DNS query trends seen at Root and JP

Kazunori Fujiwara, JPRS
fujiwara@jprs.co.jp
2016/6/14, NANOG 67 DNS Track

DNS query trends

- Are DNS queries increasing ?
- How much queries does each IP address send to Root, TLD ?
- Is IPv6 deployed ?
- Is DNSSEC deployed ?

DNS-OARC's Root Datasets

- "A Day in the Life of the Internet" (DITL) is a large-scale data collection project undertaken by CAIDA and DNS-OARC every year since 2006
  - https://www.dns-oarc.net/ditl/2011/
  - 48 hours packet capture at root DNS servers
  - Source IP addresses of b and i data are anonymized

Year	Start(UTC)	End	Analyzed data from
2011	Apr 12 1200	Apr 14 1200	a c d e f h j k l m (10/13)
2012	Apr 17 1200	Apr 19 1200	a c e f h j k l m (9/13)
2013	May 28 1200	May 30 1200	a c d e f h j k l m (10/13)
2014	Apr 15 1200	Apr 17 1200	a c e f h j k m (8/13)
2015	Apr 13 1200	Apr 15 1200	a c f h j k l m (8/13)
2016	Apr 5 1200	Apr 7 1200	a c e f h j k l m (9/13)

JP datasets

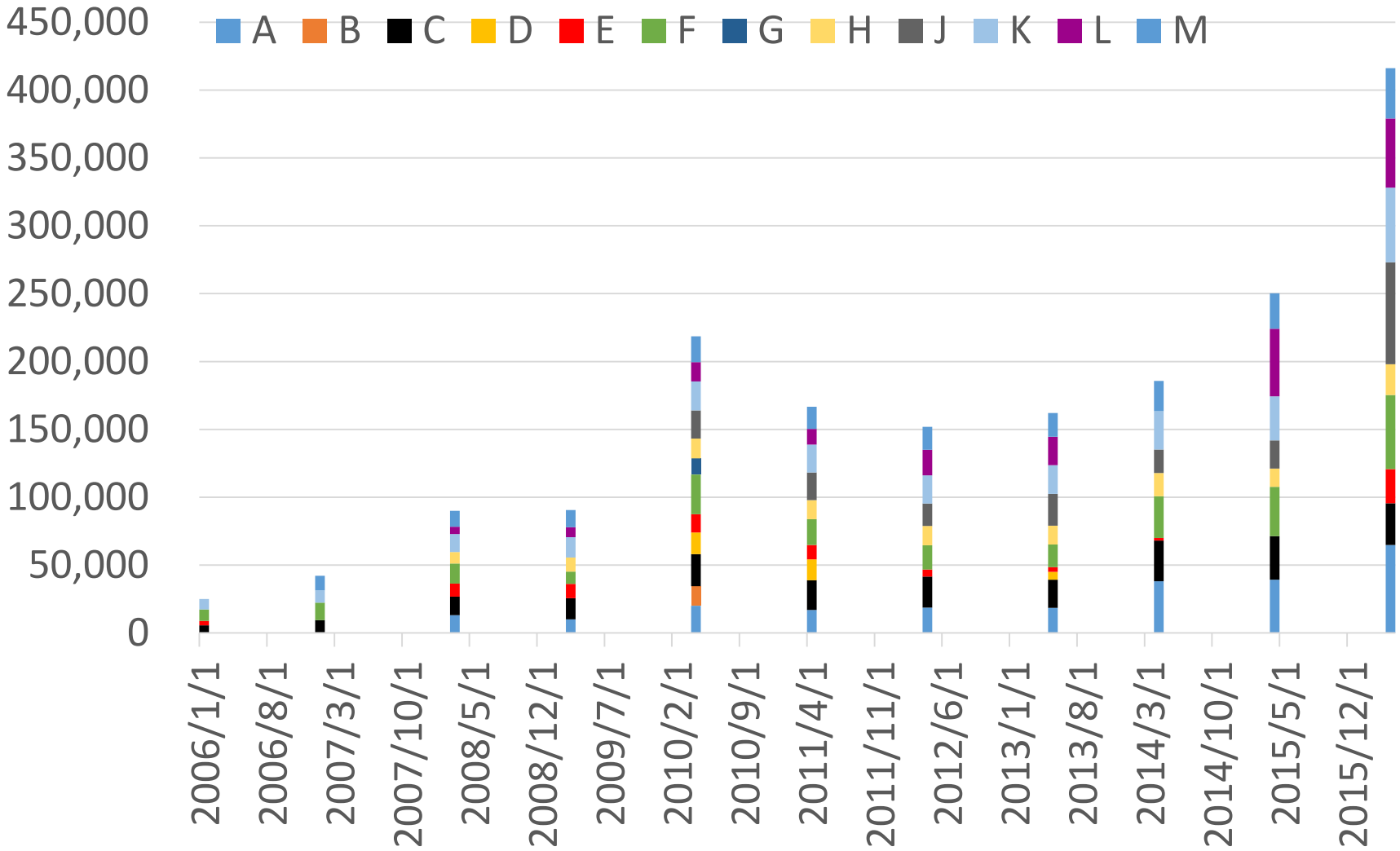
- .JP has 1,427,688 registered domain names on June 1, 2016
- JP DNS servers serve approx. 1.6 billion queries per day (2011)
- Two datasets
  - Packet captures of **all** JP DNS servers, around the same time as DNS-OARC's DITL event (and more) since 2009
  - Query logs of 2 (A and G) JP DNS servers, every day, for 13 years

Name	Operator	Location	Address (IPv4:7, IPv6:6, total 13)	Capture
A.DNS.JP	JPRS	JP*2	203.119.1.1, 2001:dc4::1	Pcap/Log
B.DNS.JP	JPNIC	JP	202.12.30.131, 2001:dc2::1	Pcap
C.DNS.JP	JPRS	See prefix	156.154.100.5, 2001:502:ad09::5	Pcap
D.DNS.JP	IIJ	JP, US, …	210.138.175.244, 2001:240::53	Pcap
E.DNS.JP	WIDE	JP, US,FR	192.50.43.53, 2001:200:c000::35	Pcap
F.DNS.JP	NII	JP	150.100.6.8, 2001:2f8:0:100::153	Pcap
G.DNS.JP	JPRS	JP	203.119.40.1	Pcap/Log

Interests of the evaluation

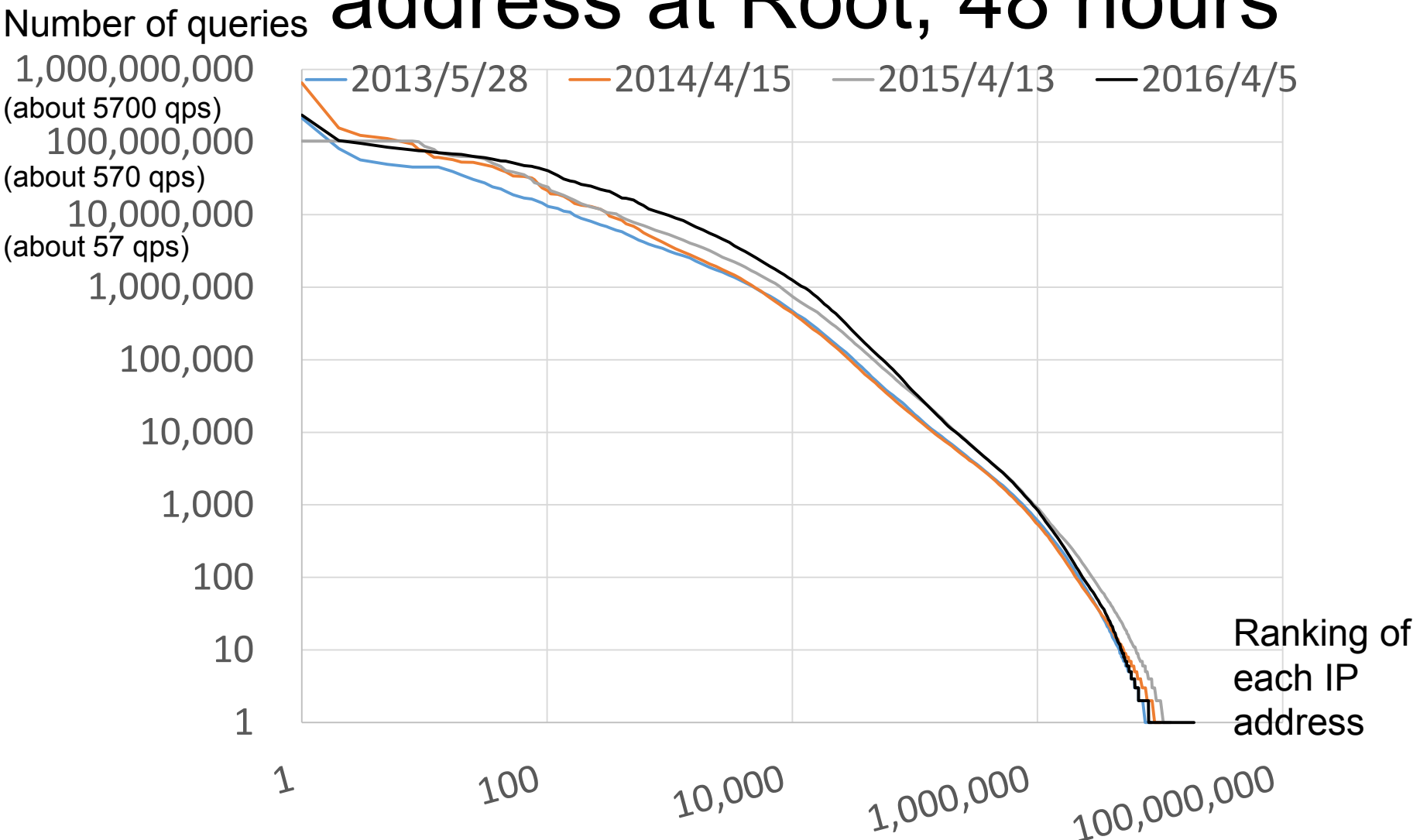
- Number of queries (average queries per second)
- Number of query source addresses
- Query ratio of IPv6, EDNS0, DNSSEC
- Ratio of IPv6 addresses
- Ratio of IP addresses that sent EDNS0, DNSSEC_OK, “.” DNSKEY
- Analyzed by countries (using Maxmind’s GeoLite Free country database)

Corrected queries at Root DITL dataset, 48 hours (average queries / second)



Increasing. However, five times increase in this 8 years. (slower than contents size)

Number of queries from each IP address at Root, 48 hours

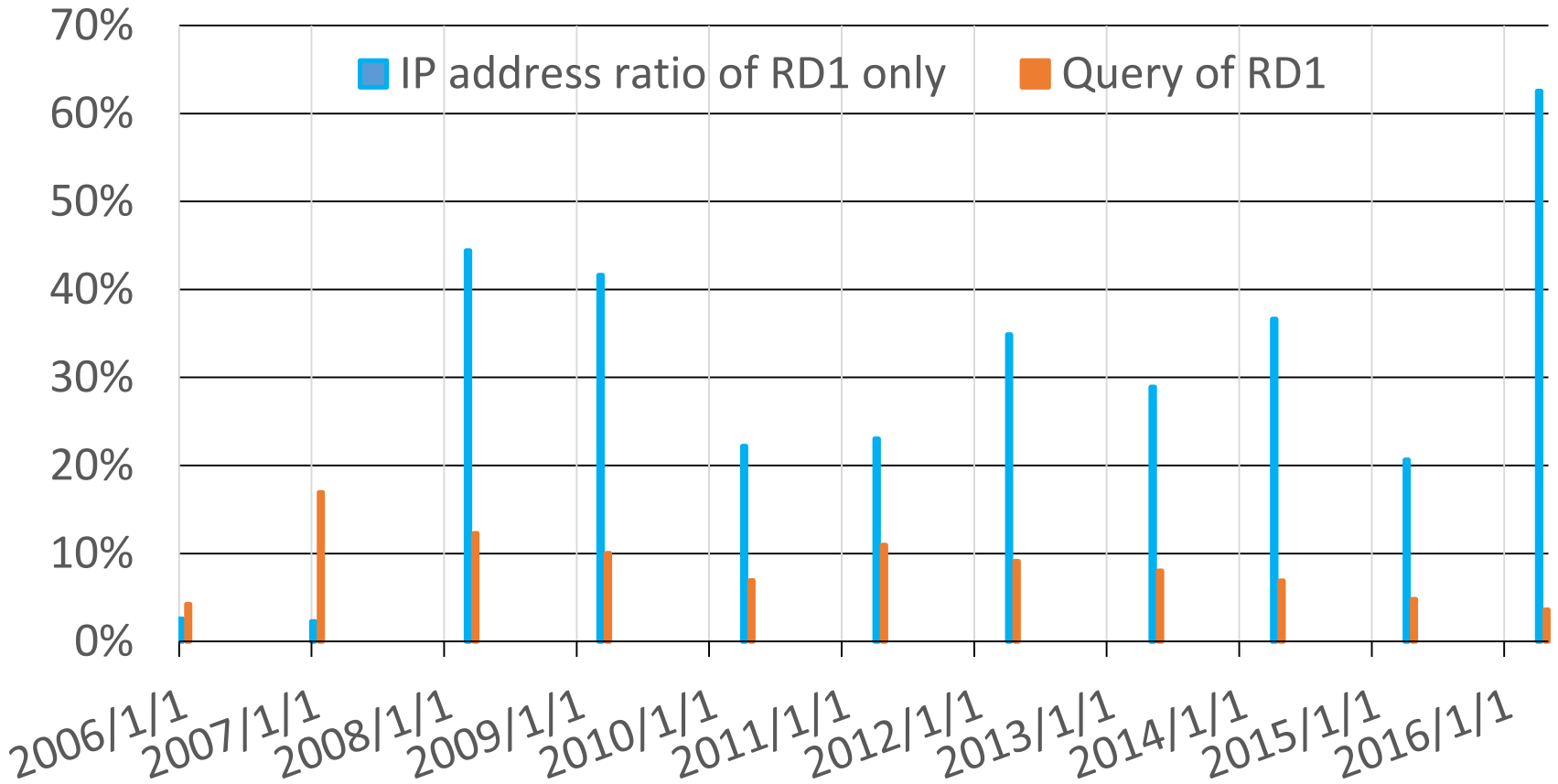


Same tendency by years. Why they send over 1,000,000 queries to Root in two days?
(Some of them comes from Typo (non-existent TLD) queries and implementation bugs.)

Strange data in 2016 at Root

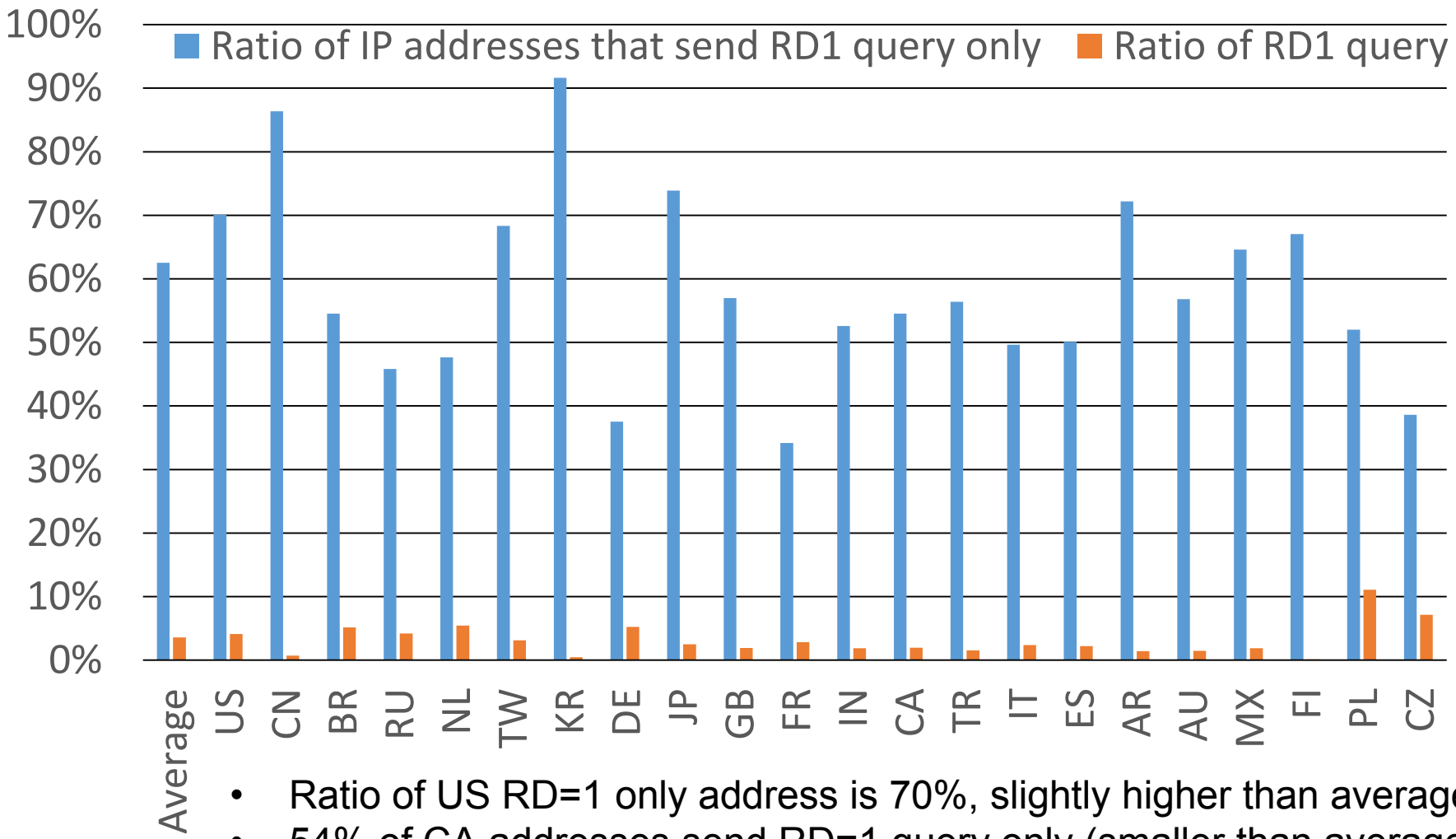
- Heavy query senders
  - No.1 address sent 235,840,481 queries /48h (1364 q/s ave.)
  - No.2 sent 104,967,673 queries/48h (607 q/s average)
    - RD=0, EDNS0, DO=1, 99% existing TLD, various query names
    - The addresses don't seem to use cache
  - No.3 sent 95,757,567 queries/48h (554 qps average)
    - RD=0, EDNS0, DO=1, [a-m].root-servers.net AAAA / A6 only
- IP addresses that send RD=1 only
  - RD bit shows Recursion Desired, set by stub resolvers
  - Recent full-service resolvers don't set RD bit
  - 62% of IP addresses send RD=1 query only in 2016.
    They may come from misconfigured hosts or bots.
  - The ratio is higher than previous: 20.6% in 2015, 36% in 2014
  - Analyzing only RD=0 queries may be useful

Ratio of RD=1 only queries at Root



- Ratio of RD=1 queries is small and relatively stable, 4% in 2016
- However, ratio of IP addresses that send RD=1 query only is not stable and very high, 62% in 2016
- RD=1 query may come from stub resolvers, or from careless bots, careless measurement for root DNS servers. These queries may not show Internet status
- Following evaluations add graphs that excluded IP addresses that send RD=1 query only
- Ratio of IP addresses that send RD=1 only is 1~2% at JP, 2014~2016, smaller than at Root

Ratio of RD=1 by country, 2016, Root



- Ratio of US RD=1 only address is 70%, slightly higher than average.
- 54% of CA addresses send RD=1 query only (smaller than average)
- 86% of CN, 91% of KR addresses send RD=1 query only.

Query source ranking by countries at Root/2016

	Country	Queries	Ratio of Qu	IP addresses	Ratio of addresses	Queries/Addresses
	ALL	7.19E+10		18,875,577		3810.2
1	US	1.69E+10	23.43%	6,183,419	32.76%	2725.1
2	CN	1.45E+10	20.18%	1,602,135	8.49%	9056.8
3	BR	3.46E+09	4.81%	563,529	2.99%	6133.8
4	RU	2.76E+09	3.84%	449,672	2.38%	6145.3
5	NL	2.62E+09	3.64%	351,136	1.86%	7455.1
6	TW	2.08E+09	2.89%	197,161	1.04%	10560.0
7	KR	2.06E+09	2.87%	423,614	2.24%	4873.8
8	DE	1.78E+09	2.47%	1,254,094	6.64%	1416.3
9	JP	1.74E+09	2.41%	886,269	4.70%	1958.1
10	GB	1.68E+09	2.33%	617,038	3.27%	2719.1
11	FR	1.53E+09	2.13%	818,889	4.34%	1866.6
12	IN	1.21E+09	1.68%	260,623	1.38%	4626.7
13	CA	1.06E+09	1.47%	404,151	2.14%	2620.5
14	TR	9.86E+08	1.37%	121,239	0.64%	8134.6

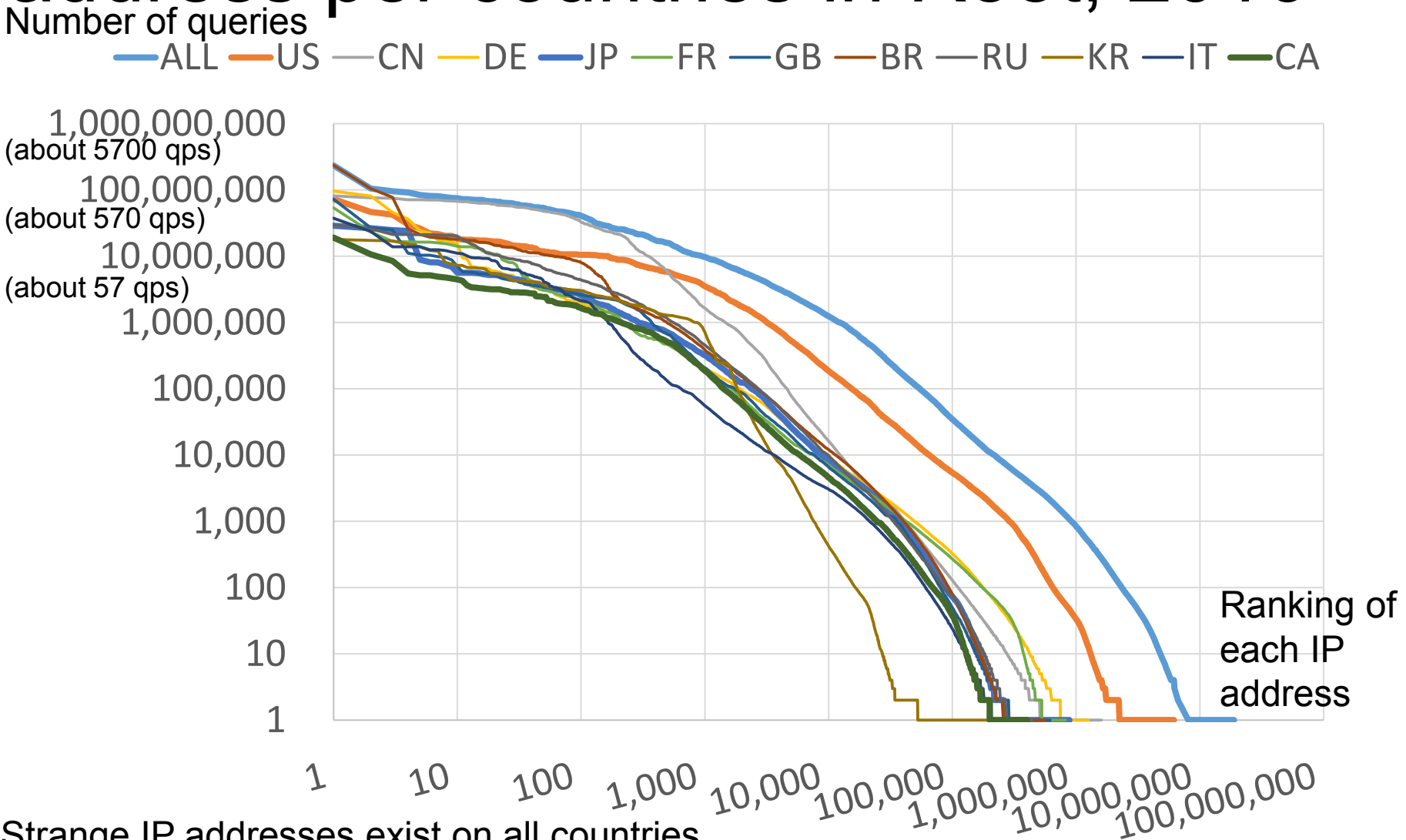
- Ratio of US IP addresses is very high (32.76%)
- US and CN addresses generate 23% and 20% of root queries
- Number of queries/ number of IP addresses is high at some countries

Query source ranking by countries at JP/2016

	Country	Ratio of Queries	IP addresses	Ratio of IP addresses
	ALL		2,019,978	
1	JP	38.79%	180,666	8.94%
2	US	25.77%	540,915	26.78%
3	TW	3.78%	25,838	1.28%
4	CN	3.10%	58,899	2.92%
5	NL	2.49%	40,977	2.03%
6	RU	2.21%	65,130	3.22%
7	DE	1.94%	174,807	8.65%
8	IT	1.50%	55,491	2.75%
9	KR	1.45%	16,096	0.80%
10	SG	1.14%	8,740	0.43%
11	GB	1.05%	64,356	3.19%
12	FR	0.97%	87,334	4.32%

- IP addresses in JP generated 38.79% of JP queries
- IP addresses in US generated 25.77% of JP queries
- US uses largest number of IP addresses to resolve

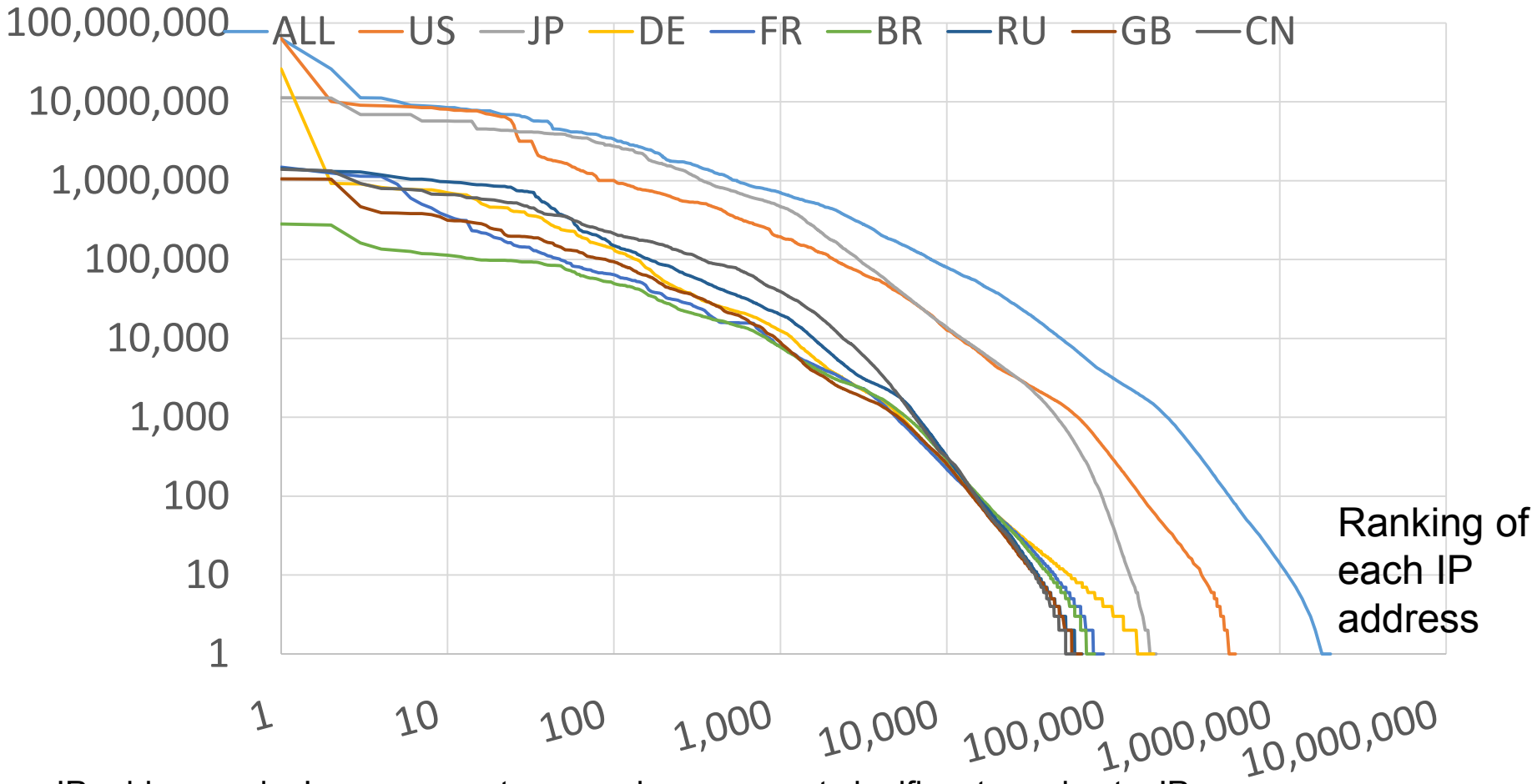
Number of queries from each IP address per countries in Root, 2016



Strange IP addresses exist on all countries

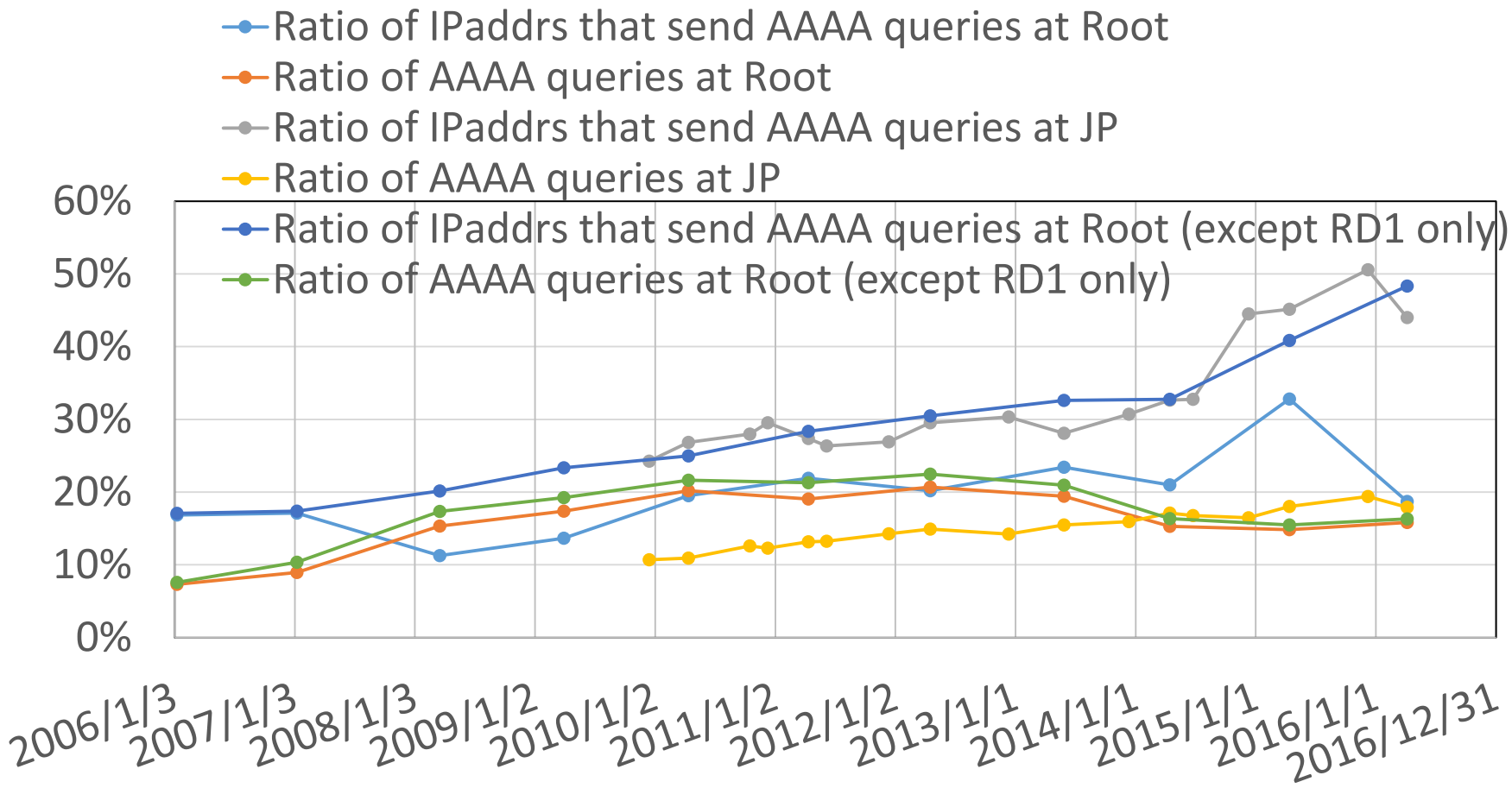
Top query generators are BR addresses.

Number of queries from each IP address per countries in JP,2016



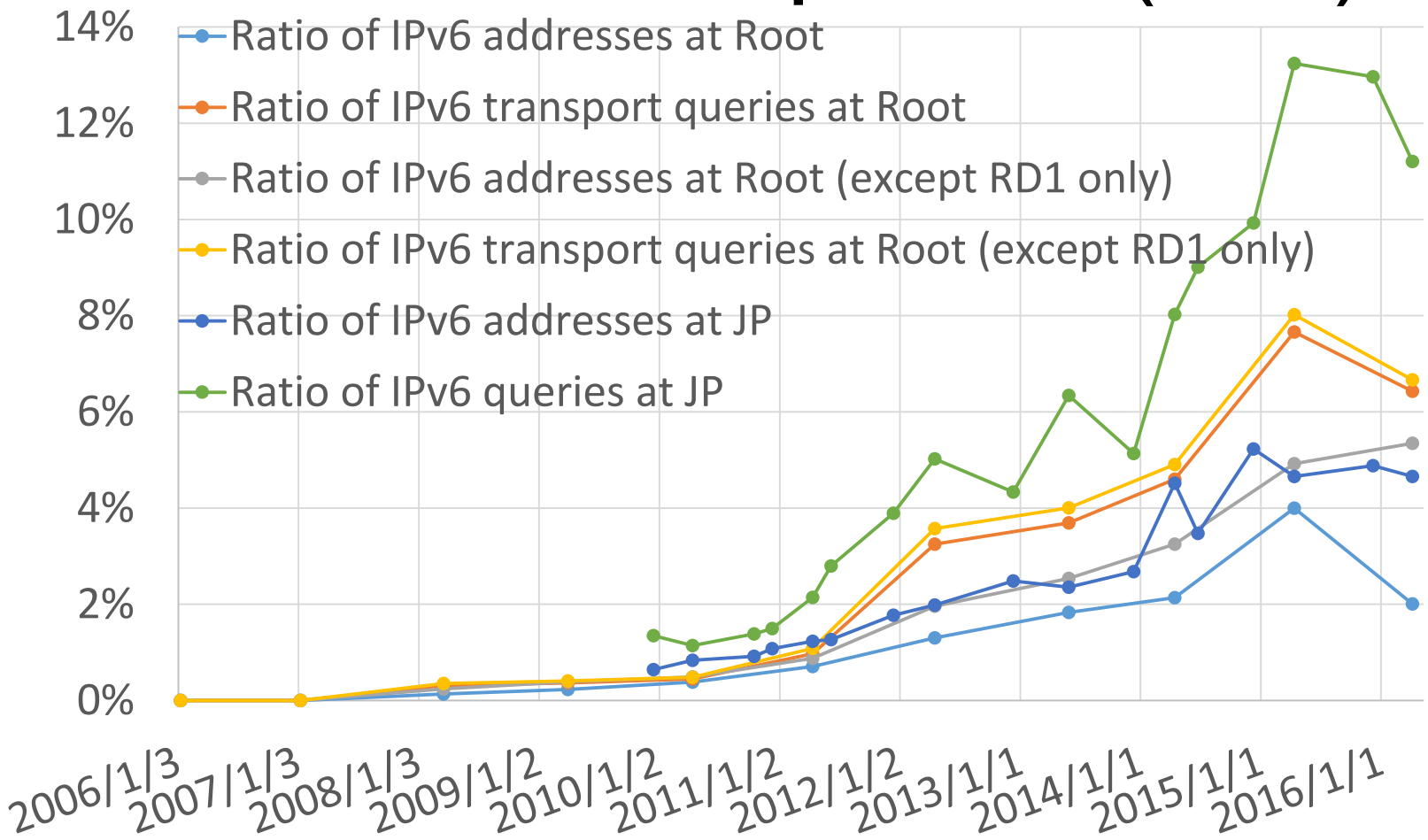
- IP addresses in Japan generate many, however, not significant queries to JP
  - JP TLD is well used from Japan
  - JP zone has about 1.4M delegations and 1 day TTL. Busy resolvers may send 2.8M queries in 2 days
- Top query senders are from US and DE: 10M or more are too large numbers

Ratio of AAAA queries



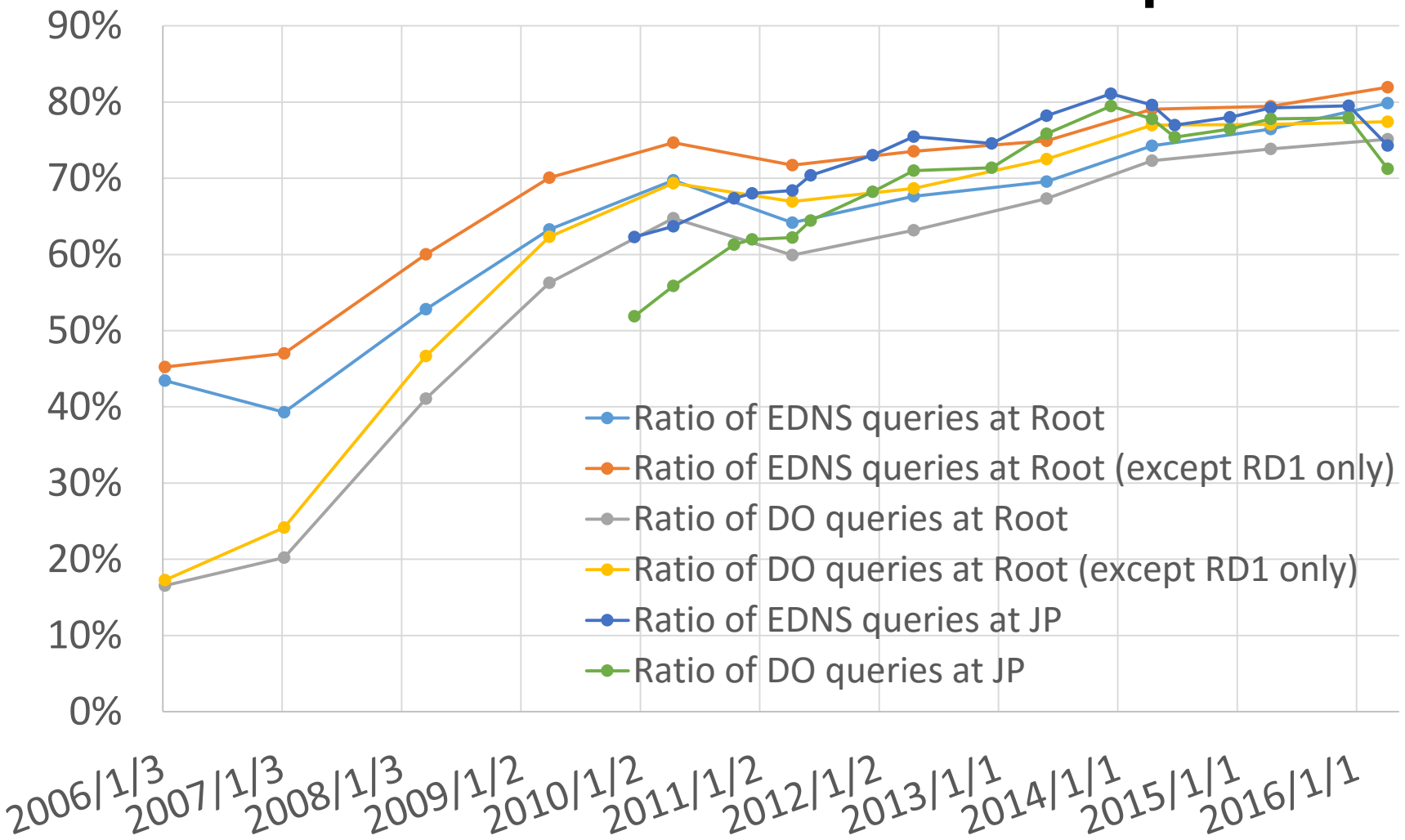
- IPv6 query type (AAAA) is well used and almost stable: about 17%
- Ratio of IP addresses that send AAAA queries is large: about 48% at root, 43% at JP
- The reason of the increase is that most of client applications and OSes send both A and AAAA queries
- These AAAA queries may be hidden by RFC 7816 query name / type minimization
- Excluding addresses that send RD=1 queries only hide irregular decrease of Ratio of IP addresses that send AAAA queries at root (light blue line → deep blue line)

Ratio of IPv6 queries (48h)



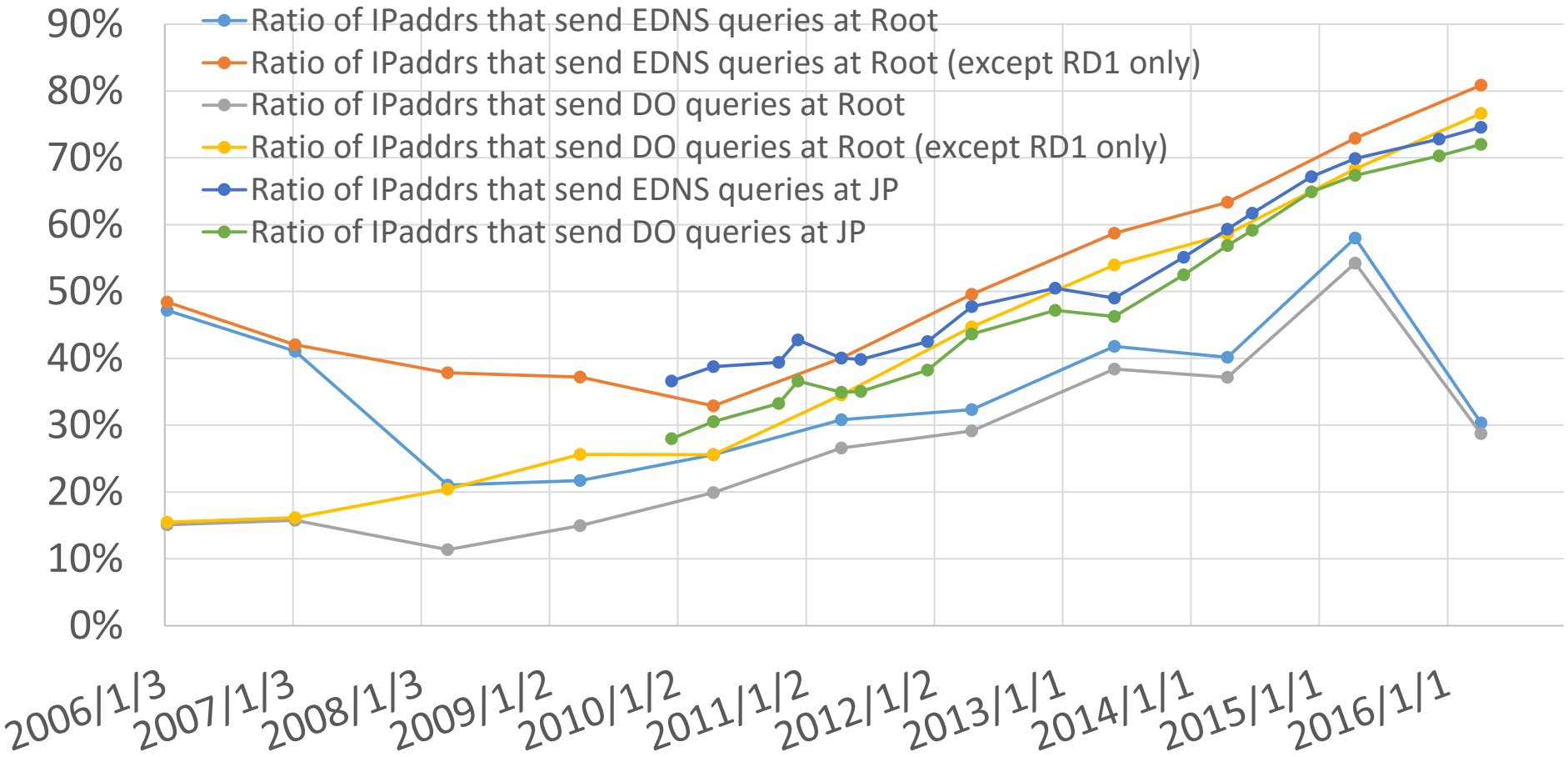
- IPv6 transport is well used: about 7% at root, 12% at JP
- Ratio of IPv6 addresses has been increased: 5% at both Root and JP
- Excluding addresses that send RD=1 queries only hide irregular decrease of Ratio of IPv6 addresses at root (light blue line)
- IPv6 usage may be still increasing, or stable.

Ratio of EDNS/DO(dnssec_ok) queries



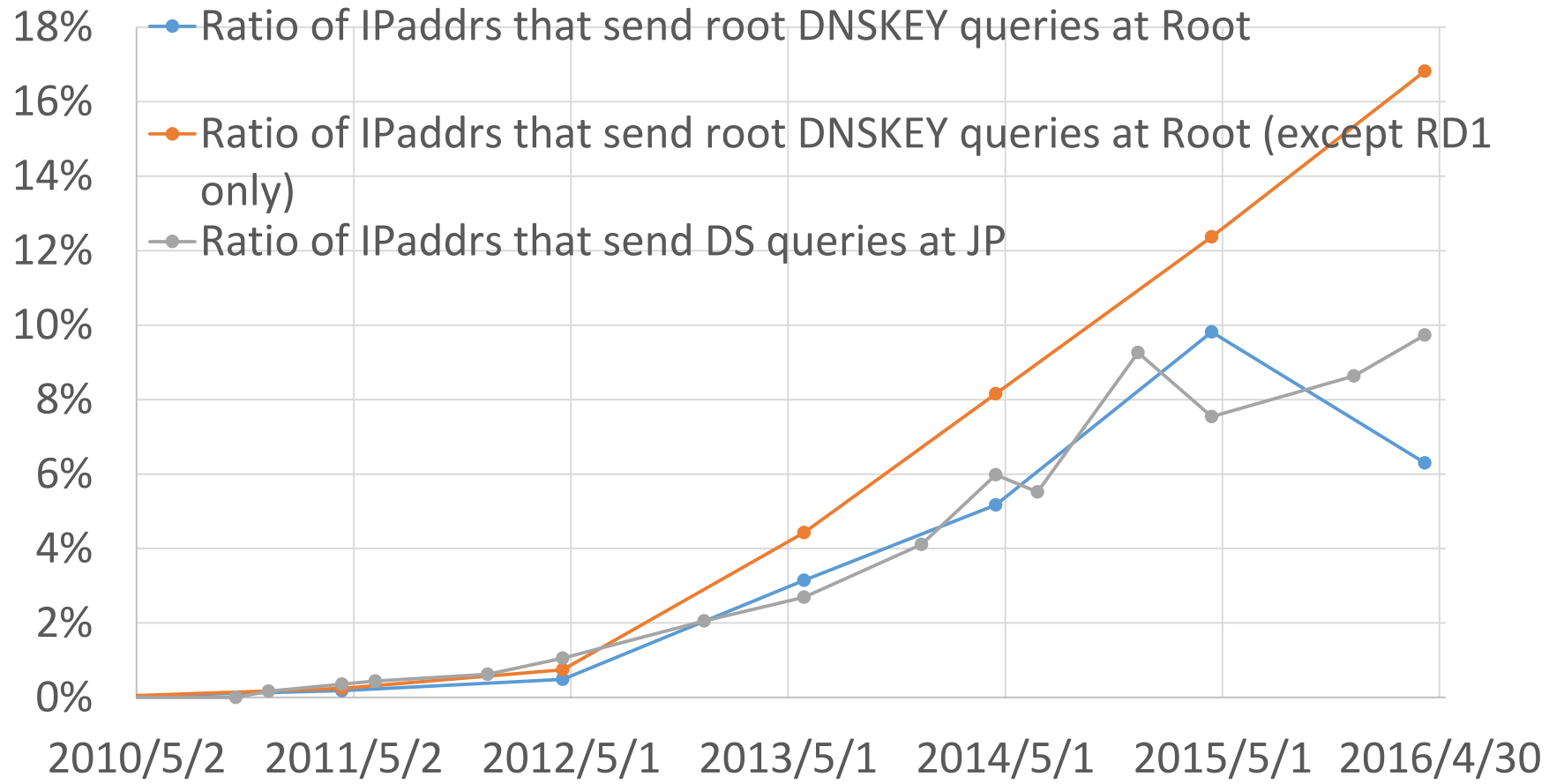
About 70~80% of all queries have EDNS0 and DO bit.
The query sources know DNSSEC (at least, software support DNSSEC)

Ratio of EDNS/DO IP addresses



DNSSEC protocol support in full-resolvers is well deployed
70% of IP addresses at both Root and JP
Excluding addresses that send RD=1 query only hides
irregular decrease of IP address ratio (light blue line and dark line)

DNSSEC validation at Root & JP



- DNSSEC validators send root DNSKEY query every day
  - Blue line shows ratio of IP addresses which send root DNSKEY queries: 6% or 9%
  - Red line shows ratio of IP addresses which send root DNSKEY queries except RD=1 only addresses: 16%
- DNSSEC validators send each jp domain name DS every 900/7200 second to JP
  - Dark line shows ratio of IP addresses which send domainname.jp DS queries: 9.7%
- Ratio of probable DNSSEC validators may be 9.7% or 16% now

Findings

- Queries to root DNS servers are increasing
  - Five times increased in this 8 years. (slower than contents size)
- Some (many) addresses generate strange many queries
  - Some hosts send over 100,000,000 queries/2days(570 qps)
  - 100,000 addrs send over 1,000,000 queries/2days (very large)

- IPv6 query type (AAAA) is well used, almost stable: (17%)
- IPv6 transport is well used: 7% at root, 12% at JP
- Ratio of IPv6 addresses is increased: 5%(Root, JP)
- Ratio of probable DNSSEC validators may be 9.7% or 16%

- DITL dataset may contain irregular data. We need to analyze them
  - Ignoring RD=1 queries is one effective idea to decrease effects of irregular (non-full-service-resolver) queries

Future analysis

- Analysis of strange IP addresses
  - To find the reason of unnecessary queries
- EDNS0 options (NSID, SIT/COOKIE, Client Subnet)
- DNSSEC, IPv6 deployment by countries
- IDN, new gTLD queries by countries

Acknowledgements

- DNS-OARC as the data source of Root dataset