

Analysis of queries from viewpoint of caching servers

Tsuyoshi Toyono, Haruhiko Nishida, and
Keisuke Ishibashi
NTT Laboratories

2007 OARC Workshop

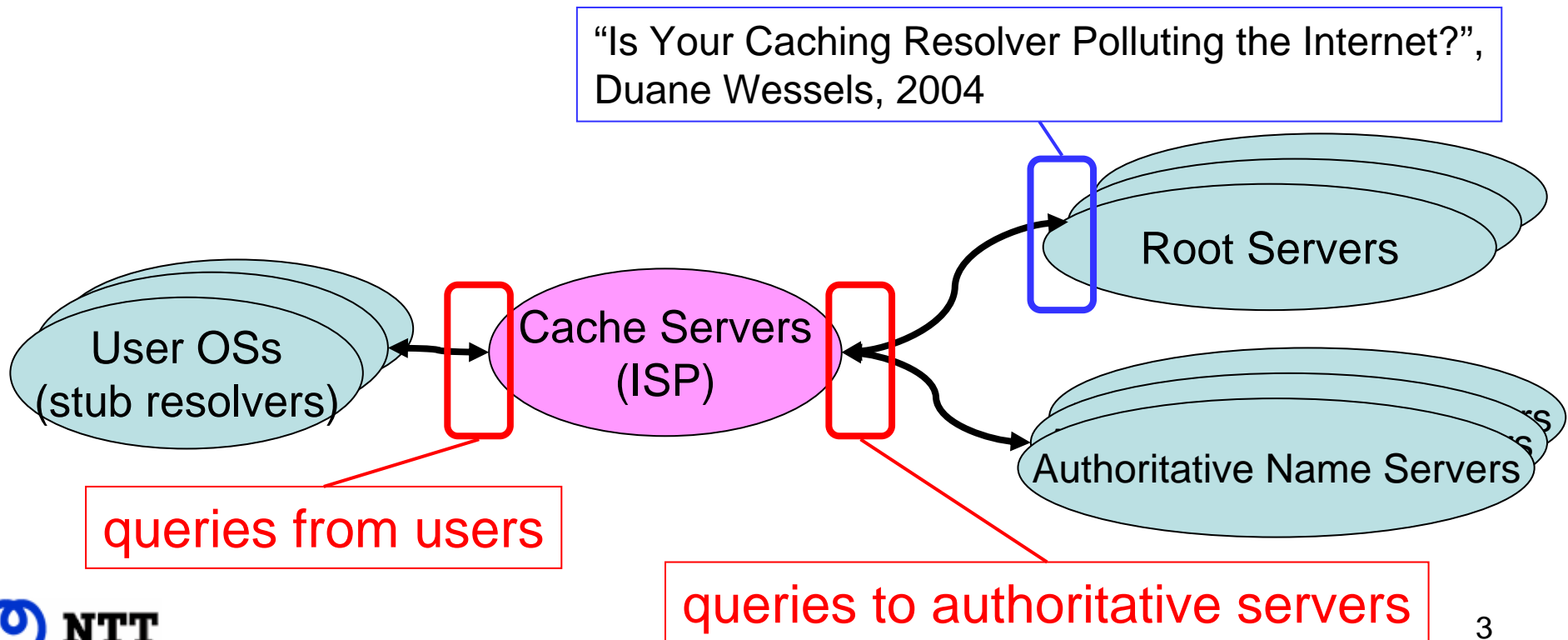
Outline

NTT Information Sharing Platform Laboratories

1. One-day summary of queries at caching servers
2. Bogus queries observed at caching servers
3. Summary & what can we do?

Focus

- DNS caching server in/out queries
 - User → Cache queries (recursive)
 - Cache → Authoritative (nonrecursive)



1) Summary of queries at caching servers

Data summary: day of cache data (May 11, 2007)

NTT Information Sharing Platform Laboratories

- Capturing queries sent to/from DNS cache servers (ISP's commercial DNS traffic)
 - Handling billions of queries/day

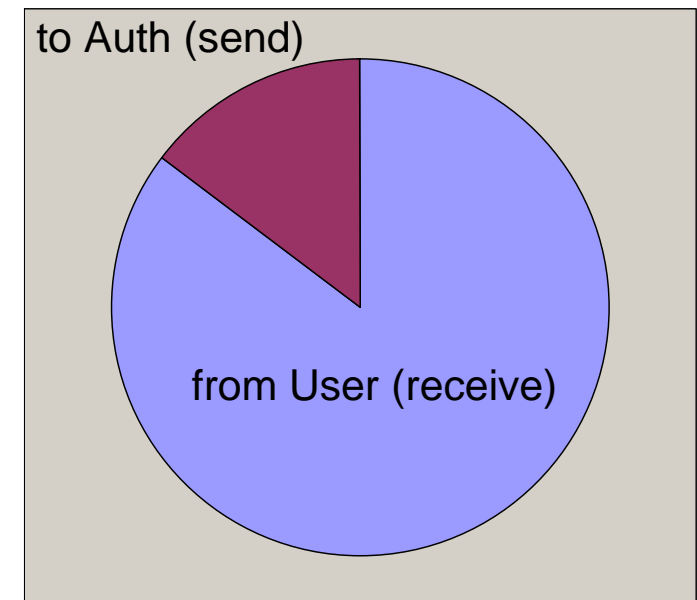
- **Cache query prevention percentage**

- About 82%

- Number of queries sent to authoritative servers
/Number of queries from users

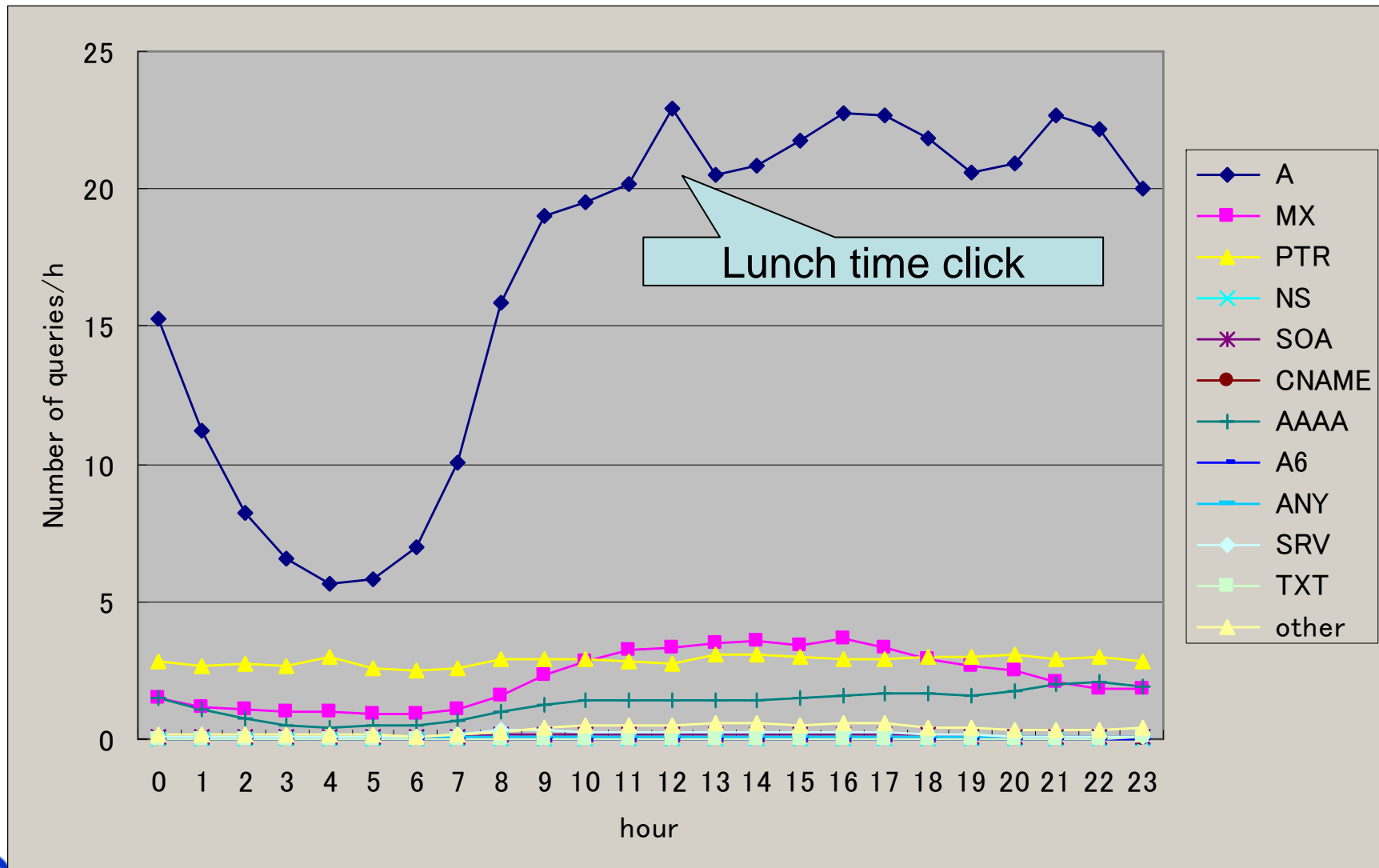
- = 0.1740...

- **82.6% prevention**



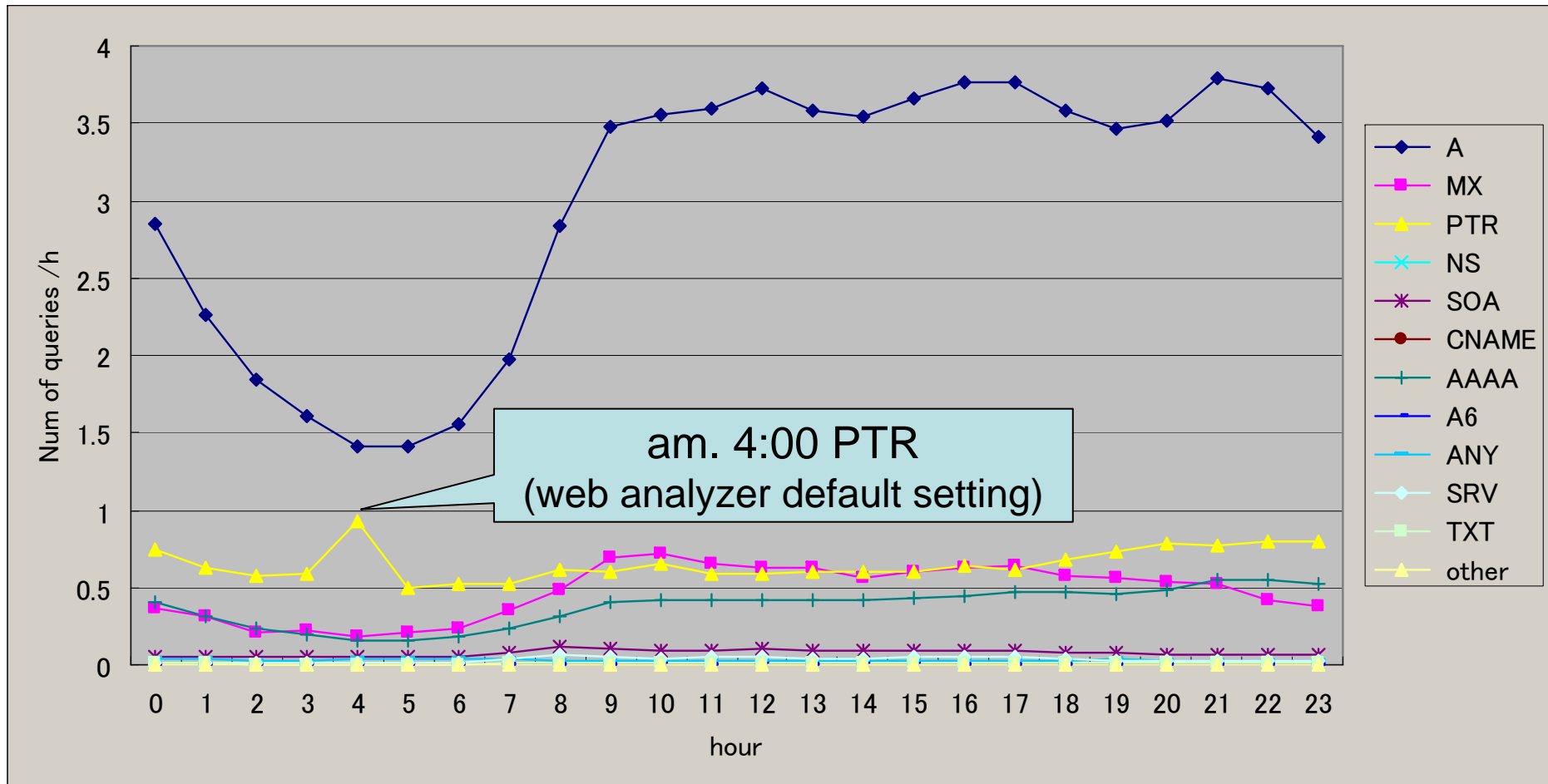
Qtype time series (queries from users/day)

NTT Information Sharing Platform Laboratories



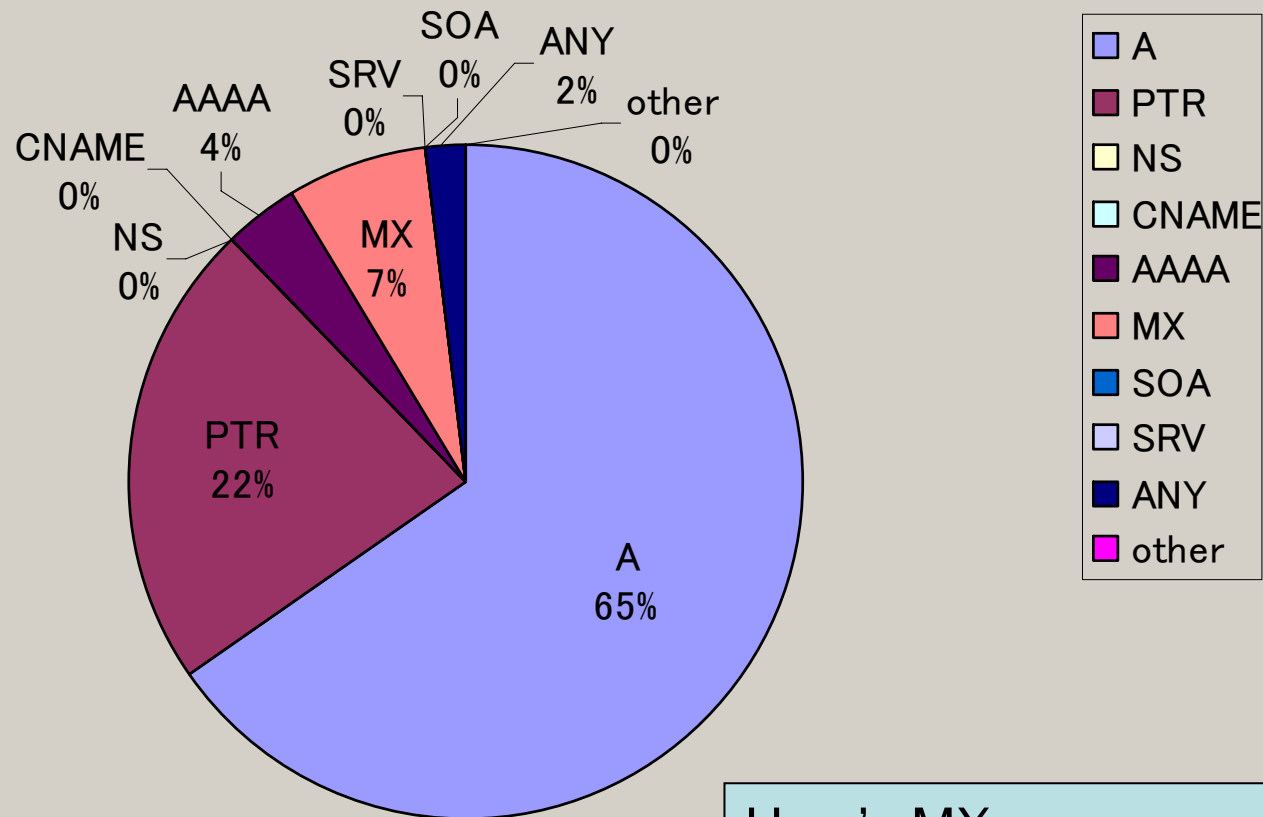
Qtype time series (queries to authoritative/day)

NTT Information Sharing Platform Laboratories



Qtype percentages (from users)

NTT Information Sharing Platform Laboratories



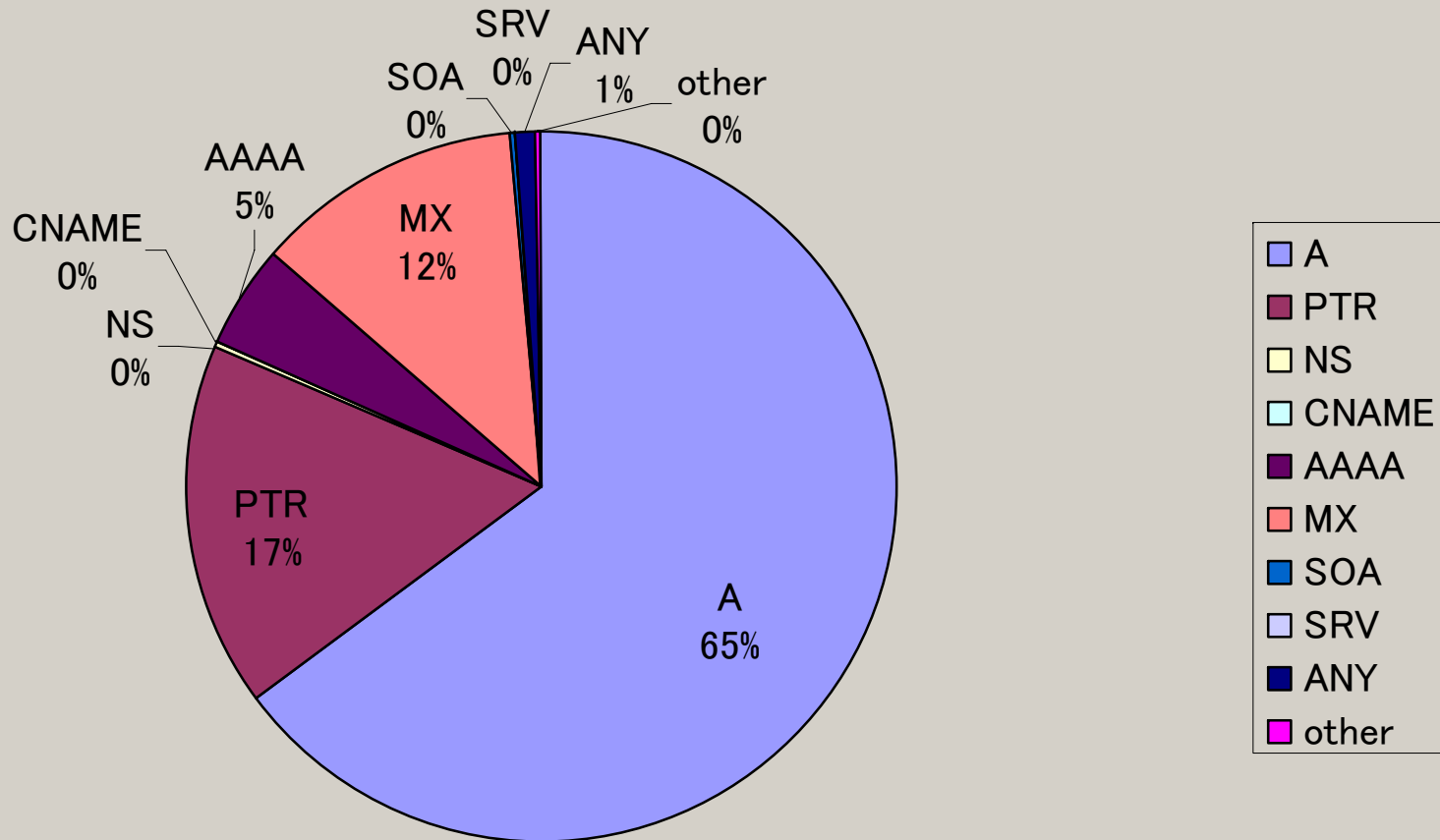
User's MX

- botnets
- SPAM sender
- Kinds of Netsky worms



Qtype percentages (to authoritative)

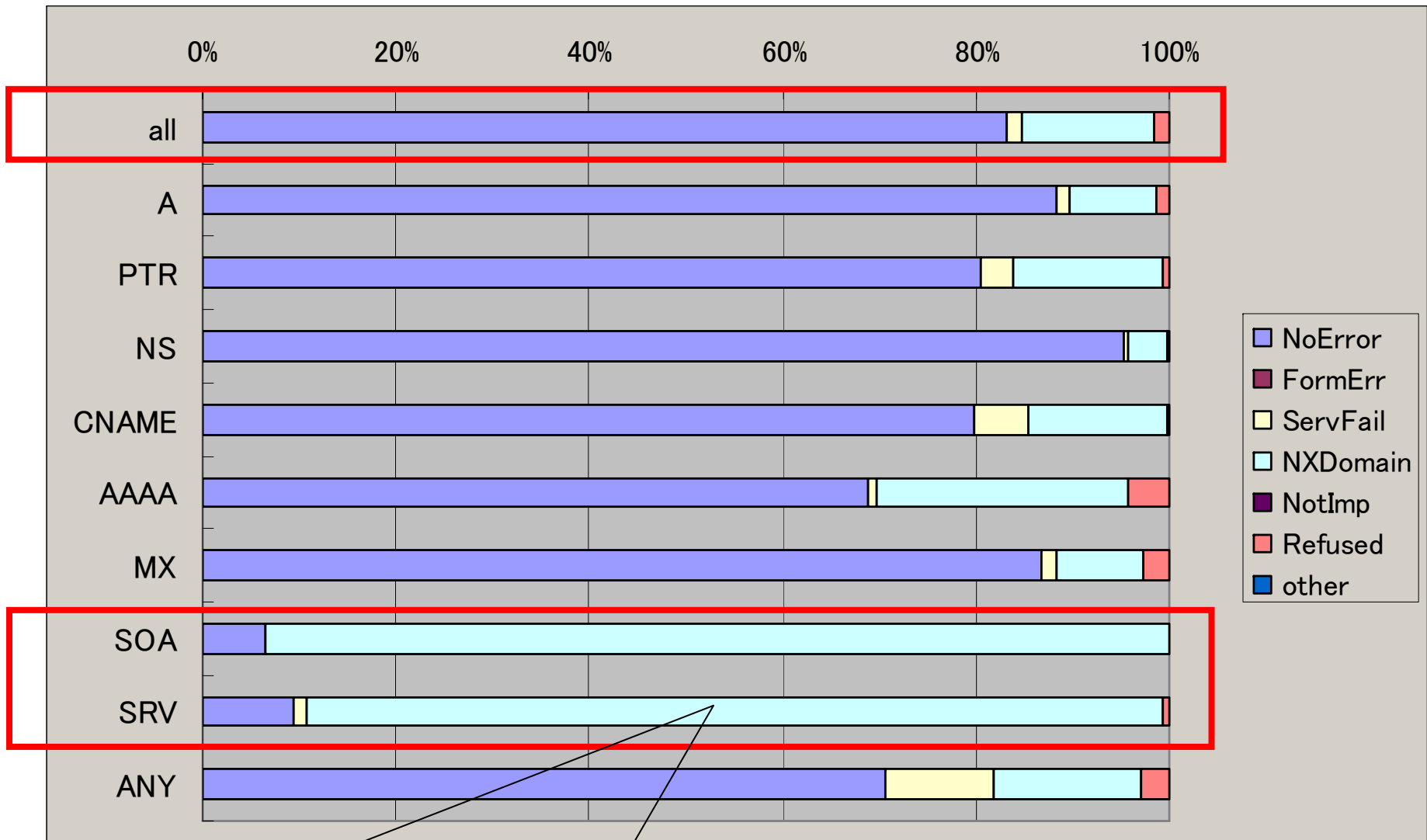
NTT Information Sharing Platform Laboratories



Distributions are similar to "from users"

Rcode-classified Qtypes (to authoritative)

NTT Information Sharing Platform Laboratories



(Ex) SRV _ldap._tcp.Default-First-Site-Name._sites.dc._msdcs.MS***.local

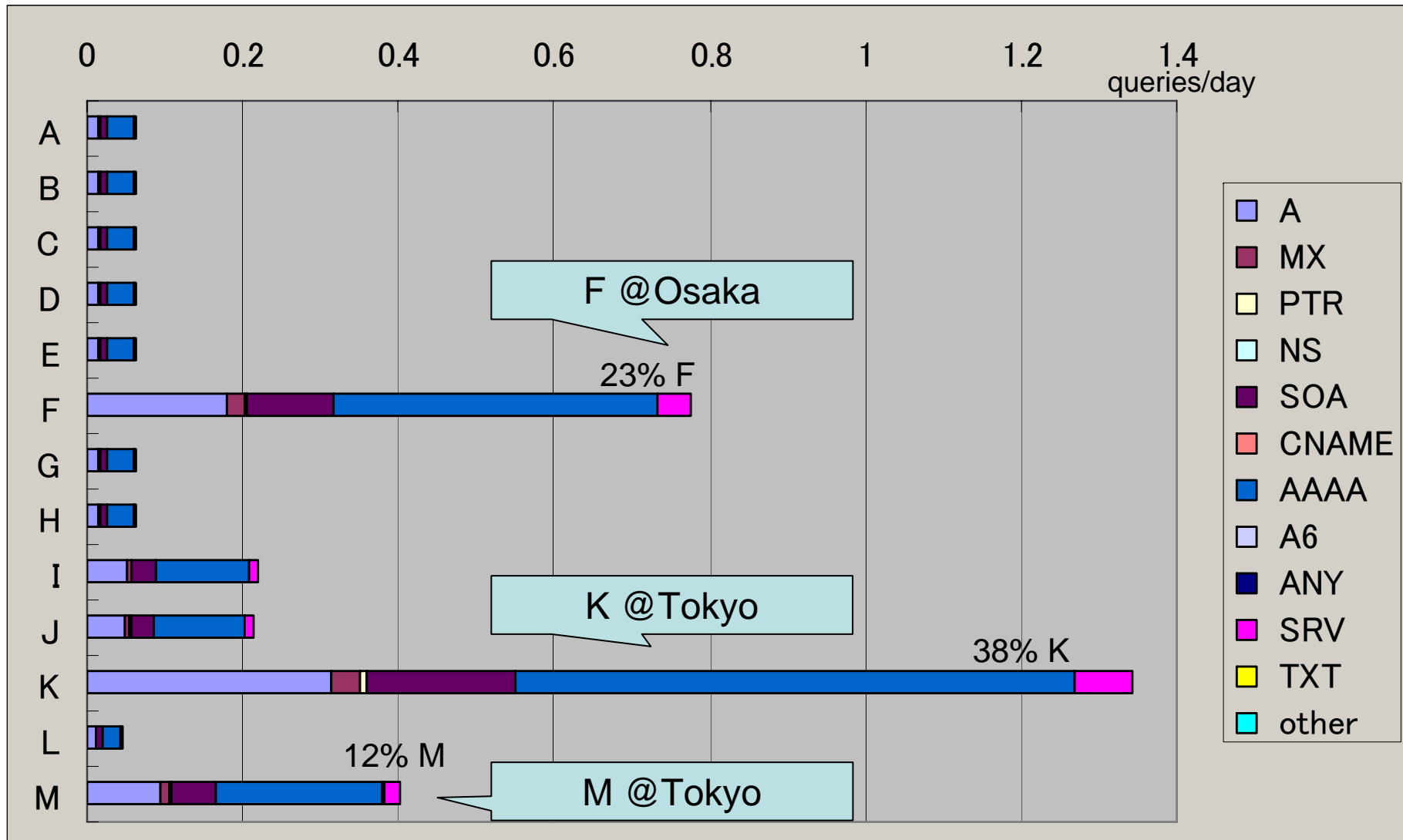
Queries sent to root servers

NTT Information Sharing Platform Laboratories

- Queries sent to 13 root servers
 - 8.23% of all nonrecursive queries

Using root servers (in Tokyo point)

NTT Information Sharing Platform Laboratories



2) Bogus queries sent
from users & sent to authoritative servers

Bogus queries: RFC1918(AS112)

NTT Information Sharing Platform Laboratories

- PTR queries for RFC1918
 - Ex. PTR “*. *. *.10.in-addr.arpa”
- From Users
 - RFC1918 PTR
 - 32.22% of all PTR queries
 - 4.42% of total queries
 - RFC1918 SOA
 - 32.27% of all SOA queries
 - 0.25% of total queries
- This is almost the same percentage of queries as that of our 2005 data.
- AS112 servers reply to these queries

Bogus queries: Invalid TLDs

NTT Information Sharing Platform Laboratories

- “localhost.”, “local.”, “localdomain.”, “workgroup.”, “wpad.” ...
- From Users
 - 6.09% of total queries
 - Percentage more than that of RFC1918 PTR queries
- To Authoritative servers
 - 7.50% of all queries (of all nonrecursive)
 - 99.5% of all queries (to root servers)
- In the end, these queries are sent to root-servers
- These queries are not prevented by DNS cache server

Bogus queries: A for A queries

NTT Information Sharing Platform Laboratories

- A queries for IP addresses
 - e.g., A “10.0.0.1”
- From Users
 - 1.53% of A queries
 - **0.88% of total queries**
- To Authoritative servers
 - 0.50% of A queries
 - 0.31% of all nonrecursive queries
 - 1.56% of total queries (to root-servers)

“TLD ranking” from user queries

NTT Information Sharing Platform Laboratories

% of total		
28.41%	PTR	arpa
25.98%	A	jp
19.94%	A	com
6.04%	A	net
2.31%	AAAA	jp
2.15%	MX	com
1.91%	MX	jp
0.98%	AAAA	com
0.96%	AAAA	localhost
0.84%	A	org
0.81%	A	local

valid
suspicious
unnecessary

“TLD ranking” to authoritative queries

NTT Information Sharing Platform Laboratories

% of total		
19.97%	A	com
15.75%	A	jp
15.61%	PTR	arpa
11.57%	A	net
5.76%	MX	com
5.43%	AAAA	localhost
2.12%	A	org
1.62%	AAAA	jp
1.51%	AAAA	com
1.32%	MX	jp
1.15%	ANY	jp

valid
suspicious
unnecessary

“TLD ranking” to root-servers queries

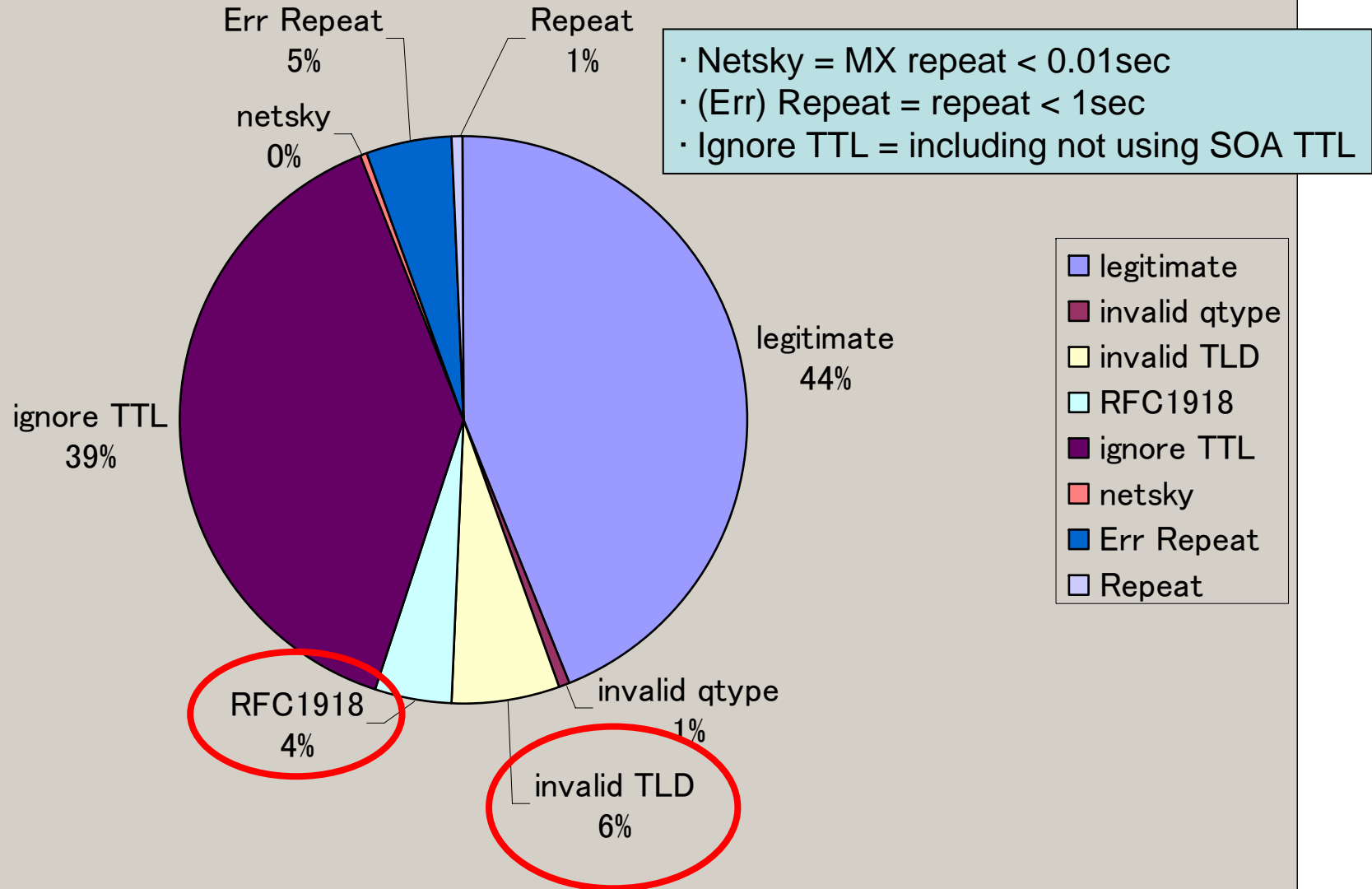
NTT Information Sharing Platform Laboratories

% of total	
70.45%	localhost
12.90%	local
1.56%	A_for_A
1.09%	localdomain
0.57%	LOCAL
0.42%	arpa
0.26%	Domain
0.22%	domain
0.19%	not-defined
0.15%	**n
0.05%	valid-ccTLDs
0.03%	WORKGROUP
11.83%	Other (almost invalid)

valid
suspicious
unnecessary

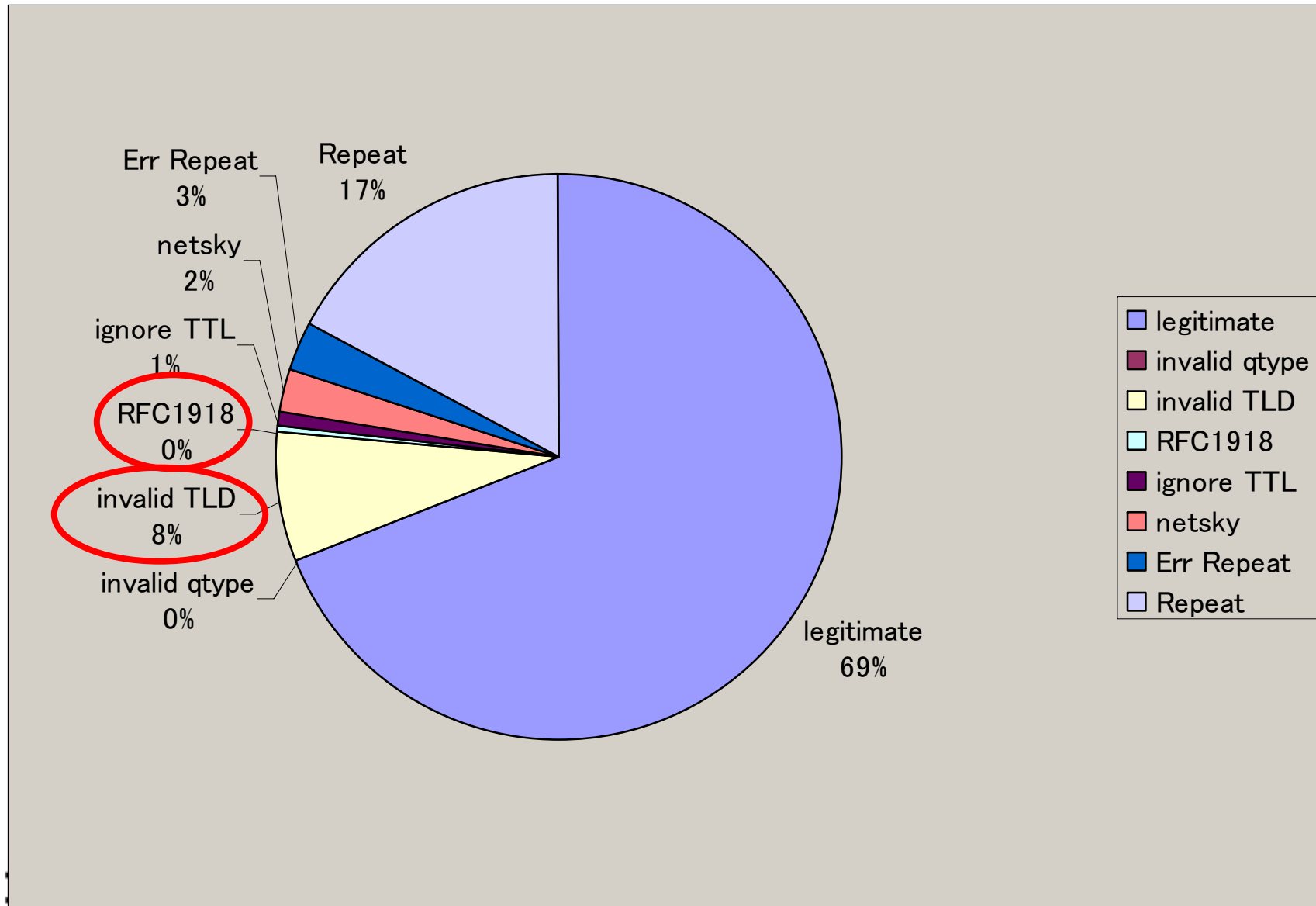
Classified queries pie chart (from users)

NTT Information Sharing Platform Laboratories



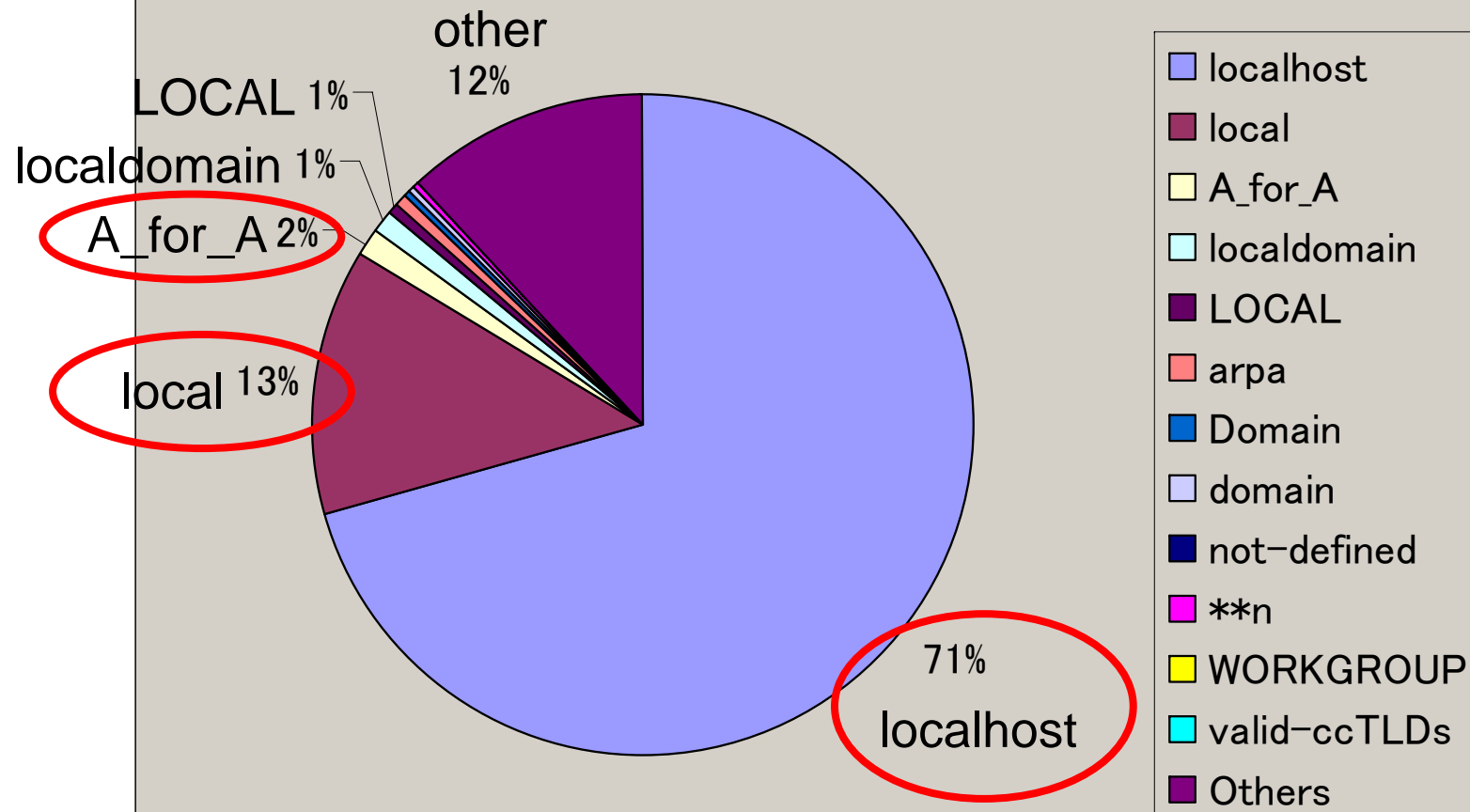
Query percentages pie chart (to authoritative)

NTT Information Sharing Platform Laboratories



“TLD ranking” to root-servers queries

NTT Information Sharing Platform Laboratories



Summary

NTT Information Sharing Platform Laboratories

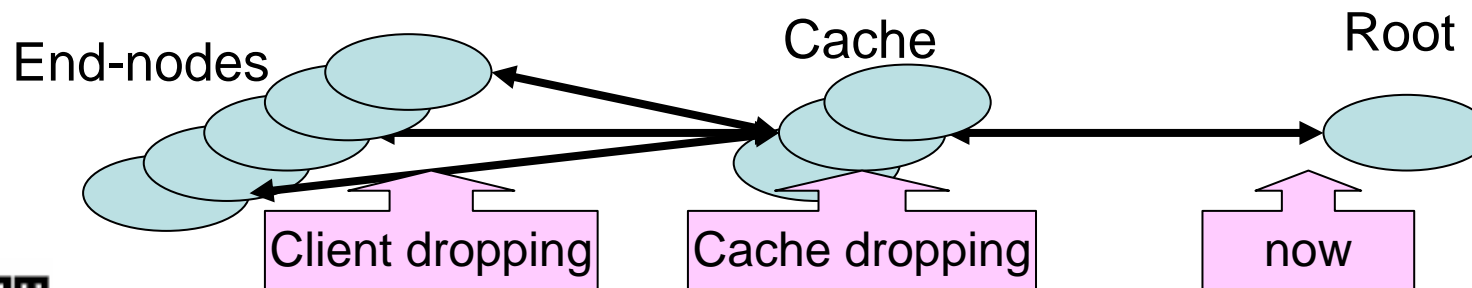
- Caching servers cache about 82% of user queries
 - Cache servers work to the best of their ability
- Number of RFC1918 PTR queries has not decreased from 2005
 - About 4% of user queries are RFC1918 queries
 - AS112 servers work to the best of their ability

But...

- The number of “Invalid TLDs” queries is more than that of RFC1918 queries
 - 6% of user queries are “invalid TLD” queries
 - 85% of these queries comprise just 3 domains
 - “.localhost.”
 - “.local.”
 - “.localdomain.”
 - Caching servers do not work effectively in some cases
 - 99.5% of to-root-servers queries are “invalid TLDs”
 - “Invalid TLDs comprise 15%~20% of F-root queries”
 - “Is Your Caching Resolver Polluting the Internet?”, Duane Wessels, 2004

What can we do?

- Is it appropriate to answer NX Domain to these queries immediately without sending authoritative queries?
 - “.localhost.” “.local.” “.localdomain.”
 - Cache server answers sent to users are same as before
- If it is, then who to answer NX Domains?
 - It's more effective if those queries are processed near end-nodes



What can we do? (cont.)

- Root (& authoritative)
 - Answer to those queries from AS112 servers
- Caching
 - Sinkhole, blackhole
 - Set up dummy authoritative servers for “.local.”
 - “default” configurations “Long negative cache TTL”
 - e.g. /etc/bind/zones.rfc1918
 - Like “draft-ietf-dnsop-default-local-zones-02”, M. Andrews
- Stub & users
 - The promotion of appropriate “default” configurations
 - The promotion of appropriate implementations

Any questions?