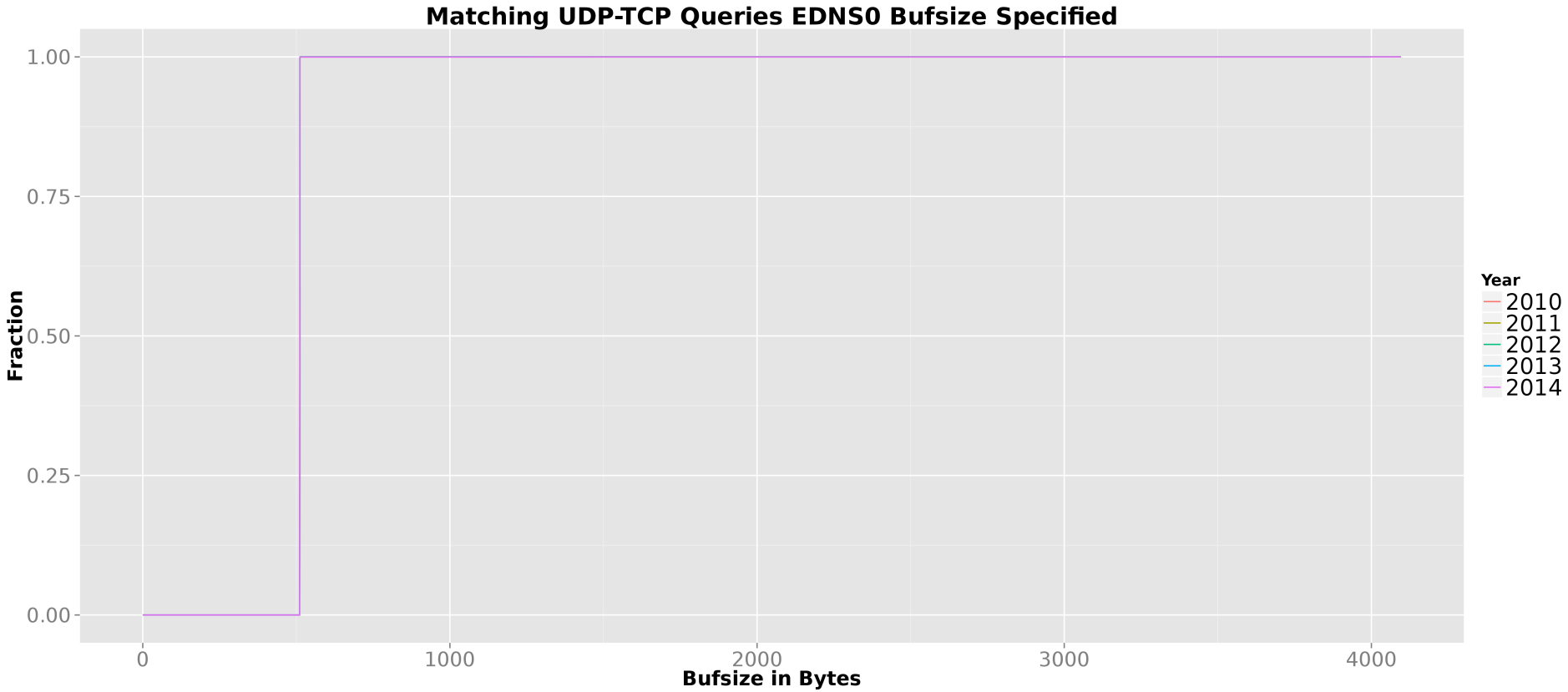
TC Bit Returned?



- Out of all UDP queries, how often do we see a subsequent TCP from the same source for the same name?*

*excluding “.”

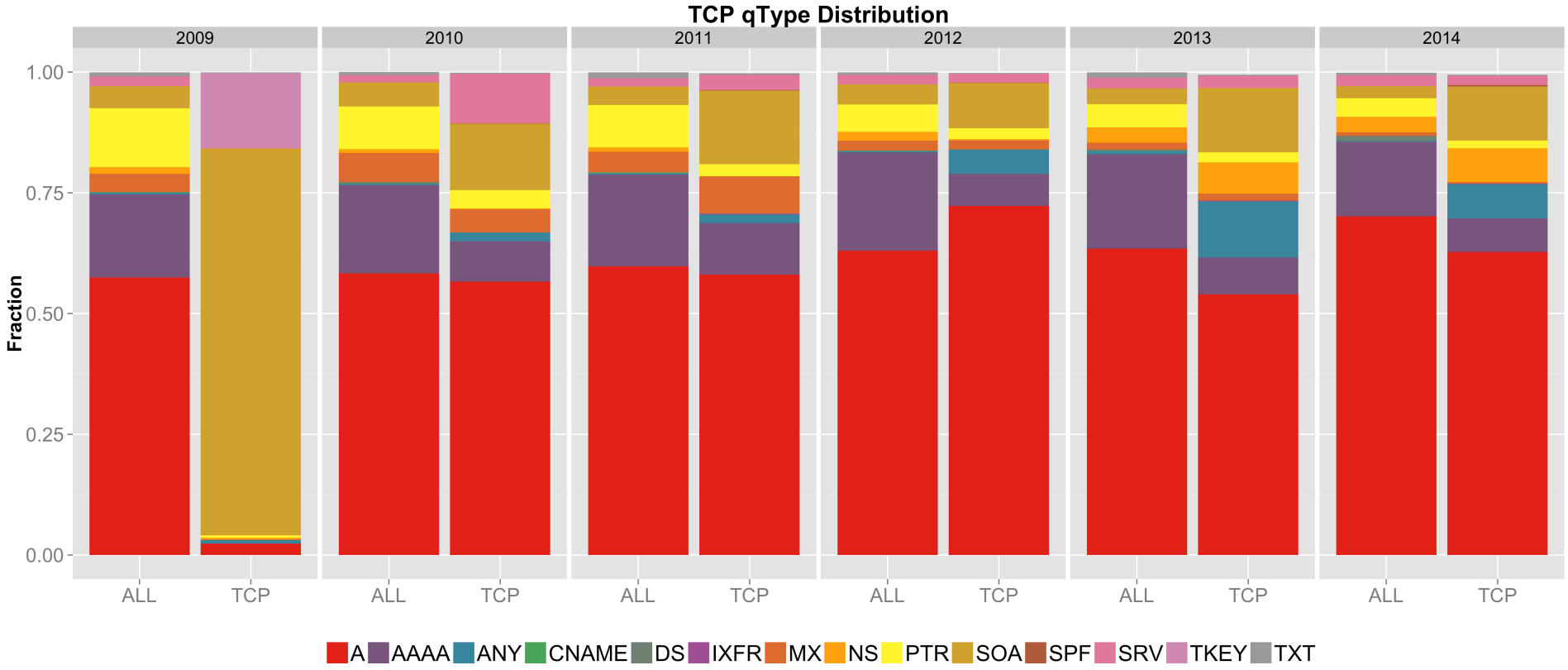
EDNS0 Bufsize for Matching UDP - TCP



- Not surprisingly the Bufsize is set to 512 bytes for all matching UDP – TCP “TC” queries

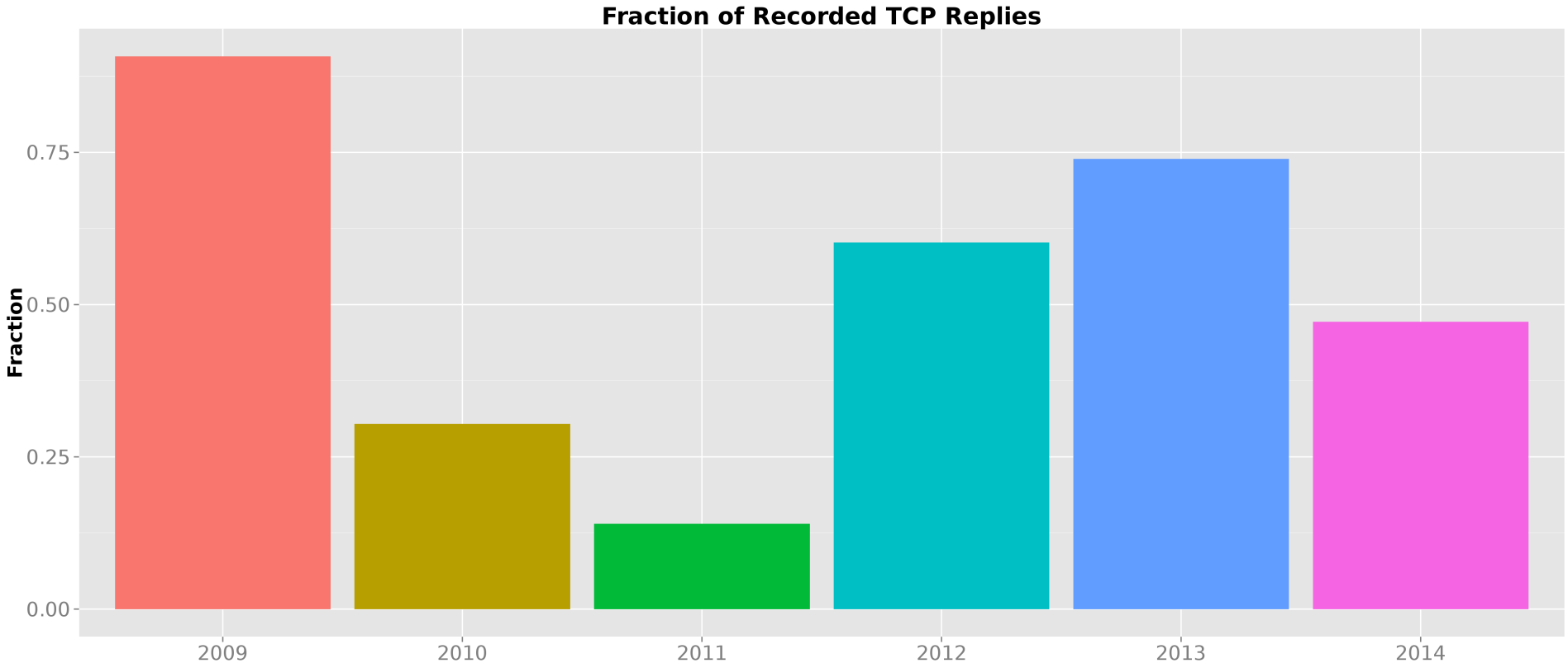
TCP Queries and Their Responses

TCP DNS Queries by qType



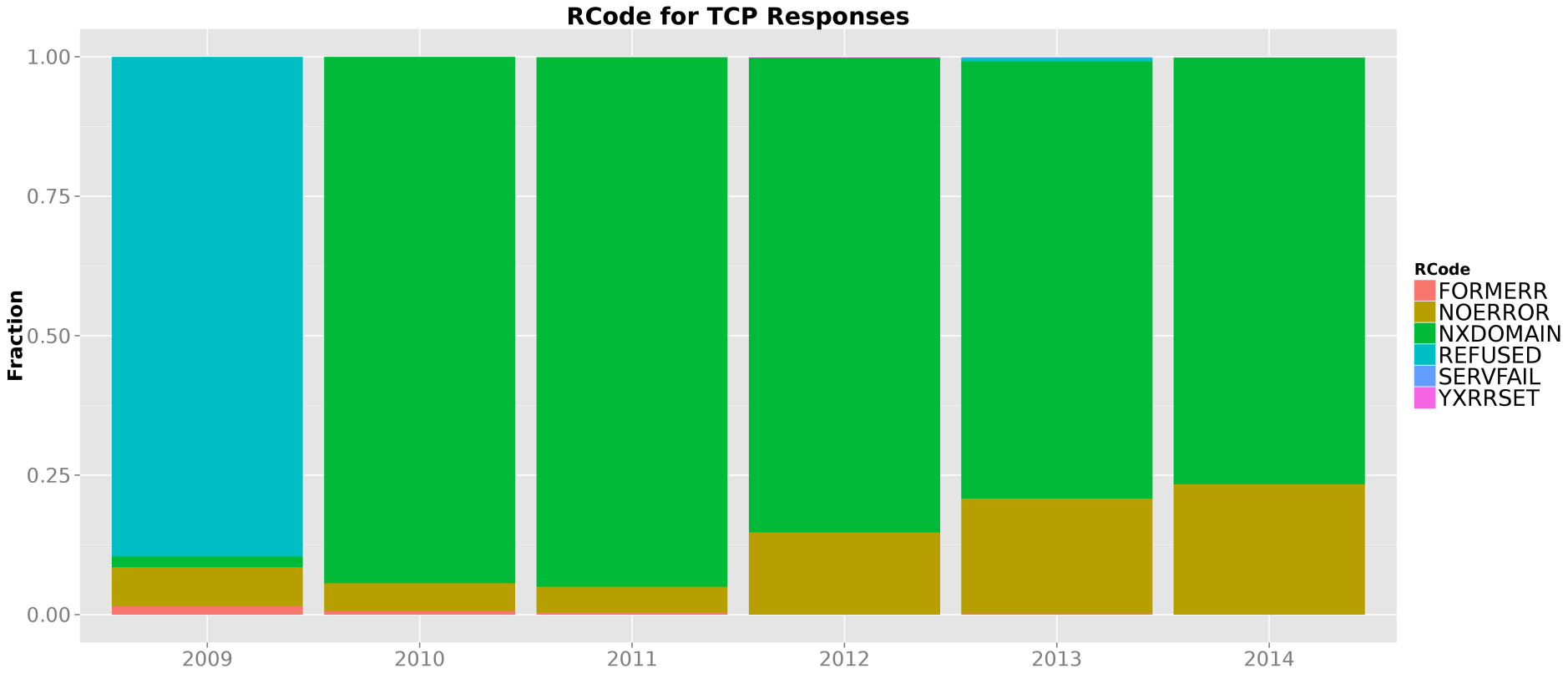
- TCP tends to have more ANY and SOA, and fewer PTR.
- 2009 TCP is largely SOA (UPDATE messages).

Fraction of TCP Responses to TCP Queries



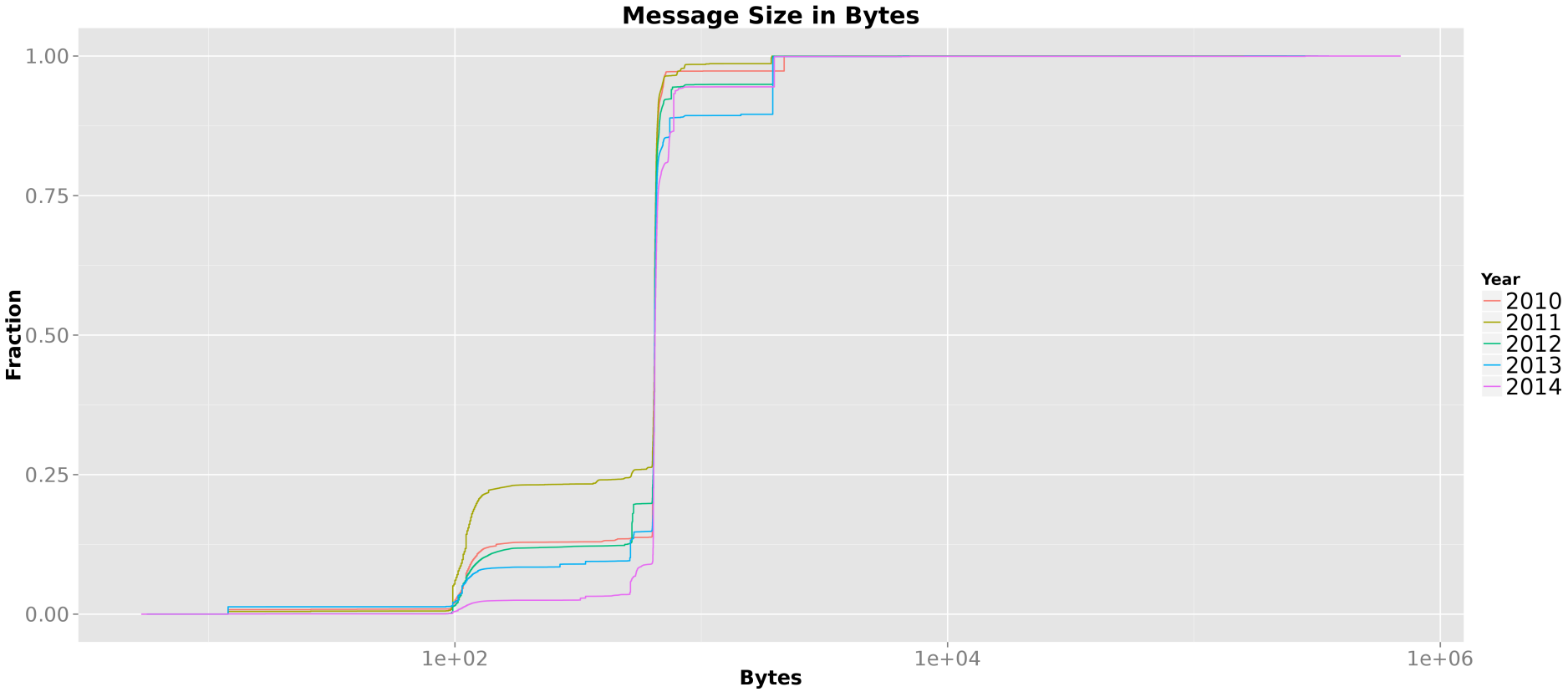
- **Various amount of the TCP query response were captured during DITL collections.**

TCP DNS Responses RCodes



- More NOERROR over time
- Fewer NXDOMAIN over time
- 2009 REFUSED due to UPDATEs

TCP DNS Responses Message Size



- **As expected the payload of a TCP response is large.**
- **Median =~ 648 Bytes**

RTT Latency Measurements

Caveats on Latency Measurements

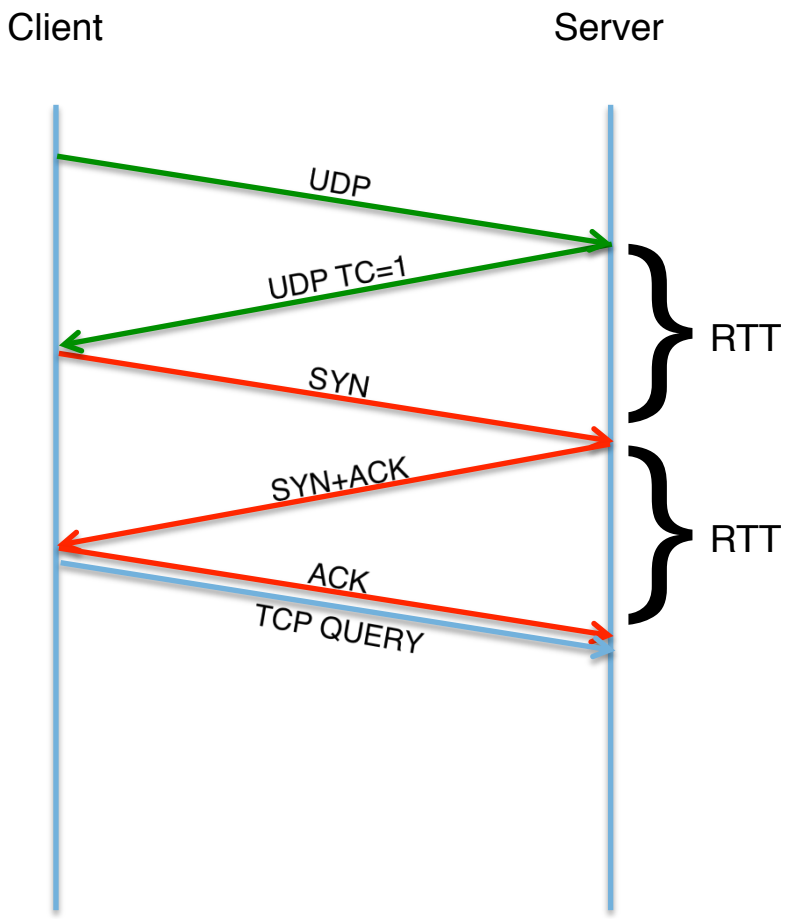
- Most importantly, note that here we report passive measurements from, essentially, the entire Internet. We have no control over the client side -- where they are, their software, their configuration, their network congestion.

- This is significantly different than active measurement infrastructure (Planet Lab, RIPE Atlas, Thousand Eyes, Catchpoint, etc).

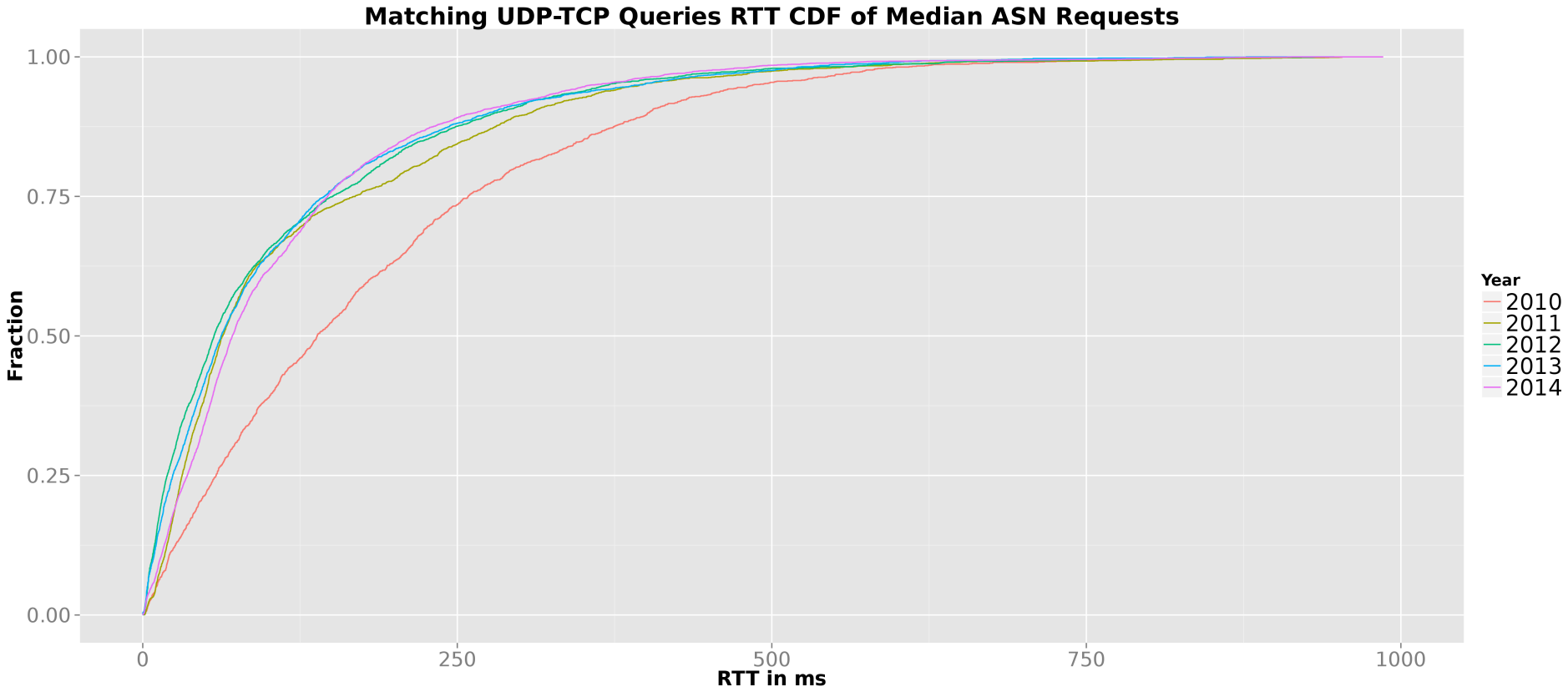
- We do not believe the two methods are comparable.

Latency based on time between TC=1 and TCP

- Earlier we talked about matching UDP queries to followup TCP queries.
- If we assume minimal processing delays, the time between UDP and TCP can be interpreted as 2x RTT.

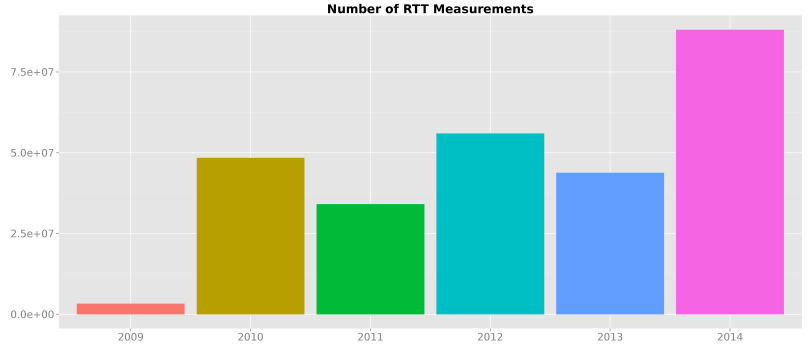


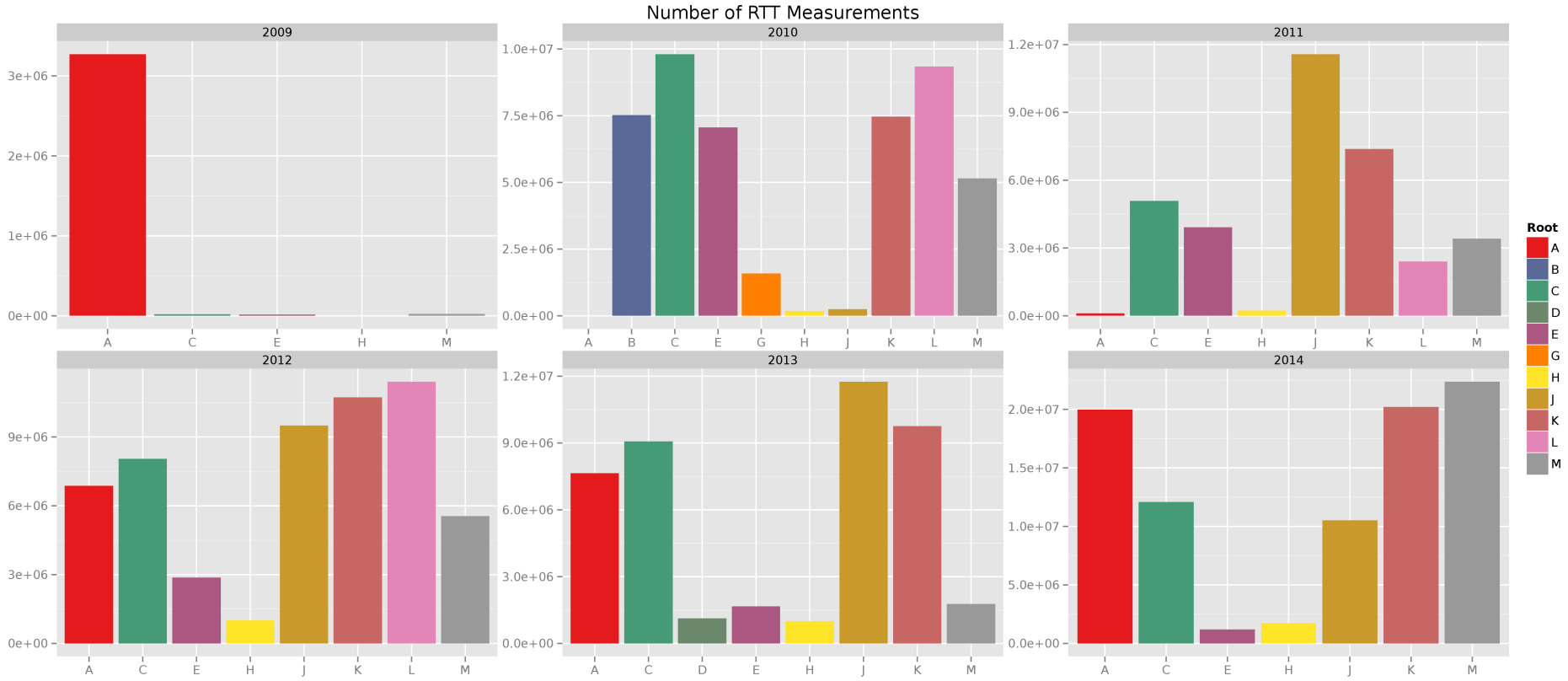
RTT Latency for Matching UDP - TCP



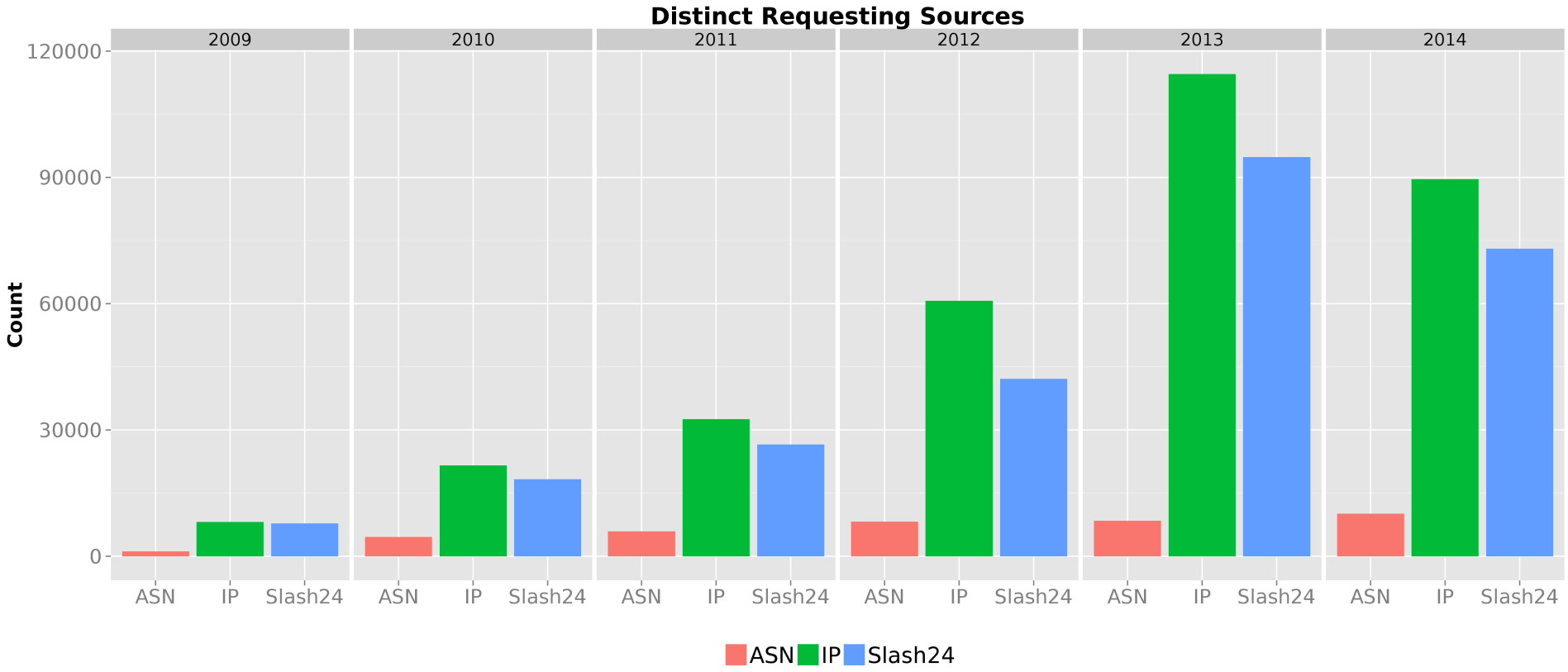
- Big change pre/post DNSSEC
  - Sample size increased
- How does this compare to TCP setup and teardown RTTs?

Roots Contributing to RTT Measurements



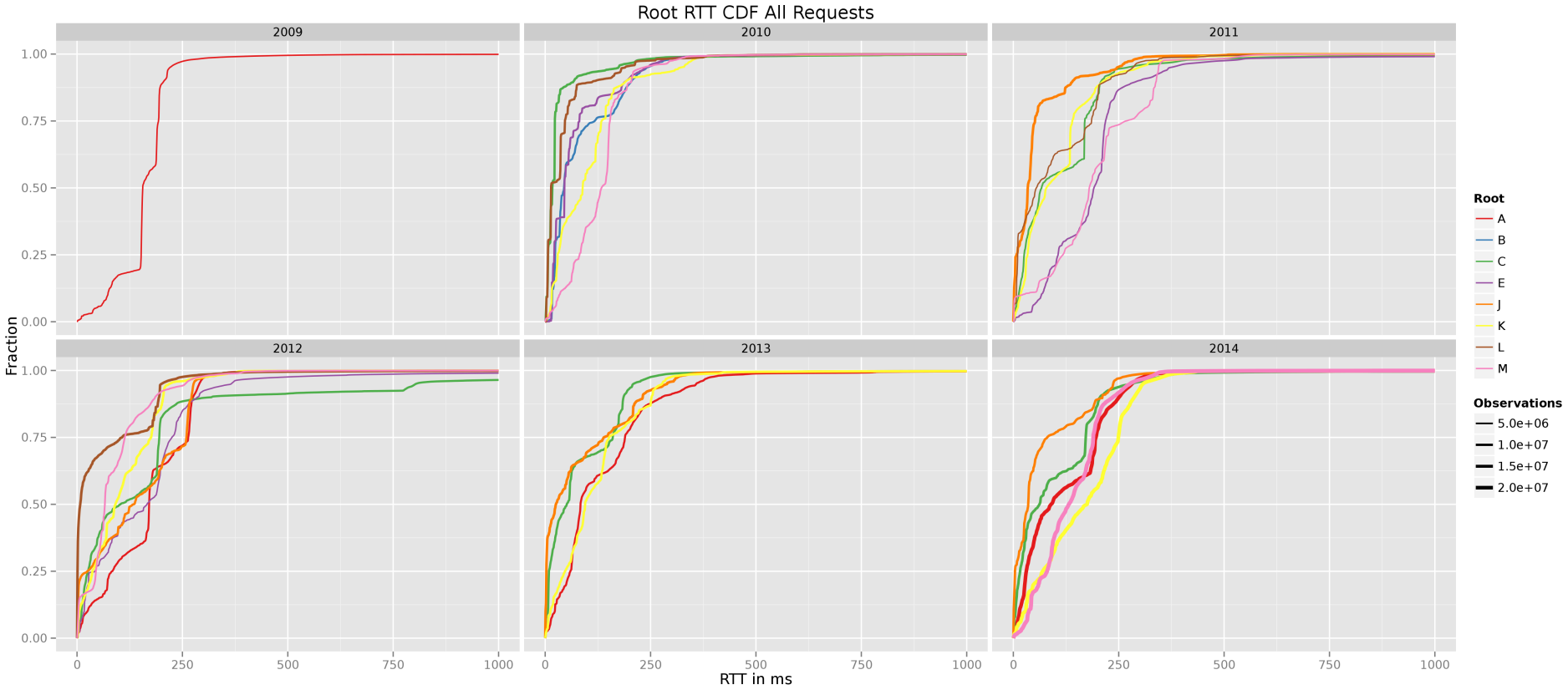


Diversity of Requesting Networks



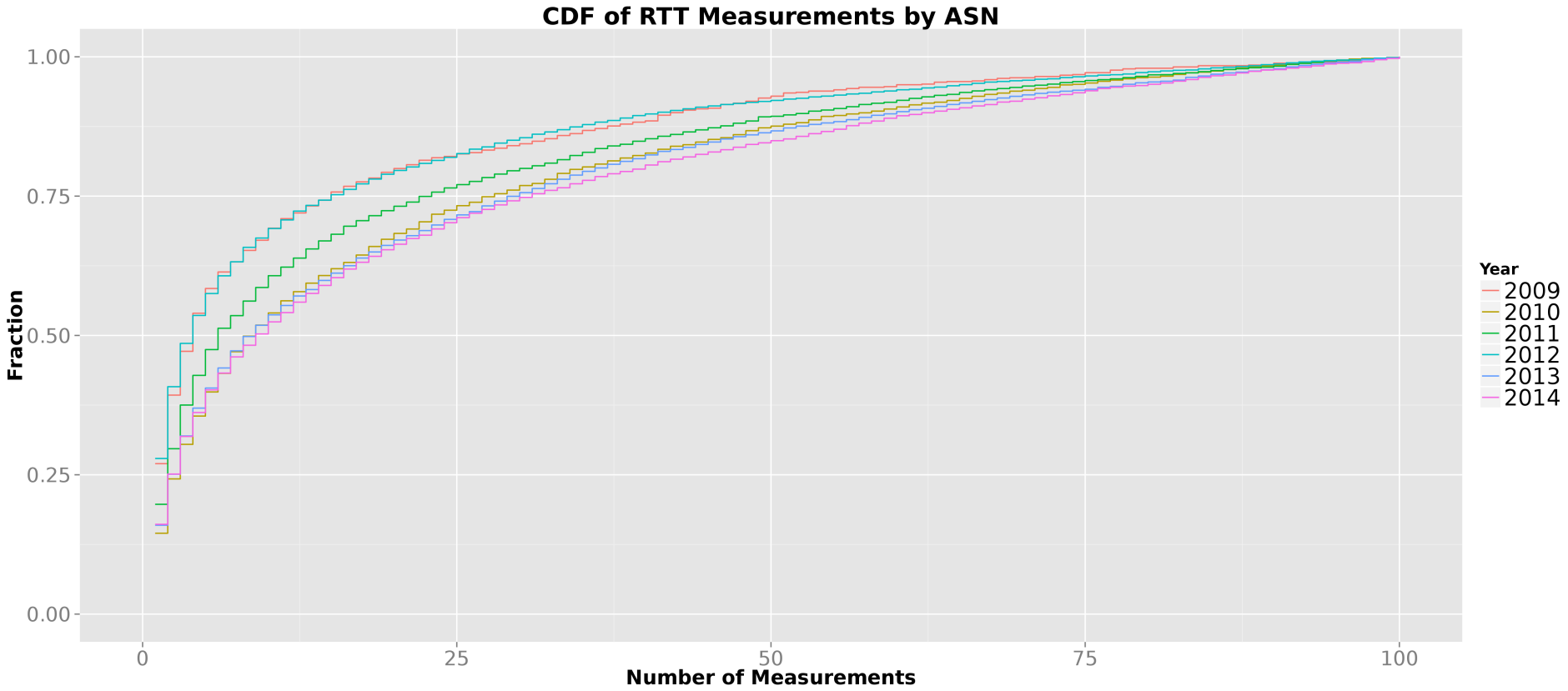
- Steady increase in the number of distinct sources over time.

RTT CDF plots



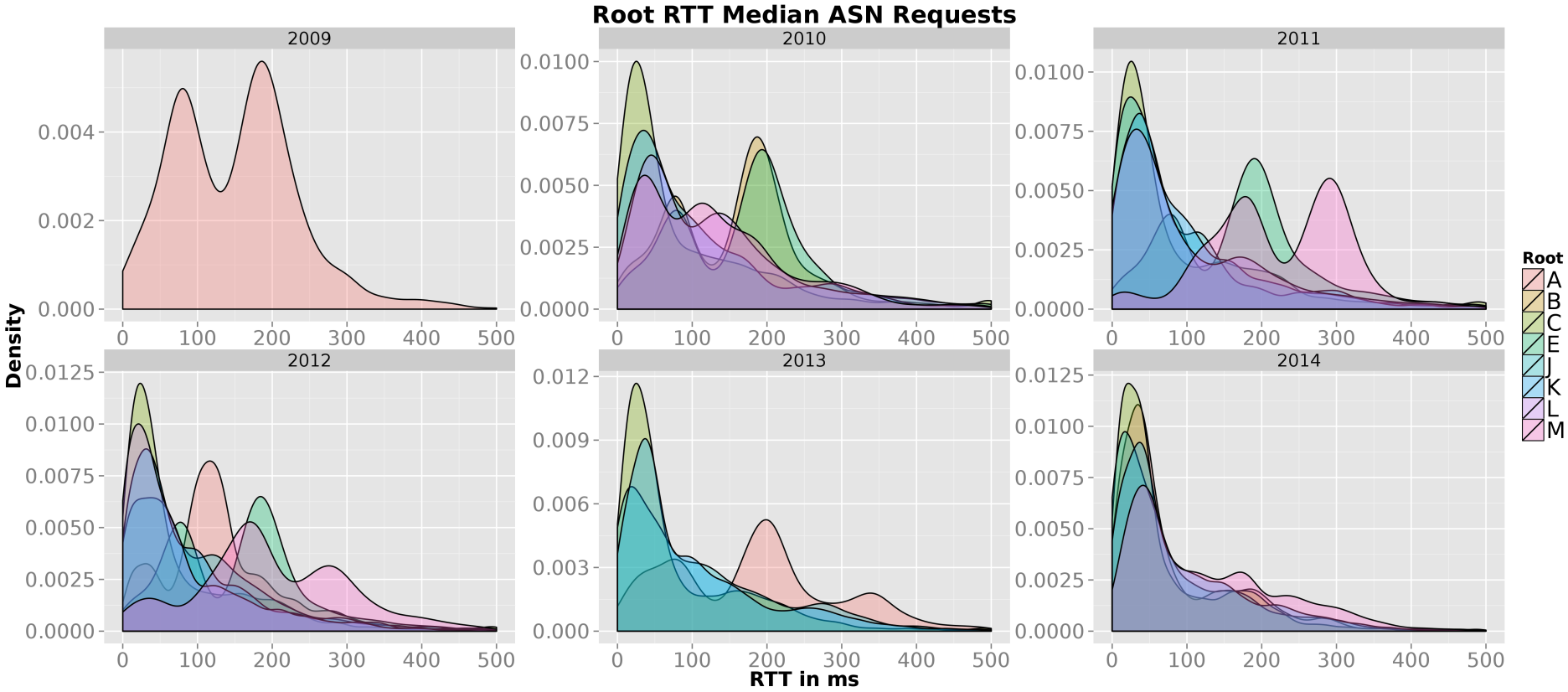
- Quite a lot of variance year-to-year and server-to-server.
- Difficult to draw conclusions.
- Results can be skewed by single sources with frequent queries, not to mention attacks…

Number of RTT Measures by ASN



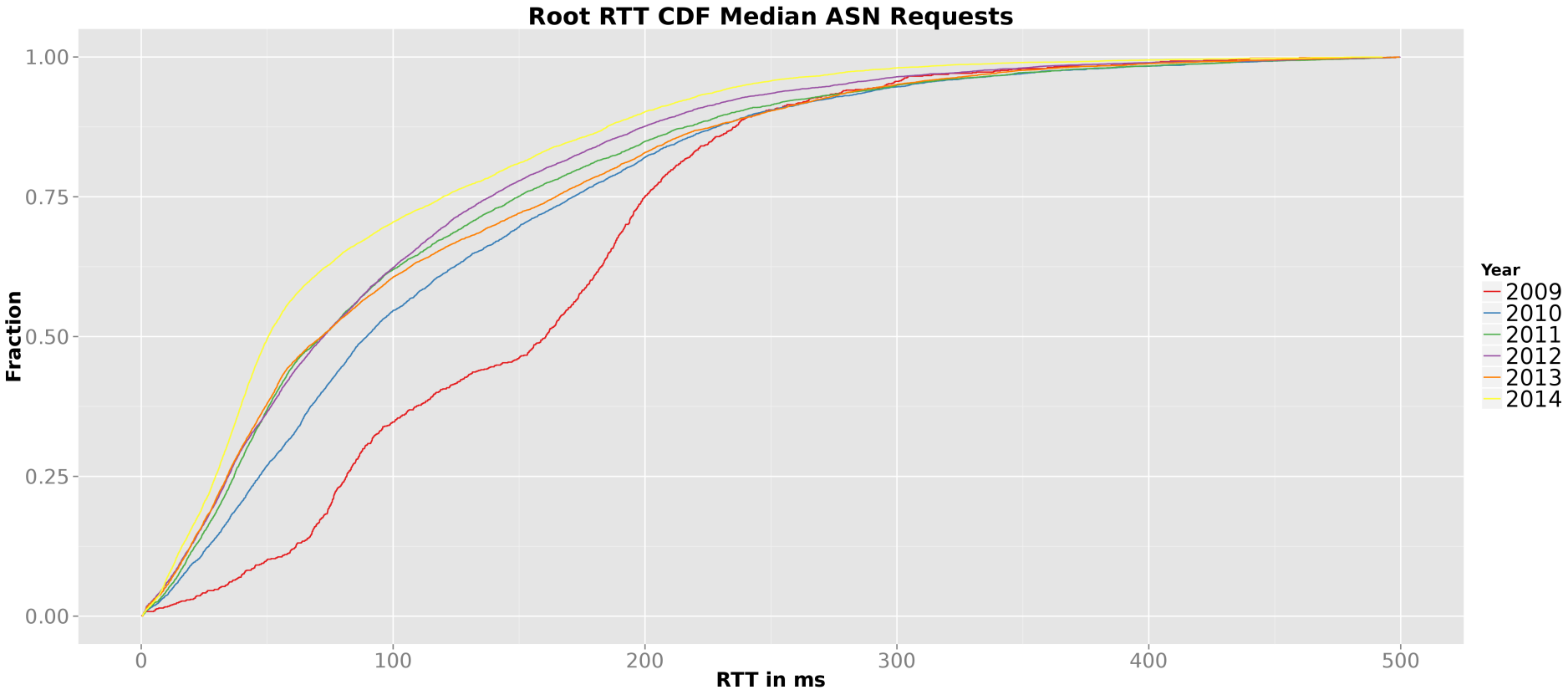
- Central Limit Theorem :: Ideally have N > 30
- Grouping RTT measures by ASN unfortunately does not approach CLT :: Just use median RTT for a given ASN

Distribution of per-ASN Medians



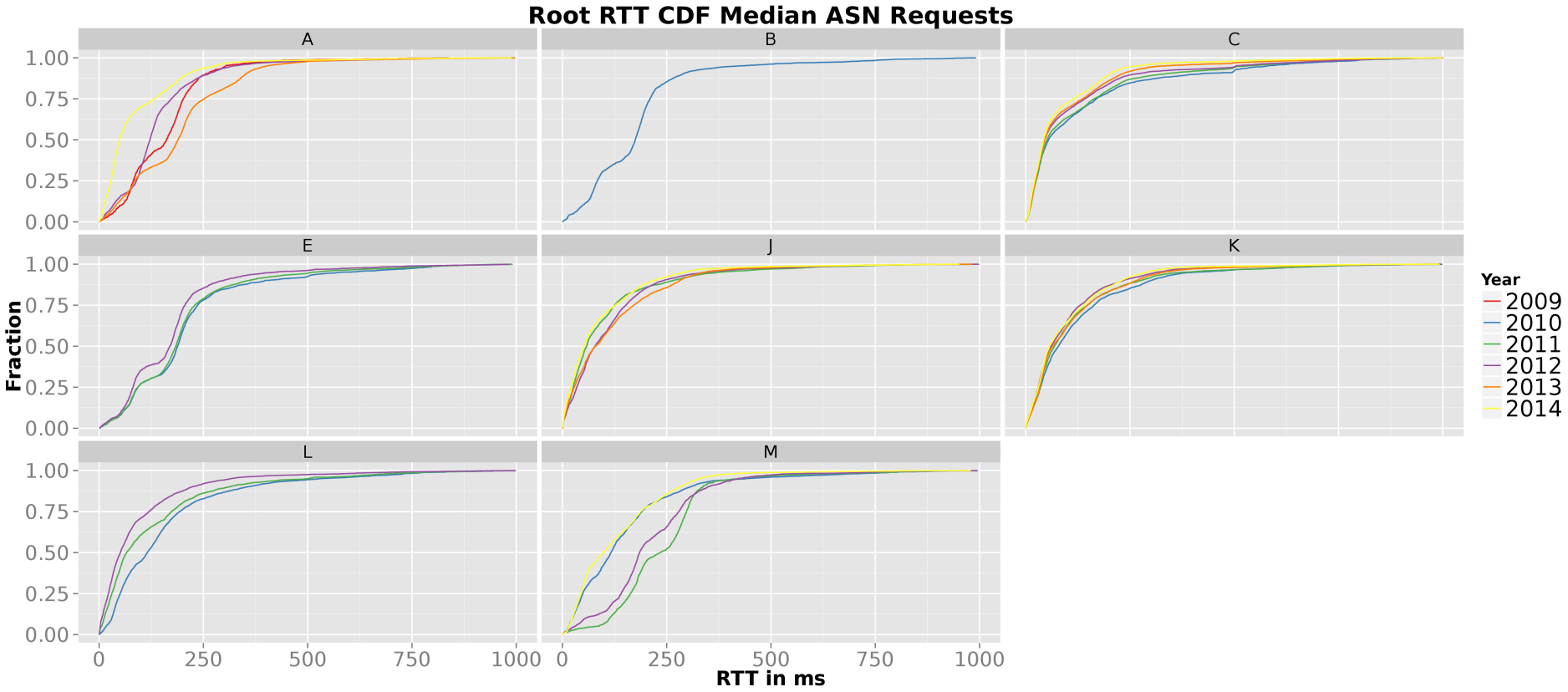
- Distributions look more like what we expect from Internet measurements.
- Shifting to lower latency over of years and long tail distribution.

Distribution of per-ASN Medians



- In general it looks like the RTT latency is improving over time.
- Similar to RTT for “TC” UDP-TCP queries.

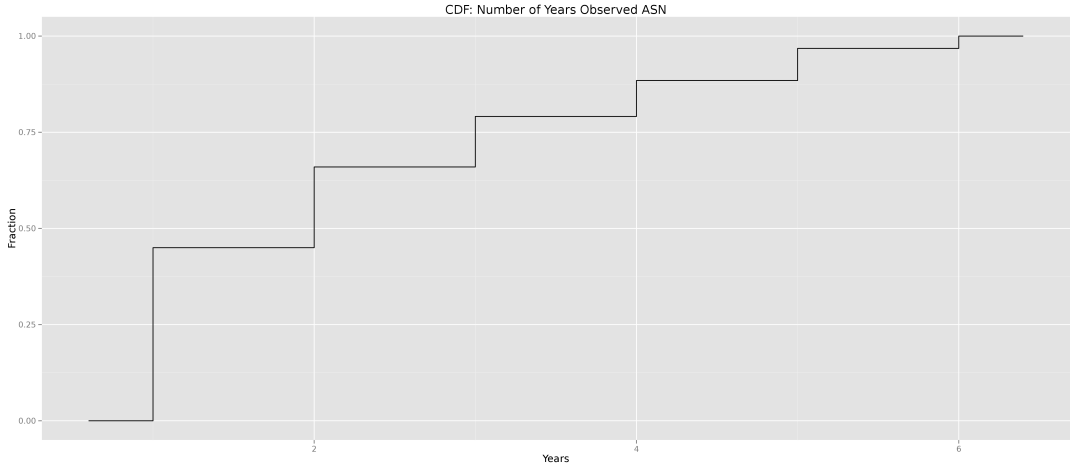
Distribution of per-ASN Medians



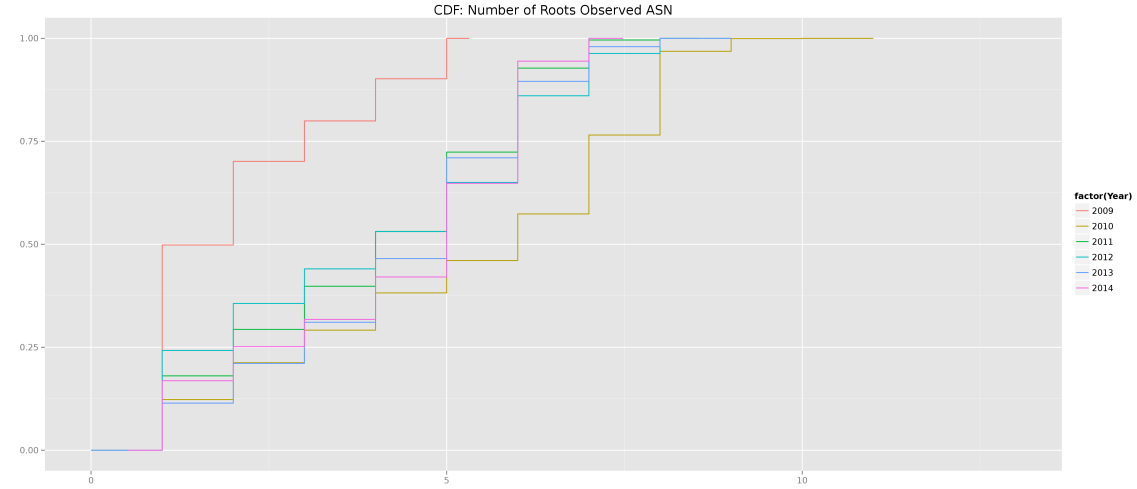
- Different roots exhibit variable amounts of variance over years.

RTT Latency Comparison Year over Year

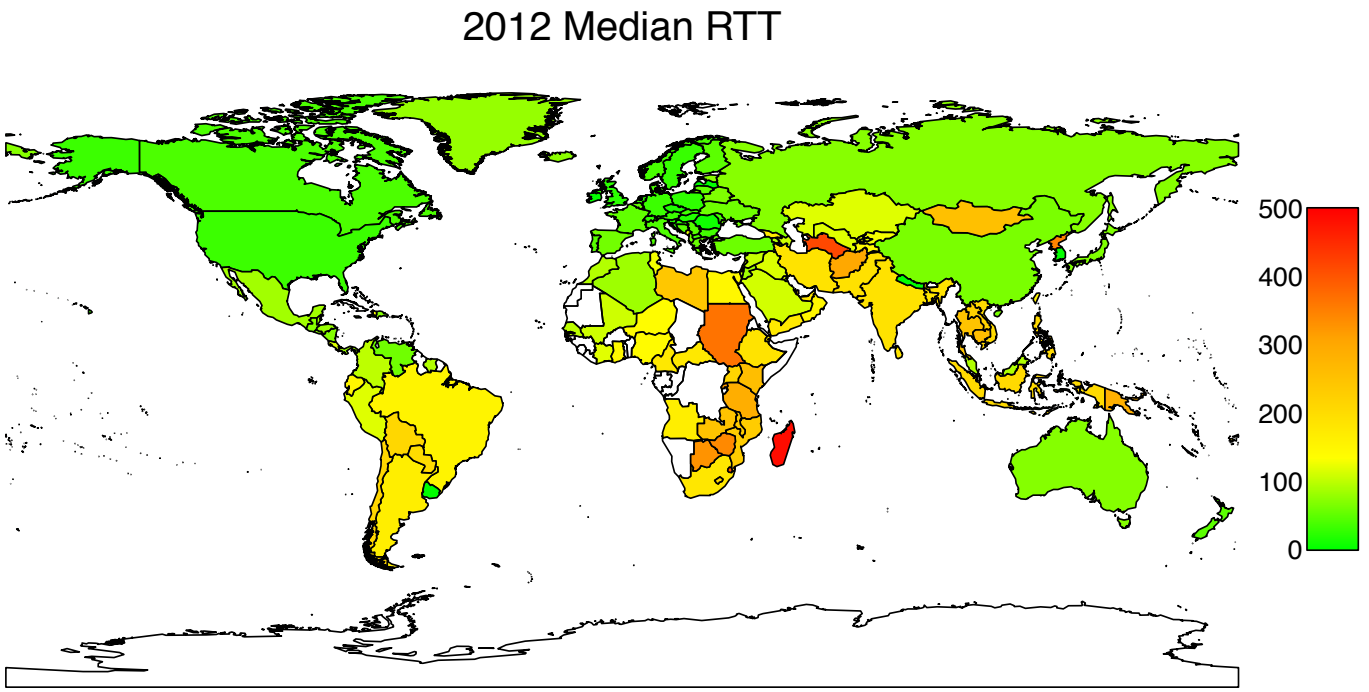
- Unfortunately lots of churn from year to year in terms of ASNs observed so difficult to compare.



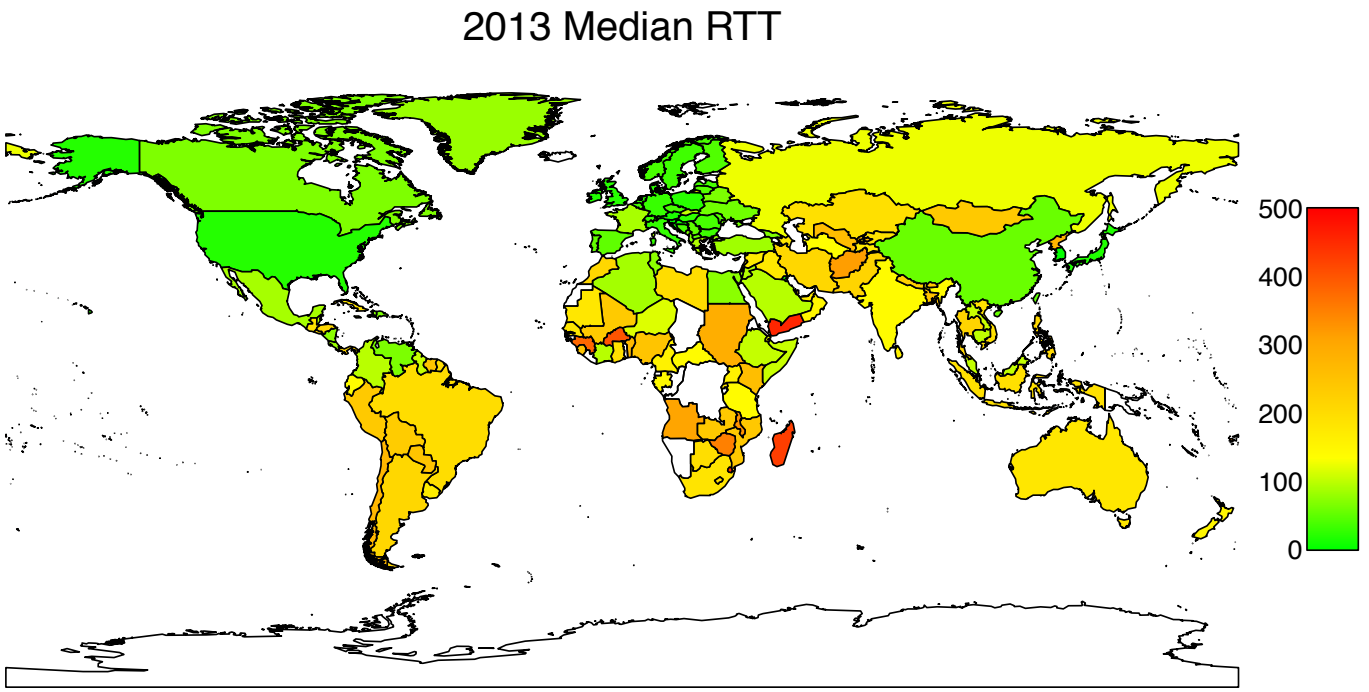
- ASN overlap between roots large enough that future comparison measurement studies possible.



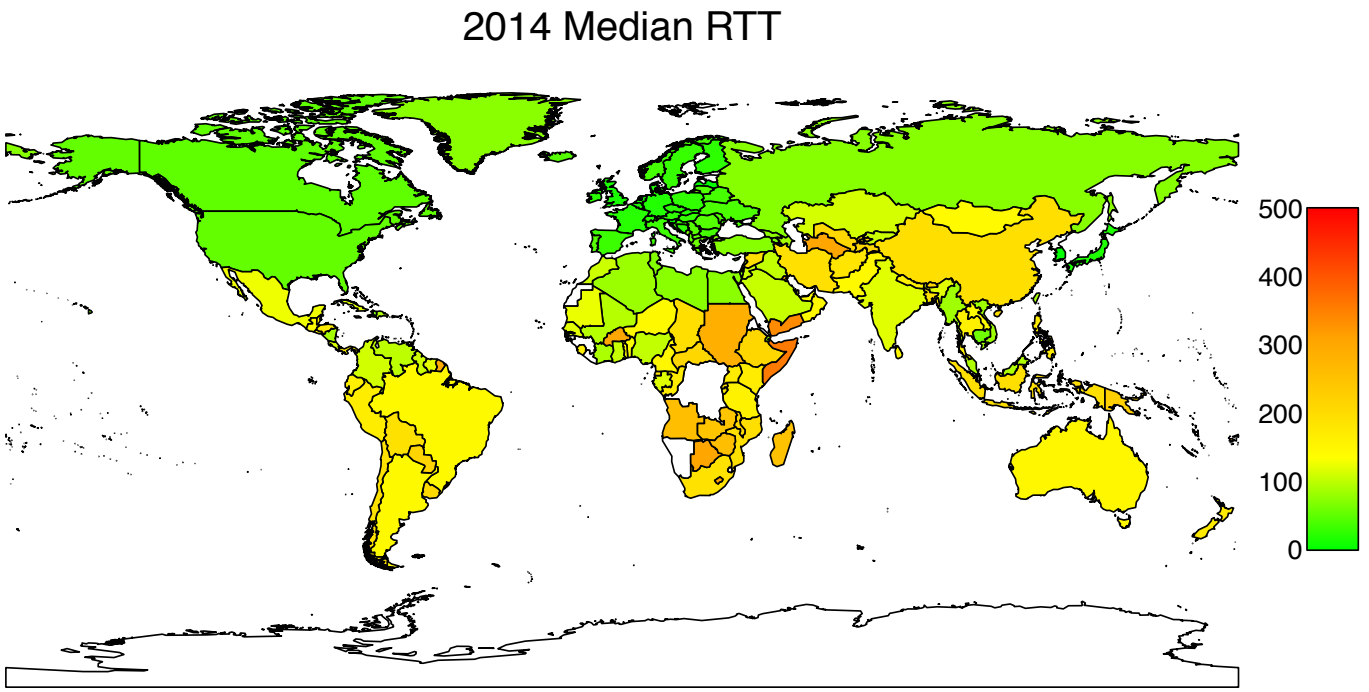
Global RTT Measurements



Global RTT Measurements



Global RTT Measurements



powered by

VERISIGN™