DNS traffic analysis
-- Issues of IPv6 and CDN --

Kazunori Fujiwara”^, Akira Sato”, Kenichi Yoshida”

”University of Tsukuba

^Japan Registry Services Co., Ltd (JPRS)

July 29, 2012 IEPG meeting at Vancouver

This material comes from

K. Fujiwara, A. Sato”, K. Yoshida,

“DNS traffic analysis -- Issues of IPv6 and CDN --”, Proceedings of 2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet (SAINT- 2012), pp.129-137, July 2012.

Outline

- Background
- Related works and recent issues
- DNS Traffic data at University of Tsukuba
- Analysis results
  - Increase of AAAA queries
  - Cache hit rate and effect to authoritative DNS servers
- Conclusion

Background: DNS and Internet

- The Internet has become a critical infrastructure
- Domain Name System (DNS)
  - a key naming system of the Internet
  - Bridge domain names & IP addresses
- DNS load is increasing

Related works and recent issues

- Related works
  - In 2002, Jaeyeon et al. reported on DNS performance and the effectiveness of caching
- Recent Issues from 2002
  - IPv6 aware clients has been spread
    - They send both A (IPv4) and AAAA (IPv6) queries
  - Content Delivery Networks (CDNs) use complicated DNS configurations

CDN and Web service issues

1. Cache: DNS answers are cached for next query at Full-resolvers
   - TTL: Cacheable time (in second)
   - Root, TLDs set 86400(1day) or 172800(2days)
   - CDN and large web services use small TTL value to control their traffic (20, 30, 60, 300)

2. CNAME: Alias mechanism in DNS
   - Full resolvers need to restart name resolution process from root when it confronts CNAME

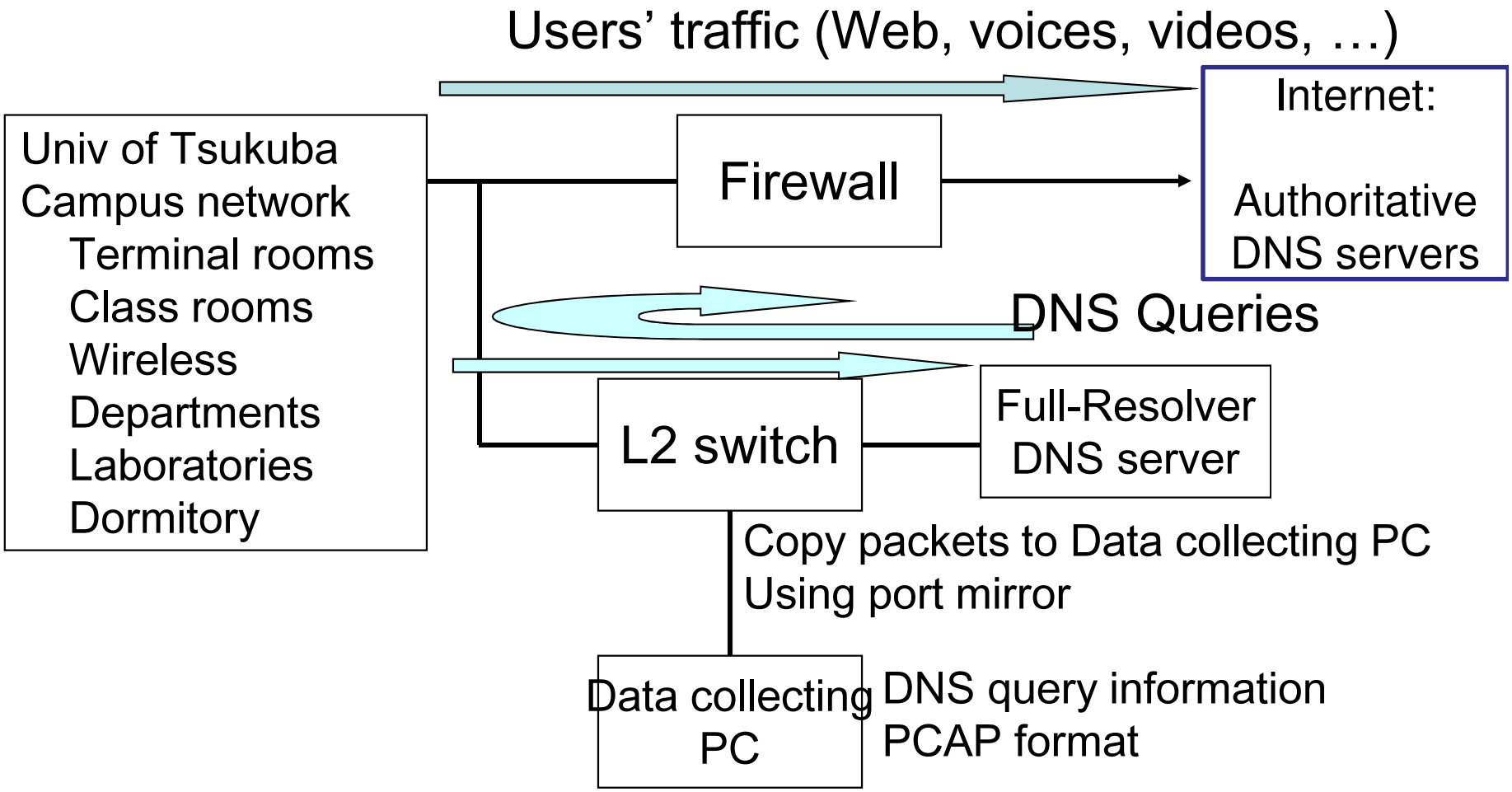
3. Use of out-of-bailiwick DNS server name
   - DNS server name resolution

DNS Traffic data at Univ. of Tsukuba

- Compared 2002 Jaeyeon's data with recent Tsukuba's data
  - Lookup rates are 20 times larger than MIT 10 years ago
  - Query names are 12 times increased
  - Number of clients are similar

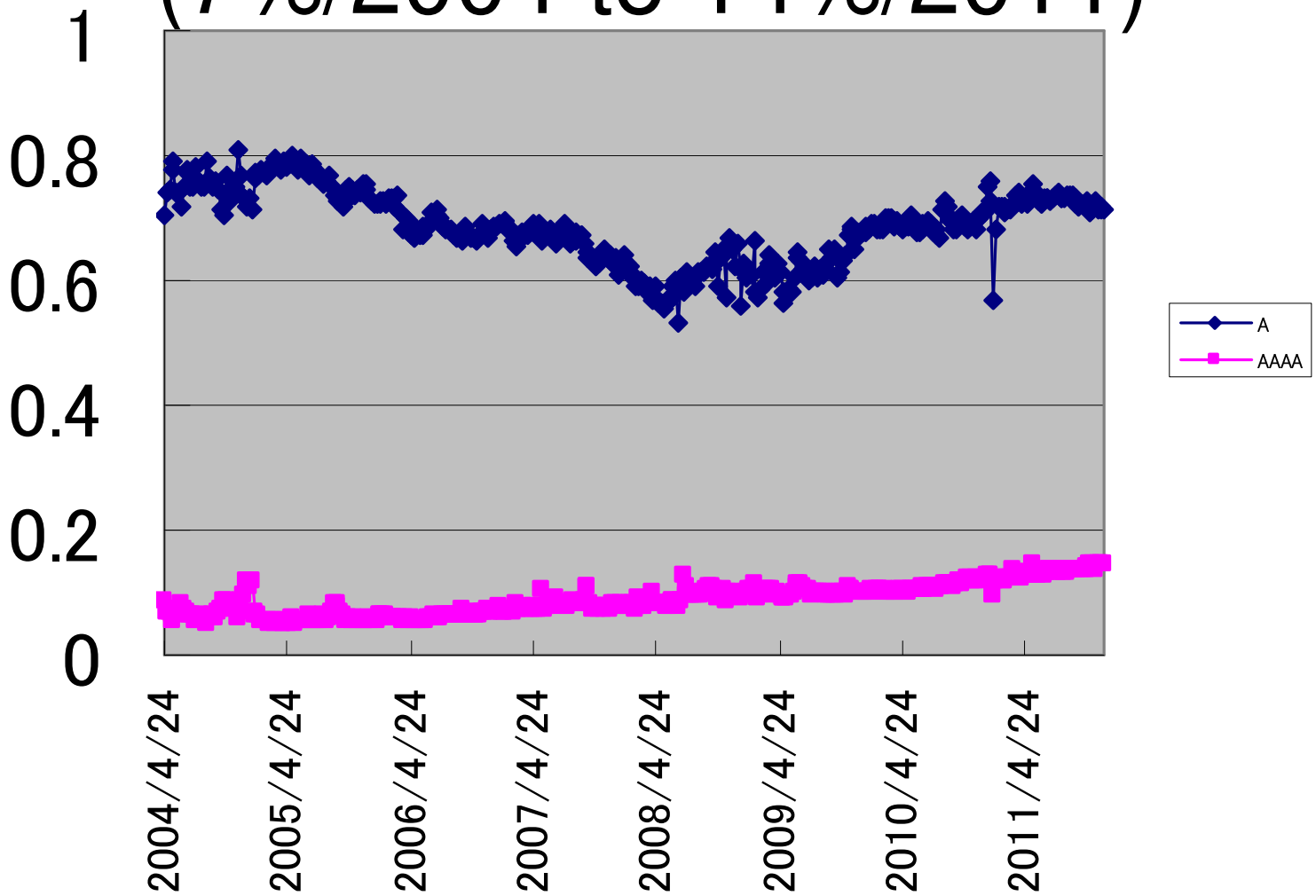
Origin	jung2002dns	jung2002dns	Author	Author
Date (from-to)	2000/12/04-11	2001/05/18-24	2010/11/1-30	2011/11/1-30
Place	MIT	KAIST	Tsukuba	Tsukuba
lookups	4,160,954	4,339,473	234,308,393	366,489,499
Queries/sec	6.88	8.37	90.40	141.39
query names	302,032	219,144	3,375,088	4,015,966
clients	1,216	8,605	6556	8815
query/sec/clients			0.01378	0.01603

DNS Data collection at Univ. of Tsukuba



Academic Computing & Communication Center offers Full-Resolver DNS servers for Campus network of Univ. of Tsukuba

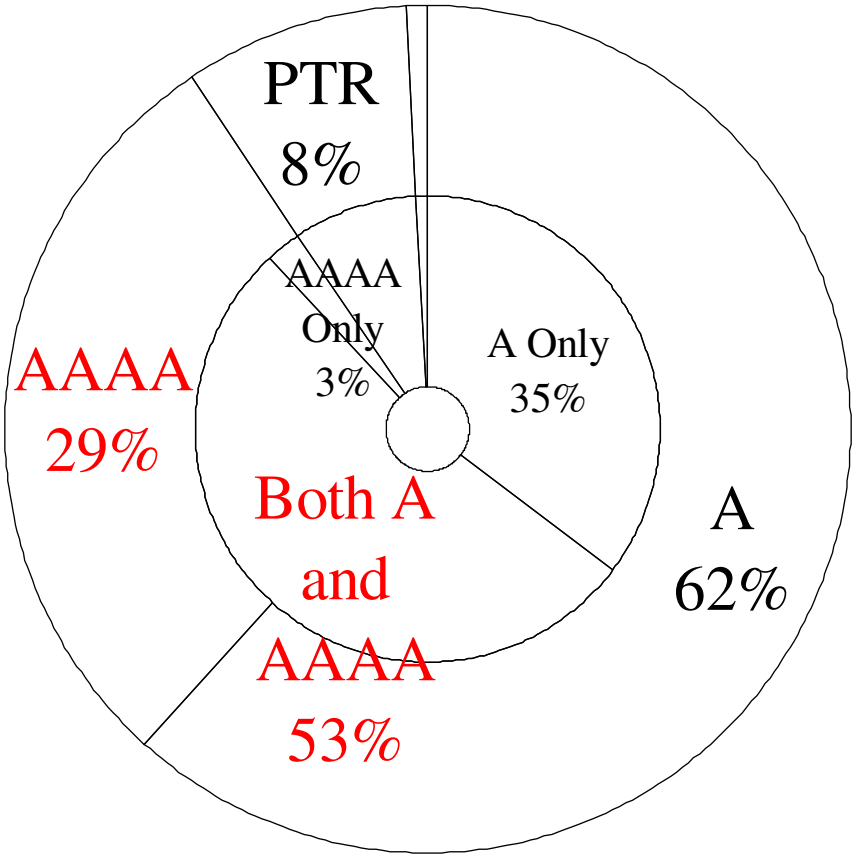
Query types seen at A.DNS.JP
AAAA queries are increasing slightly (7%/2004 to 14%/2011)



Increase of AAAA queries
Query types from clients

- Recently, 29% of queries are IPv6 address (AAAA) queries

  even if they don't have IPv6 connectivity



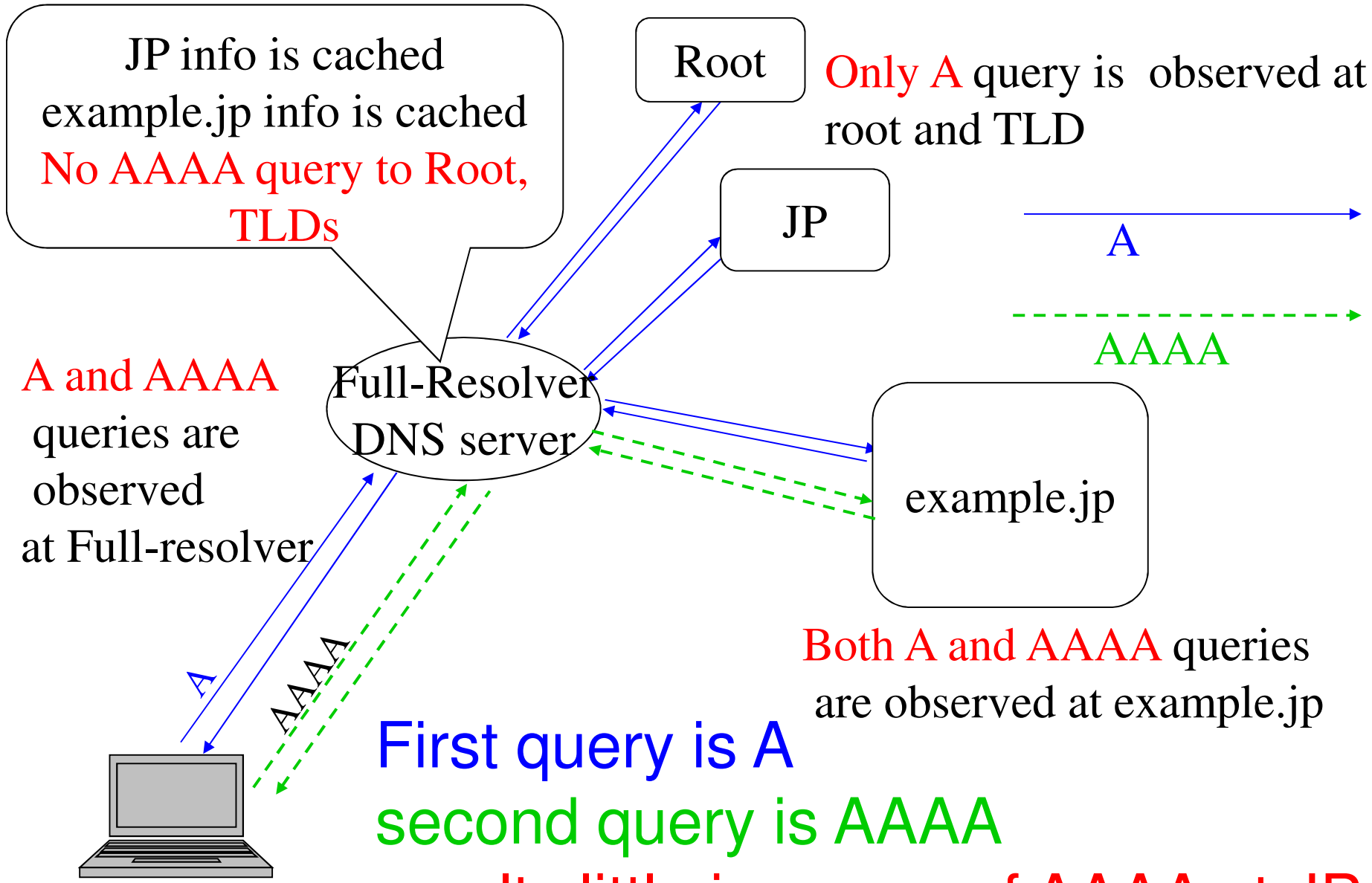
Query types observed by Full-resolver, Nov.2011

Num. of clients which send A & AAAA

Case		Hosts	[%]
Total Hosts		8815	100
Hosts, send A queries		8707	98.8
Hosts, send AAAA queries		7772	88.2
Hosts, send both A and AAAA		7720	87.6
	send A before AAAA	6993	79.3
	send AAAA before A	100	1.1
	mix (undetermined)	627	7.1

- 88.2% of clients send AAAA queries even if they don't have IPv6 connectivity
- 79.3% of clients send A queries before AAAA queries.

79.3% of clients send A before AAAA



Answer types for clients
92.6% of AAAA answers: NO DATA

Query Type	Total	A	AAAA	PTR	Other
Number of queries	366.5	225.7	106.6	31.1	3
[million]	100%	62%	29%	8%	1%
Answer Type	%	%	%	%	%
Server Failure	2.9	0.6	1.8	23.3	1.3
Name Error	11.5	8.1	3.4	59.1	67.7
Refused	0	0	0	0	0.3
Normal Answer	57.2	89.3	2.0	17.4	17.1
No Data	27.5	0.7	92.6	0.1	12.2
Timeout	0.9	1.3	0.2	0.1	1.4

Evaluation Indexes

1. Cache hit rate
   - number of cache hit stub query / number of all stub query
   - Cache hit = no queries for authoritative

2. How many queries sent to authoritative DNS servers does one stub query generates?
   - Number of authoritative DNS server queries / Number of stub queries
   - Root: root DNS server queries / Stub queries
   - TLD: TLD DNS server queries / Stub queries
   - All Auth: Authoritative queries / Stub queries
3. Latency to clients

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
with CNAME	53.9	73.8	0.0007	0.025	0.34	30.4
without CNAME	43.2	76.7	0.00096	0.028	0.26	24.9
Normal Answer	57.2	72.3	0.00071	0.039	0.37	31.4
Error	42.8	78.9	0.0009	0.007	0.22	23.4
ServerFailure	2.9	77.8	0.00007	0.002	0.54	89.5
NoData	27.3	75.4	0.00006	0.007	0.24	23.4
NameError	11.5	93.5	0.00315	0.008	0.06	5.3
NameError (woTLDerror)	11.1	93.5	0.00008	0.008	0.06	5.3
TLDerror	0.5	92.6	0.07303	0	0.07	5.4
TTL<=300	44	67.3	0.00033	0.021	0.4	31.7
TTL>300	56	81.3	0.00115	0.029	0.23	25.1

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
Error	42.8	78.9	0.0009	0.007	0.22	23.4
ServerFailure	2.9	77.8	0.00007	0.002	0.54	89.5
NoData	27.3	75.4	0.00006	0.007	0.24	23.4
NameError	11.5	93.5	0.00315	0.008	0.06	5.3
NameError (woTLDerror)	11.1	93.5	0.00008	0.008	0.06	5.3
TLDerror	0.5	92.6	0.07303	0	0.07	5.4
TTL<=300	44	67.3	0.00033	0.021	0.4	31.7
TTL>300	56	81.3	0.00115	0.029	0.23	25.1

The average of all data

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
with CNAME	53.9	73.8	0.0007	0.025	0.34	30.4
without CNAME	43.2	76.7	0.00096	0.028	0.26	24.9
NameError	11.5	93.5	0.00315	0.008	0.06	5.3
NameError (woTLDerror)	11.1	93.5	0.00008	0.008	0.06	5.3
TLDerror	0.5	92.6	0.07303	0	0.07	5.4
TTL<=300	44	67.3	0.00033	0.021	0.4	31.7
TTL>300	56	81.3	0.00115	0.029	0.23	25.1

Use of CNAME makes response slow (5.5 ms)
increases auth queries 130% (0.34/0.26)
cache hit rate low (3%)

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
with CNAME	53.9	73.8	0.0007	0.025	0.34	30.4
without CNAME	43.2	76.7	0.00096	0.028	0.26	24.9
Normal Answer	57.2	72.3	0.00071	0.039	0.37	31.4
Error	42.8	78.9	0.0009	0.007	0.22	23.4
ServerFailure	2.9	77.8	0.00007	0.002	0.54	89.5
NoData	27.3	75.4	0.00006	0.007	0.24	23.4

A lot of Errors, but cache hit rate and latency are not bad
Especially, No Data caused by AAAA queries.
Well cached, low latency

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
with CNAME	53.9	73.8	0.0007	0.025	0.34	30.4
without CNAME	43.2	76.7	0.00096	0.028	0.26	24.9
Normal Answer	57.2	72.3	0.00071	0.039	0.37	31.4
Error	42.8	78.9	0.0009	0.007	0.22	23.4
ServerFailure	2.9	77.8	0.00007	0.002	0.54	89.5
NoData	27.3	75.4	0.00006	0.007	0.24	23.4
NameError	11.5	93.5	0.00315	0.008	0.06	5.3
NameError (woTLDerror)	11.1	93.5	0.00008	0.008	0.06	5.3
TLDerror	0.5	92.6	0.07303	0	0.07	5.4

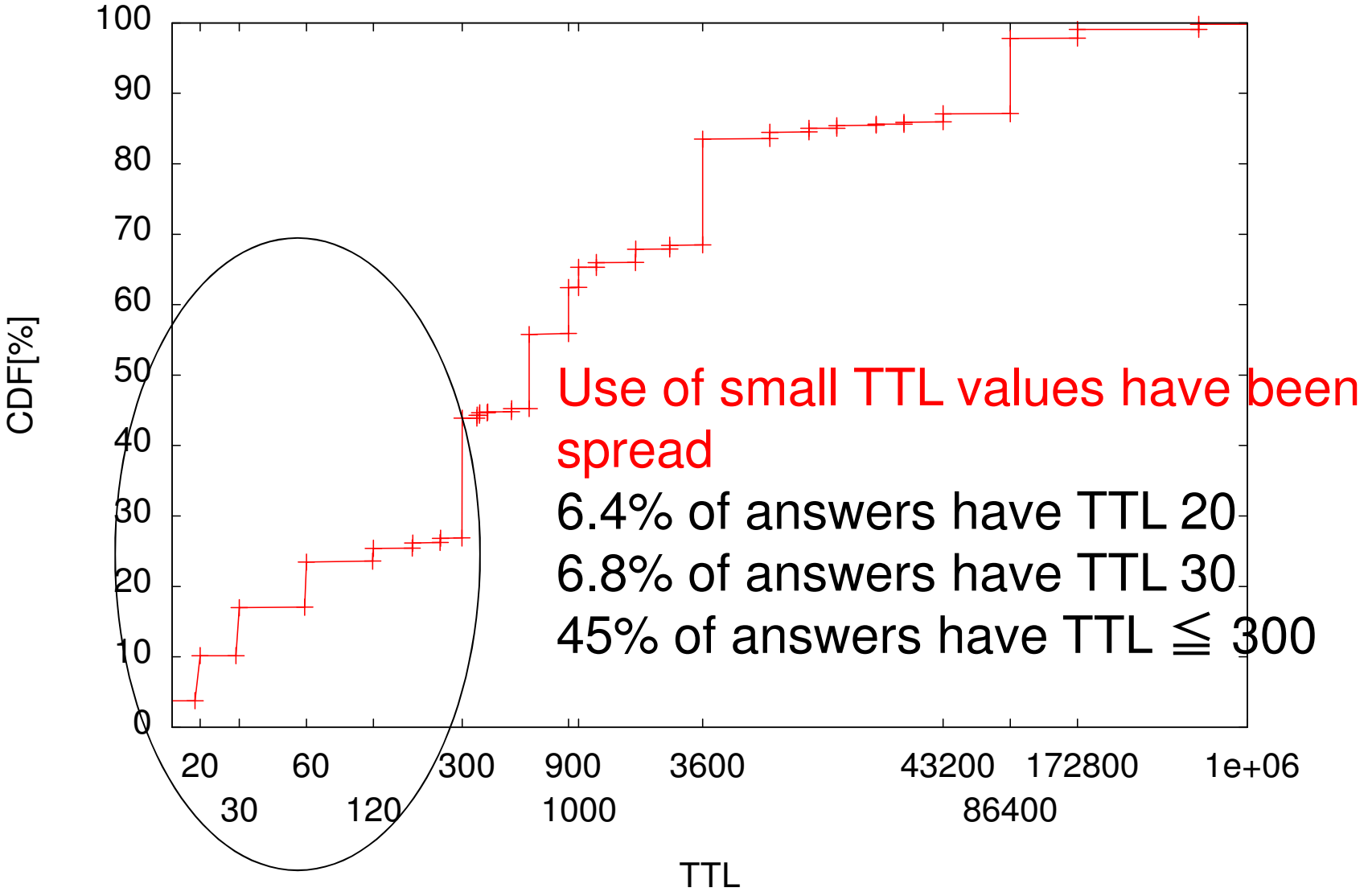
45% of root DNS server queries are caused by TLD error which specifies non-existing TLDs. (Typos?)

Cache hit rate and effect to auth.

Case	Ratio in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub			Latency (ave) [ms]
			Root	TLDs	All	
ALL	100	75.1	0.00079	0.025	0.31	28
with CNAME	53.9	73.8	0.0007	0.025	0.34	30.4
without CNAME	43.2	76.7	0.00096	0.028	0.26	24.9
Normal Answer	57.2	72.3	0.00071	0.039	0.37	31.4
Error	42.8	78.9	0.0009	0.007	0.22	23.4
ServerFailure	2.9	77.8	0.00007	0.002	0.54	89.5
TTL<=300	44	67.3	0.00033	0.021	0.4	31.7
TTL>300	56	81.3	0.00115	0.029	0.23	25.1

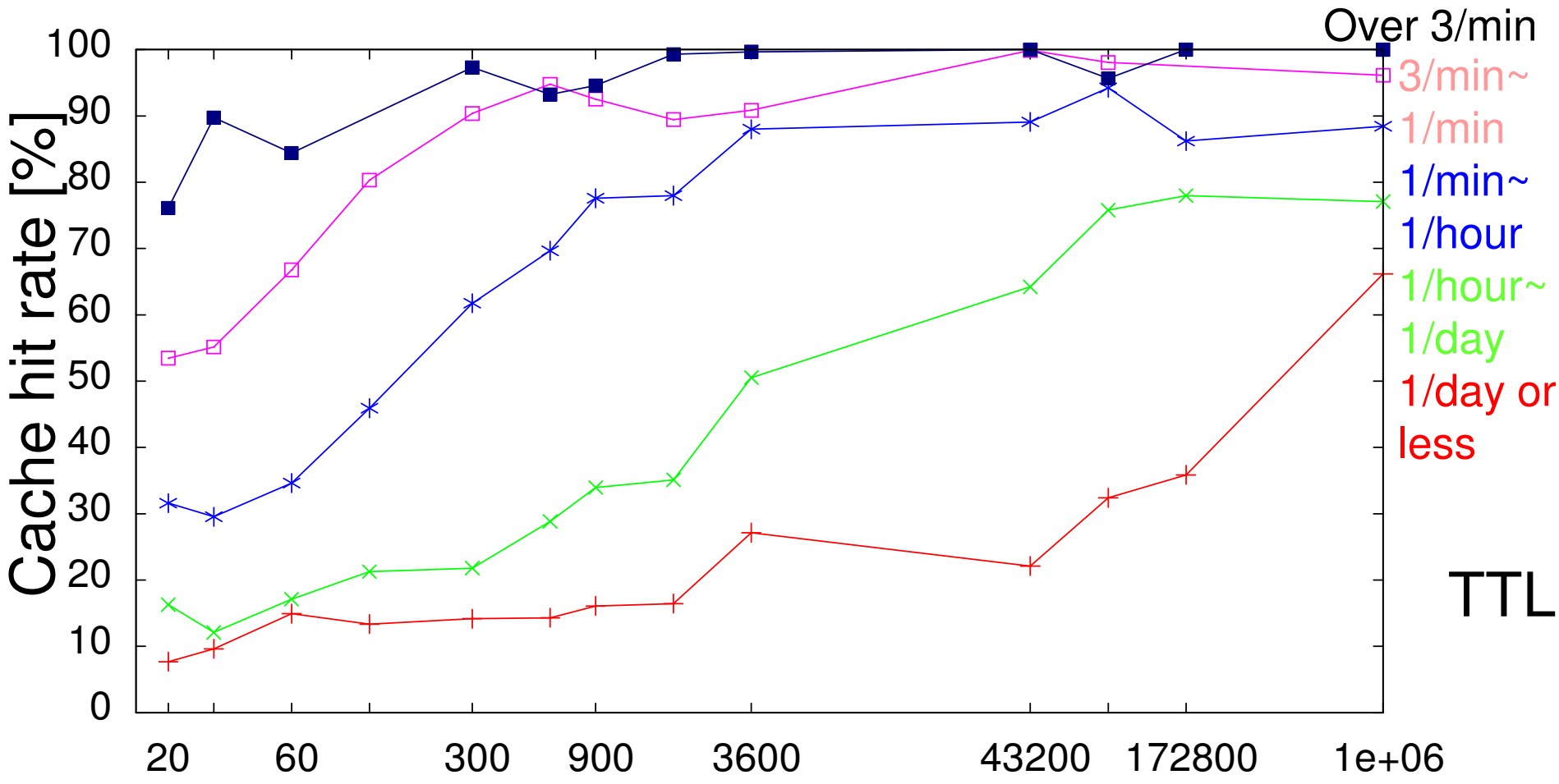
Use of small TTLs makes cache hit rate low
increases queries to authoritative DNS servers
response slow

cumulative distribution of TTL values



The TTL value for each query is chosen as the minimal TTL value for the query

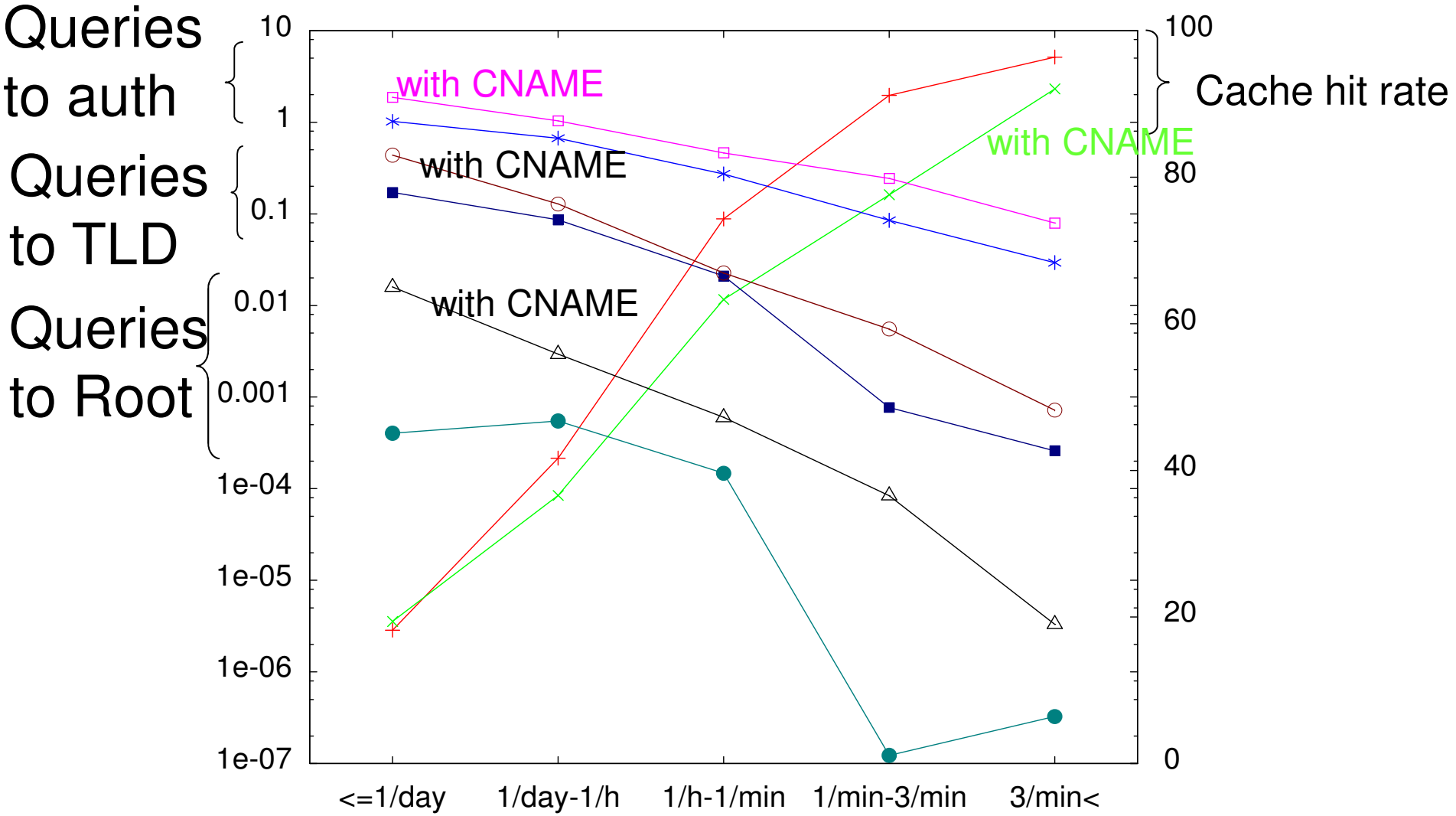
Cache hit rate at given average frequency & TTL



Short TTLs does not mean low cache hit rate.
If frequent, small TTL is OK. But not, use longer TTL.

Frequency: number of queries / period (1 month) for each qname & qtype

Authoritative queries at given query frequency & CNAME



The use of CNAMEs lowers cache hit rate about 10%

increases queries to root & TLD DNS servers

Cache hit rate and use of out-of-bailiwick names
Use of out-of-bailiwick name slow down responses

Case	Rate in the whole [%]	Cache hit rate [%]	Authoritative queries devided by Stub queries			Latency to clients [ms]
			Root	TLDs	All	
Total/Average	100	75.1	0.0008	0.025	0.31	28
All servers are In-bailiwick	27.5	78.1	0.0004	0.009	0.24	22.5
All servers are Out-of-bailiwick (except.arpa)	60.1	71.5	0.0004	0.032	0.36	33.1
Some servers are Out-of-bailiwick	3.7	69.4	0.0012	0.039	0.4	28.6
.arpa	7.3	95.7	0.0004	0.006	0.07	6.1
Undetermined	1.5	81.2	0.0256	0.116	0.29	31.5

We observed

- 29% of queries are AAAA（IPv6) queries
  even if clients do not have IPv6 connectivity
  - Most of answers are No DATA, but well cached
- CDNs and web services use low TTL values
  - It makes DNS cache hit rate low
  - High-frequency accesses hides this phenomenon.
  - TTLs<31 is 14.0%, it creates significant DNS load.
- 45% of queries to the root DNS servers came from wrong name queries, 0.5% of all queries.
- 60% of DNS traffic uses out-of-bailiwick DNS server names. Slows down the DNS response from 22.5[ms] to 33.1[ms] on average.

Countermeasures against inefficient use

- All of them are well known
- Increasing TTL makes cache hit rate high.
- Small use of CNAME makes cache hit rate high
- Changing out-of-bailiwick DNS server names to in-bailiwick improves cache hit rate & delay
- Reducing unnecessary AAAA (or A) queries reduces queries at full-resolves.
  - If they don't have IPv6 or IPv4 connectivity.

Conclusion

- Considering the recent deployment of IPv6 and CDNs, we have analyzed current DNS traffic and have revealed findings

- We have discussed countermeasures for unnecessary AAAA queries, out-of-bailiwick DNS server names, the use of CNAMEs and short TTLs.

Future works

- Analyze after World IPv6 Launch (2012/6/6)
  - Quick results
    - 20% of AAAA answers have IPv6 addresses
    - Increased from 2%(Nov 2011)
    - 72% of AAAA answers are still NODATA
- Analyze effect of DNSSEC
- More detailed CDNs and web service issues