Large-scale DNS Caching Servers Hot Topics/An Analysis of Anomalous Queries

Shintaro NAKAGAMI†, Tsuyoshi TOYONO‡
Keisuke ISHIBASHI‡, Haruhiko NISHIDA‡, and Haruhiko OHSHIMA†

† NTT Communications, OCN
‡ NTT Laboratories

Outline

1. **Hot Topics about OCN DNS Caching Servers**
   - Introduction of OCN
   - Query Trend on OCN DNS Caching Servers
   - Problems with DNS Caching Servers

2. **An Analysis of Anomalous Queries on Large-scale Caching Servers**

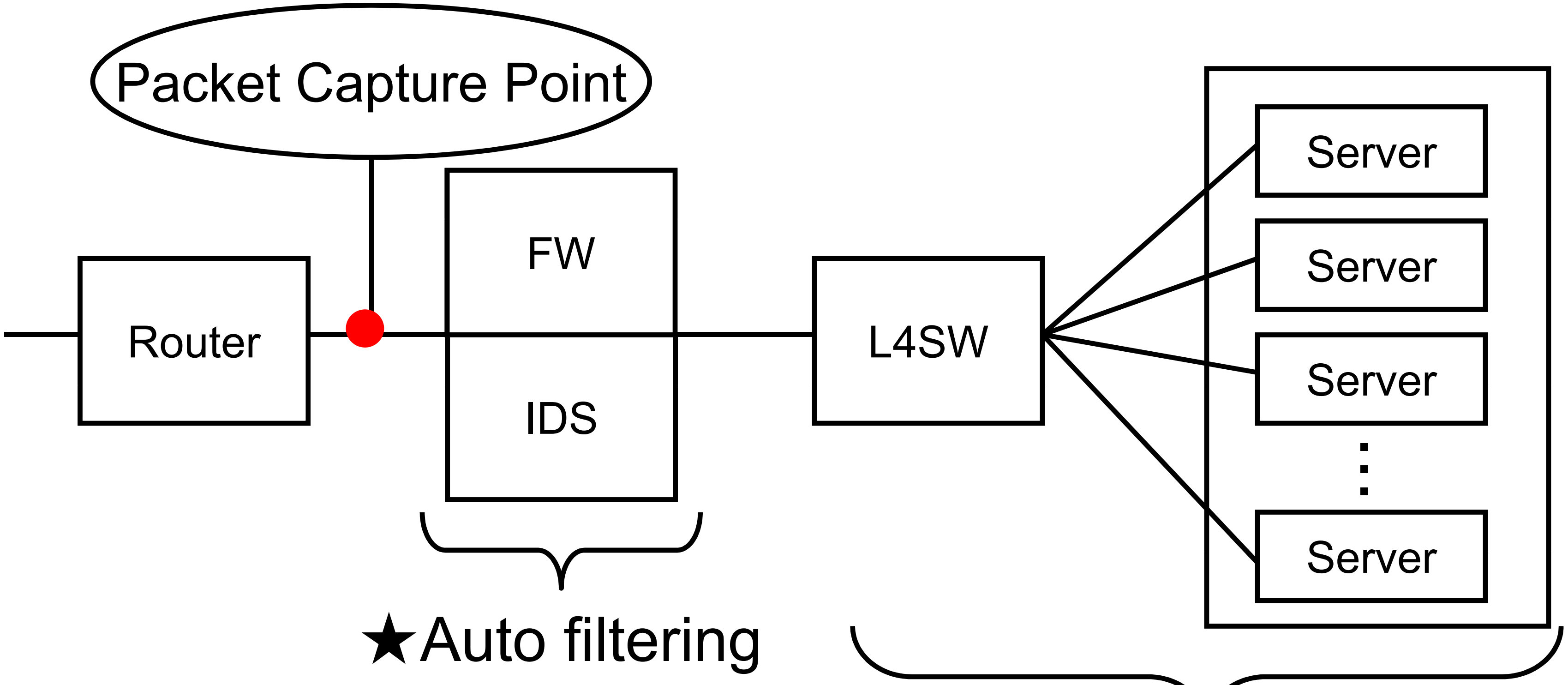
Introduction of OCN

- **OCN (AS4713)**
  - The largest ISP in JAPAN
  - 7 million customers
- **DNS operation**
  - 150 DNS servers

    -50 name servers / 100 caching servers
  - 2 kinds of DNS application

    -BIND9 / CNS (CNS has 6 times performance of BIND)
  - 6 billion queries/day (70 thousand queries/sec)

OCN Cache DNS Structure



★Auto filtering

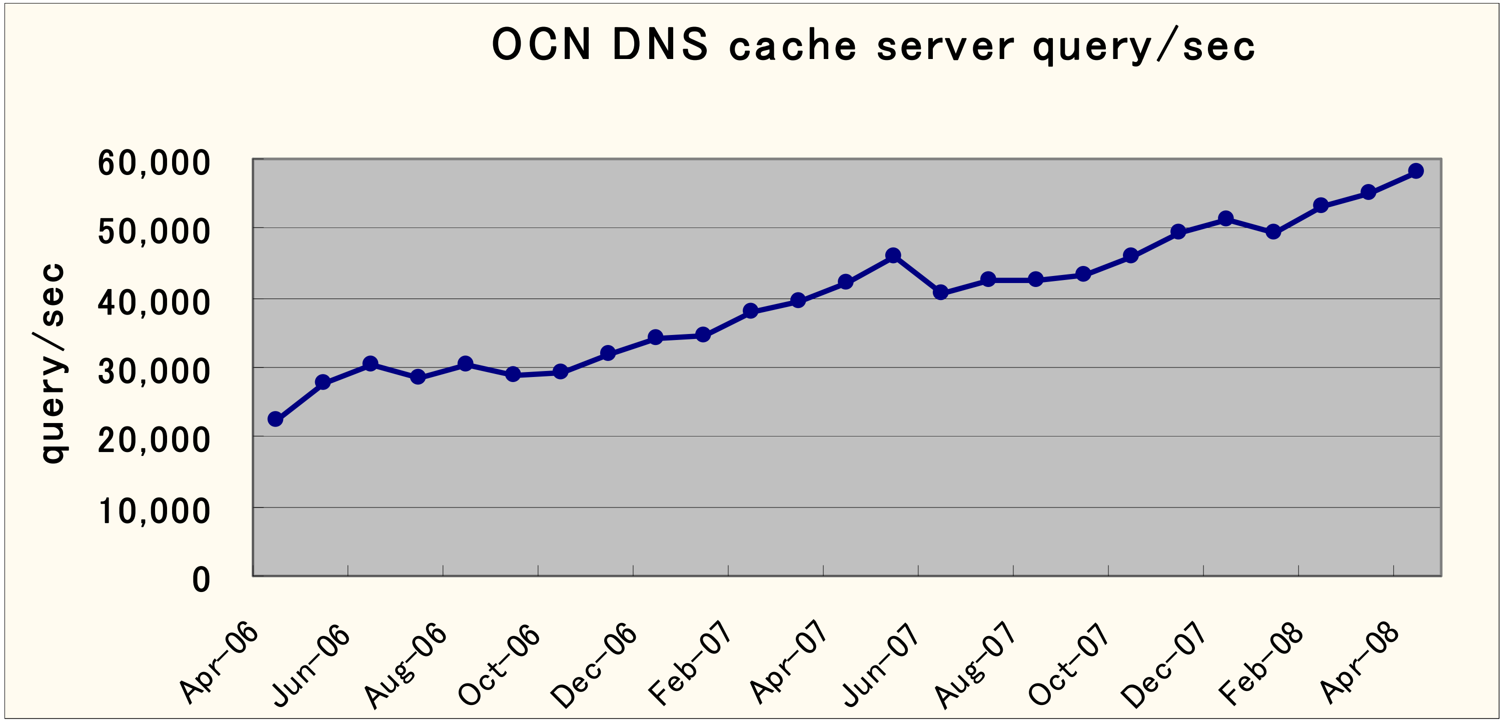
★Load Sharing

⇨ **Almost 100% Service Availability**

Query Trend on OCN DNS Caching Servers

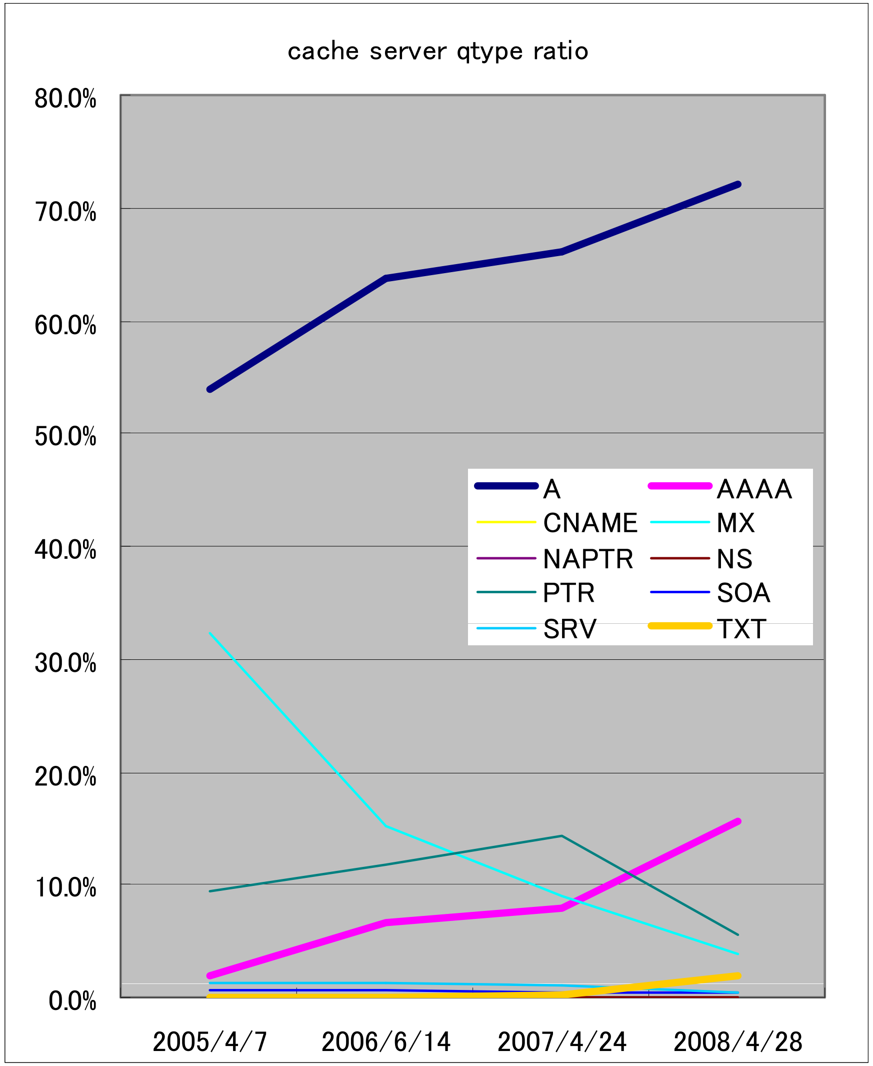
- The number of queries is increasing rapidly.
- The annual query increase rate is 150%.

  The query increase rate is much higher than the customer increase rate.



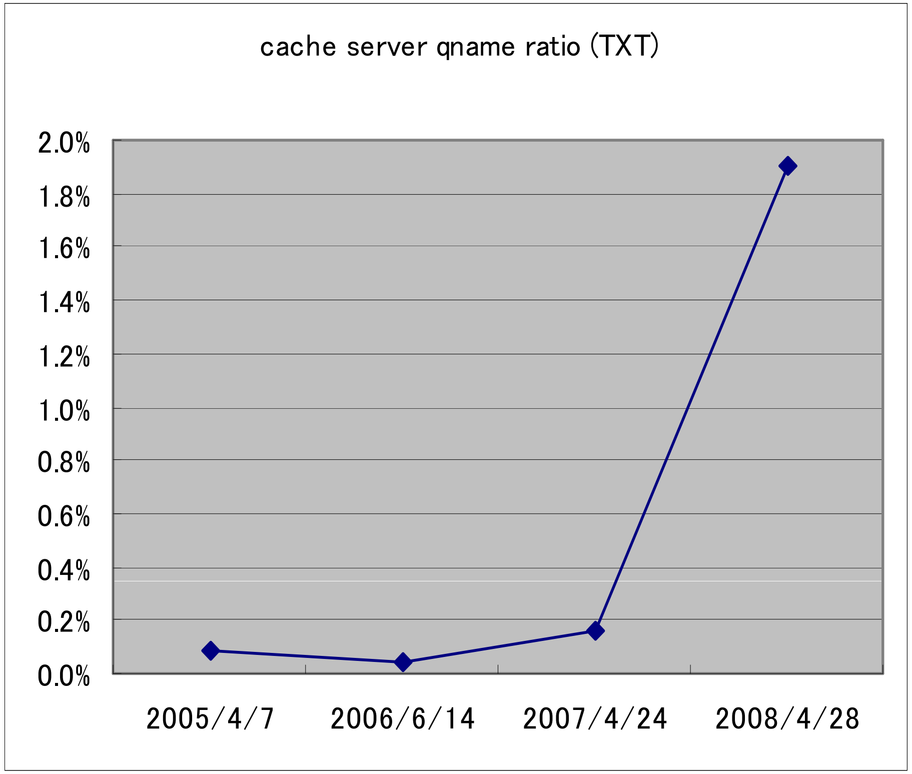
What types of Query?

- A>>AAAA>PTR>MX>TXT>>others
  - A record queries are increasing.
    - The number of customers and the number of queries per one person are increasing.
  - MX record queries are decreasing.
    - Repeat MX queries by spammer, by botnets or by worms are decreasing.
  - AAAA and TXT record queries increased rapidly this year.



TXT Record Queries

- TXT record is used for reputation check, SPF, DNSBL and so on.
- Queries for reputation check are increasing.
- SPF queries from mail servers are also increasing.
- There were only a few queries for DNSBL check until last year.

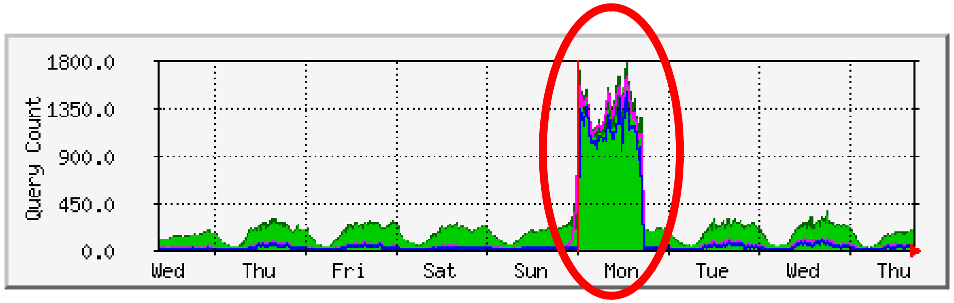


Problems with DNS Caching Servers

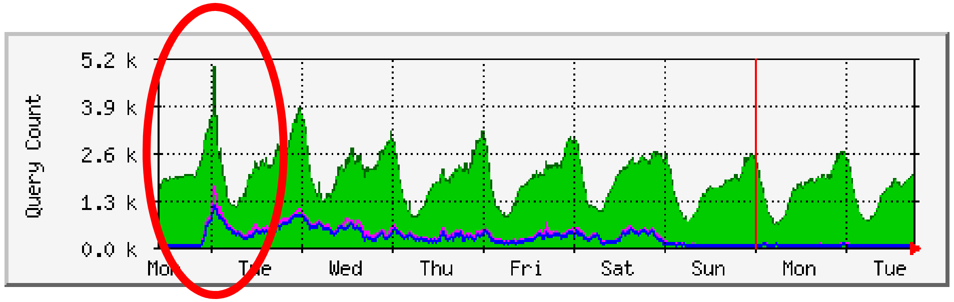
- The load of caching servers is higher than that of name servers.
- Problem queries
  - DDoS attack queries
  - Bogus queries
  - Queries for Short TTL records
- Birthday attack and Amp attack aren’t observed so much.

DDoS Attack Queries

- Attacks by worms (2004/04)
  - □ The number of queries at this time is 6 times more than usual.
  - □ Forward operation was effective in this attack.



- Attacks by botnets (2007/10)
  - □ The number of queries at this time is 2 times more than usual.
  - □ Auto filtering by IDS worked effectively in this attack.



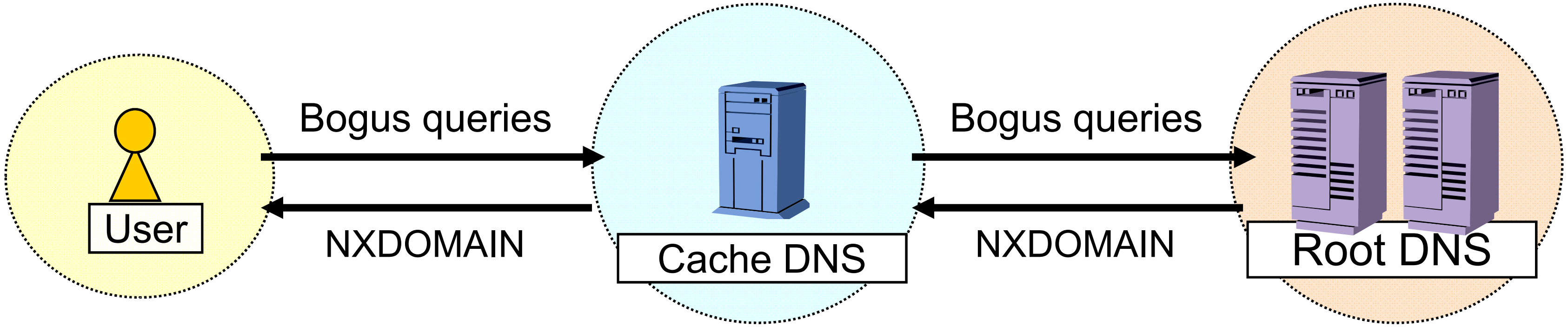
- In these case, there were a lot of SERVFAIL queries .
  - □ SERVFAIL queries cause a heavy load in caching servers.

Bogus Queries

- Caching servers receive a lot of Bogus queries.
  - □ PTR queries for RFC1918 (private IP address)

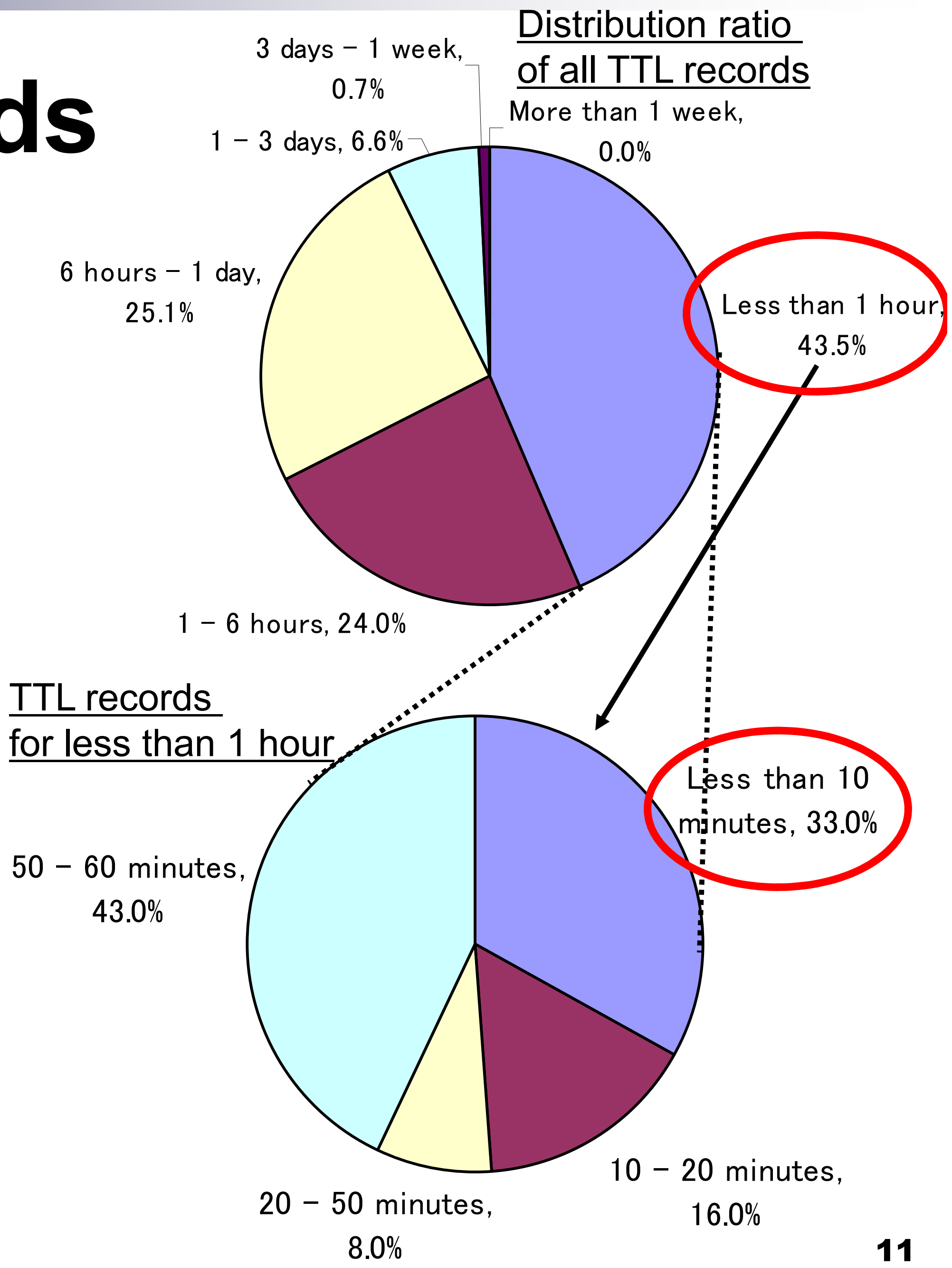
    -PTR "*.*.*.10.in-addr.arpa."
  - □ Invalid TLD

    -*.localhost, *.local
- These queries are sent to root-servers as well as cache-servers. -> Useless traffic and processing



Short TTL Records

- The Distribution ratio of TTL records in OCN caching servers.
- TTL records for less than 1 hour account for 43.5%.
- TTL records for less than 10 minutes account for 14%.
- There are also 1 second TTL records.
- If it isn't necessary, long TTL is desirable.



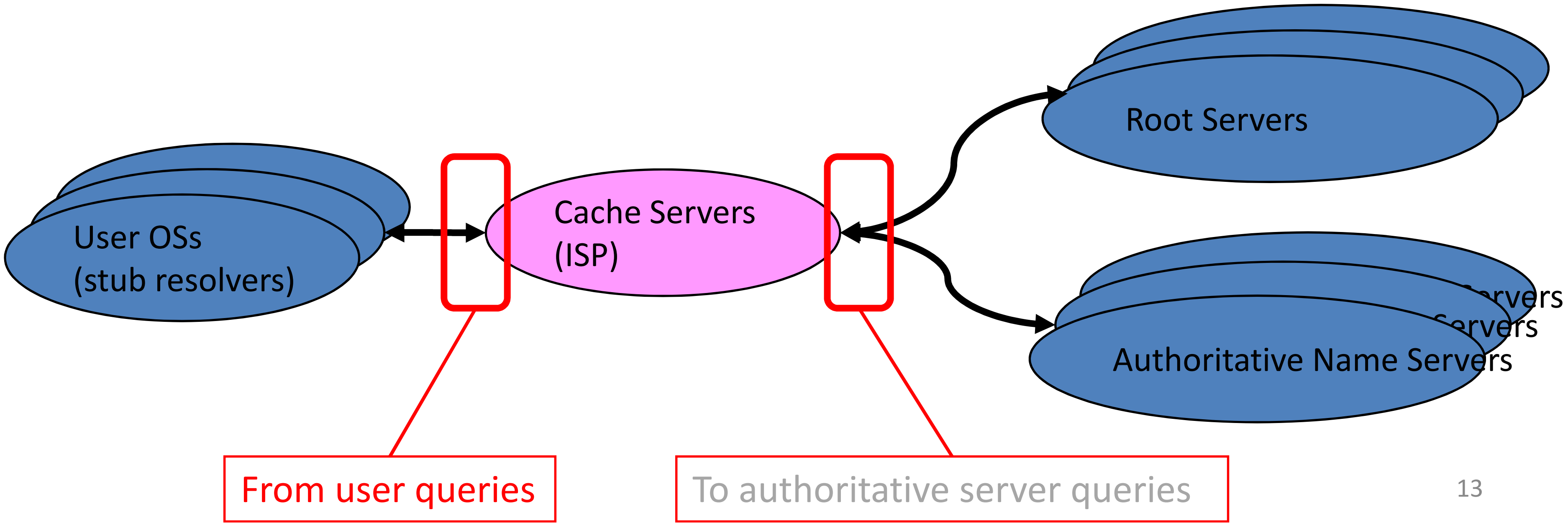
Part 2.
An Analysis of Anomalous Queries on Large-scale Caching Servers

Tsuyoshi TOYONO

NTT Lab.

Focus on

- DNS caching servers' in/out queries
  - User -> Cache queries (recursive)
  - Cache -> Authoritative (non-recursive)



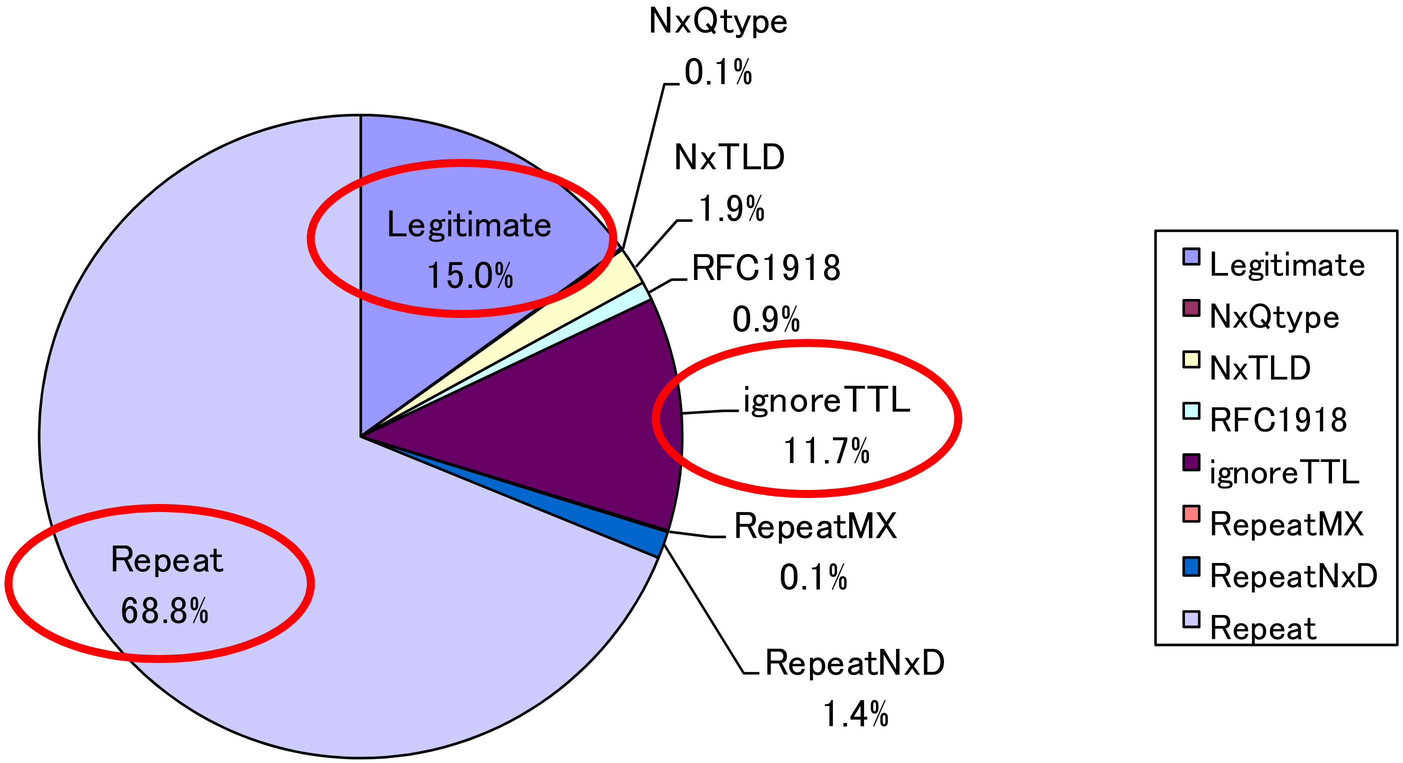
What are “Anomalous queries” ? (1/2) Invalid queries

1. Nx-Qtype（Non-existent Qtype)
   - Invalid or broken Qtype
   - (Ex.) Type 0, Type 990 …
2. Nx-TLD（Non-existent Top Level Domain)
   - (Ex.) “.localhost.”, “.localdomain.”, ”.workgroup.” …
3. RFC1918 PTR
   - PTR queries for RFC1918
   - (Ex.) PTR “1.0.0.10.in-addr.arpa”

What are “Anomalous queries” ? (2/2) They ignore our answers …

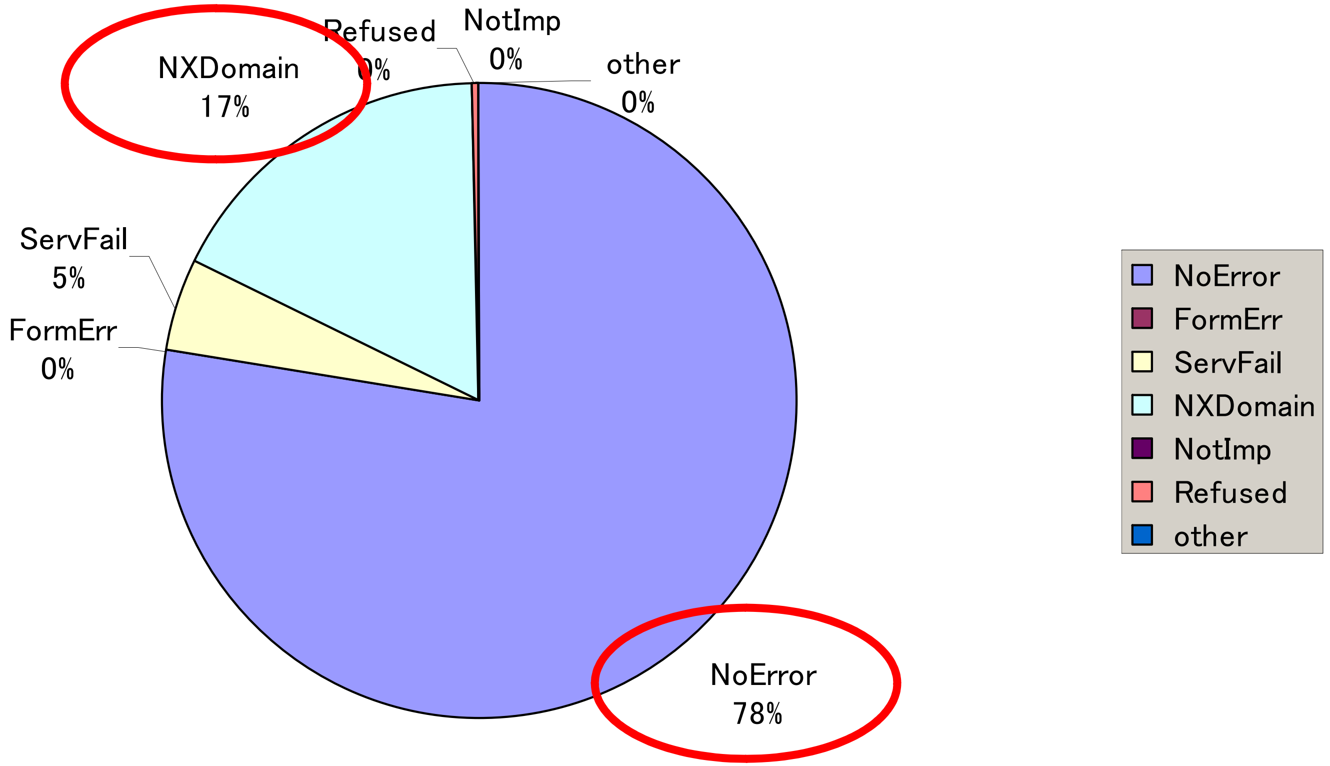
4. Repeat queries
   – Repeat same “Qtype, Qname” queries from same IP address within very short time (1 sec)
5. Other repeat queries
   – 5-1. Ignore TTL
      - Repeat same queries that ignored TTL
   – 5-2. Repeat MX
      - Repeat “MX” queries within very short time (0.1 sec)
      - Characteristic behavior in some worms (Ex.) Netsky
   – 5-3. Repeat Error
      - Error status answers (ServFail, FormErr, Refused) are replayed, but query is repeated

User queries (to caching servers)



- Legitimate queries: only 15% of all queries
- "Repeat" and "Ignore TTL" are 80% of all queries

Server answers (to users)



- Most answers are normal
  - 78% of total answers are “No Error”
  - 17% of total answers are “NXDomain”
- Few error answers (Server Fail, Format Err, Refused)

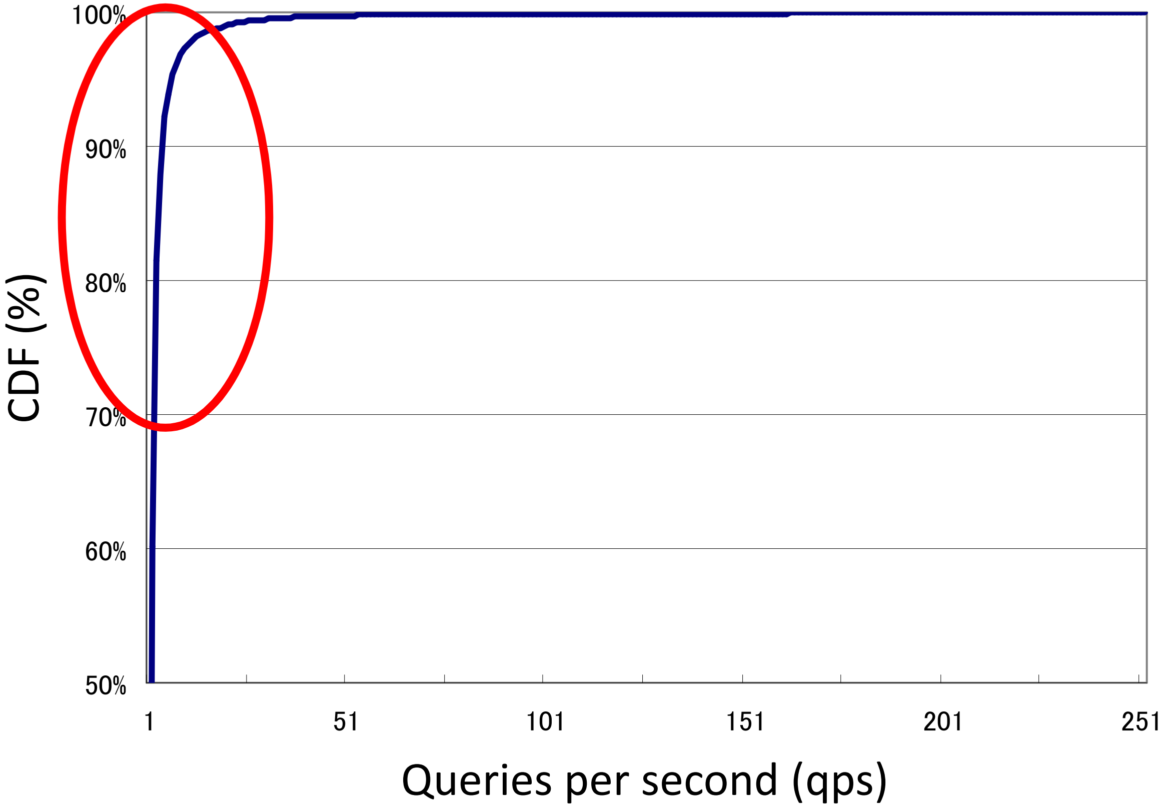
First question …

- We receive …
  - 80% anomalous queries
  - Only 15% legitimate queries

- … But do all users behave like that ?

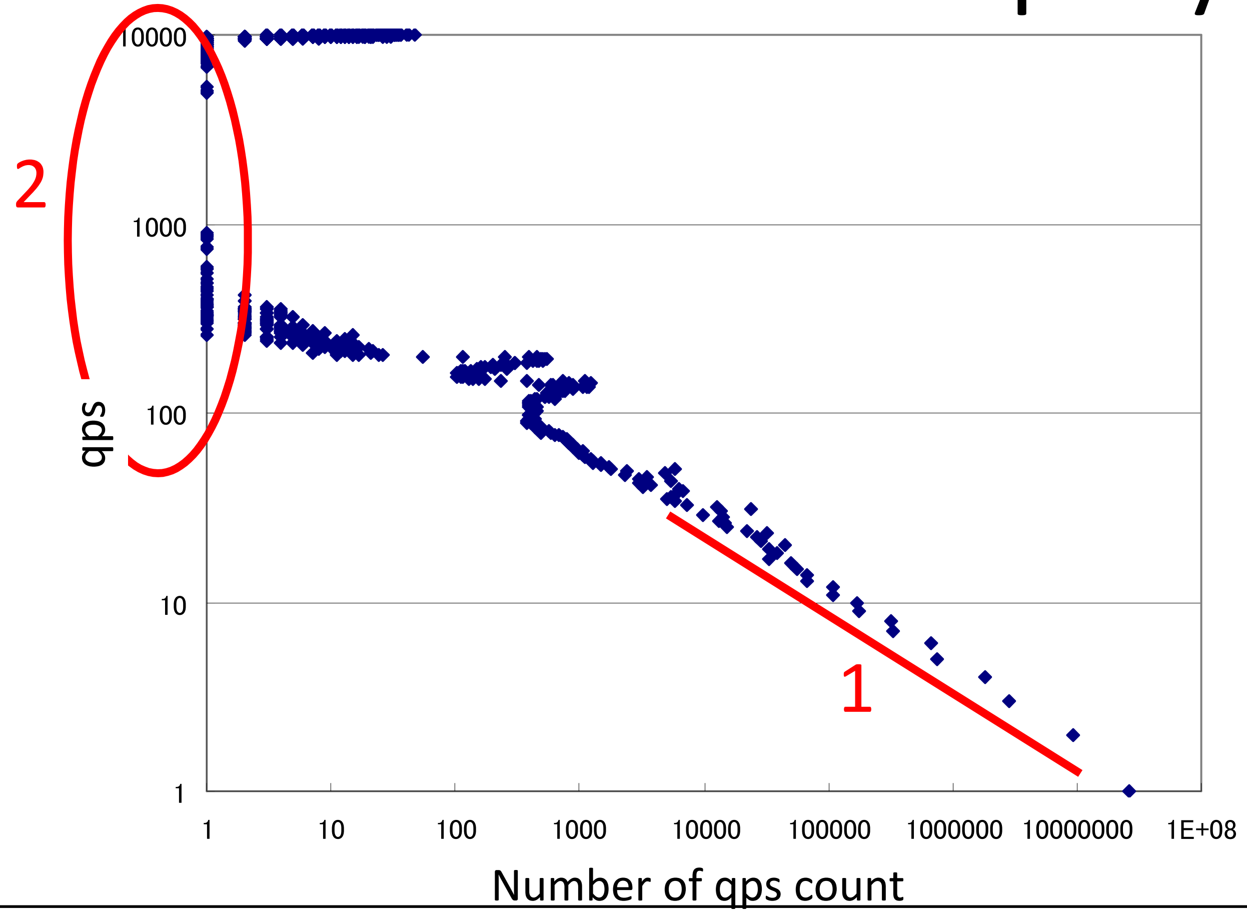
- Analysis of per user queries

Number of queries per user per second (CDF)



- Most users sent a few queries (1 ~ 10 qps)
- Only 0.07% of all users sent over 100 qps at some point

Distribution chart of user query rates



1. Obeys Zipf’s law
   - Most users sent a few queries, a few users sent most of the queries
2. Exceptions of “over-10 000-qps users” !

Percentage of anomalous queries by query rate

(Percentage of total queries)

type \ rate	100qps	200qps	300qps	400qps	500qps
Legitimate	**0.09%**	**0.01%**	0%	0%	0%
NxQtype	0%	0%	0%	0%	0%
NxTLD	0%	0%	0%	0%	0%
RFC1918	0.80%	0%	0%	0%	0%
ignoreTTL	1.63%	0.05%	0.01%	0%	0%
RepeatMX	0.01%	0%	0%	0%	0%
RepeatNxD	0.64%	0%	0%	0%	0%
Repeat	59.69%	59.69%	59.69%	59.69%	59.69%

- Most queries from high query rate users are “repeat” and “ignore TTL”
- NO legitimate queries from users sending over 300qps

Second question …

- A few users send most repeat queries

- What do they want to know so much?

- Close analysis of details of repeat queries

Analysis of details of repeat queries (1/3)

- We observed 4 characteristic types in high query rate users

- (Type A) NTP servers
  - 3.9% of high query rate users,
    but 70% of high query rate queries
  - "I want to know the correct time!"
  - Repeated public NTP servers over 10 000qps continuously
    - (Ex.) "time.stdtime.gov.tw."

Analysis of details of repeat queries (2/3)

- (Type B) Mail servers
  - 76.4% of high query rate users
  - “I want to find good SPAM servers!”
  - Repeated “A” and “MX” record queries including strings such as “mail”, “mx”, “smtp”

- (Type C) Messenger servers
  - 7.8% of high query rate users
  - Repeated major messenger service servers
    - (Ex.) AOL AIM, MSN, Windows Live, Yahoo …
  - What is their purpose?

Analysis of details of repeat queries (3/3)

- (Type D) PTR queries
  - 7.8% of high query rate users
  - Repeated “PTR” record for many IP addresses
  - Perhaps due to web log analyzer or related tools
- Others (Unclassified)
  - Repeated queries for SNS web site domains
  - Repeated queries including strings “pic” “img” “photo” …

Summary

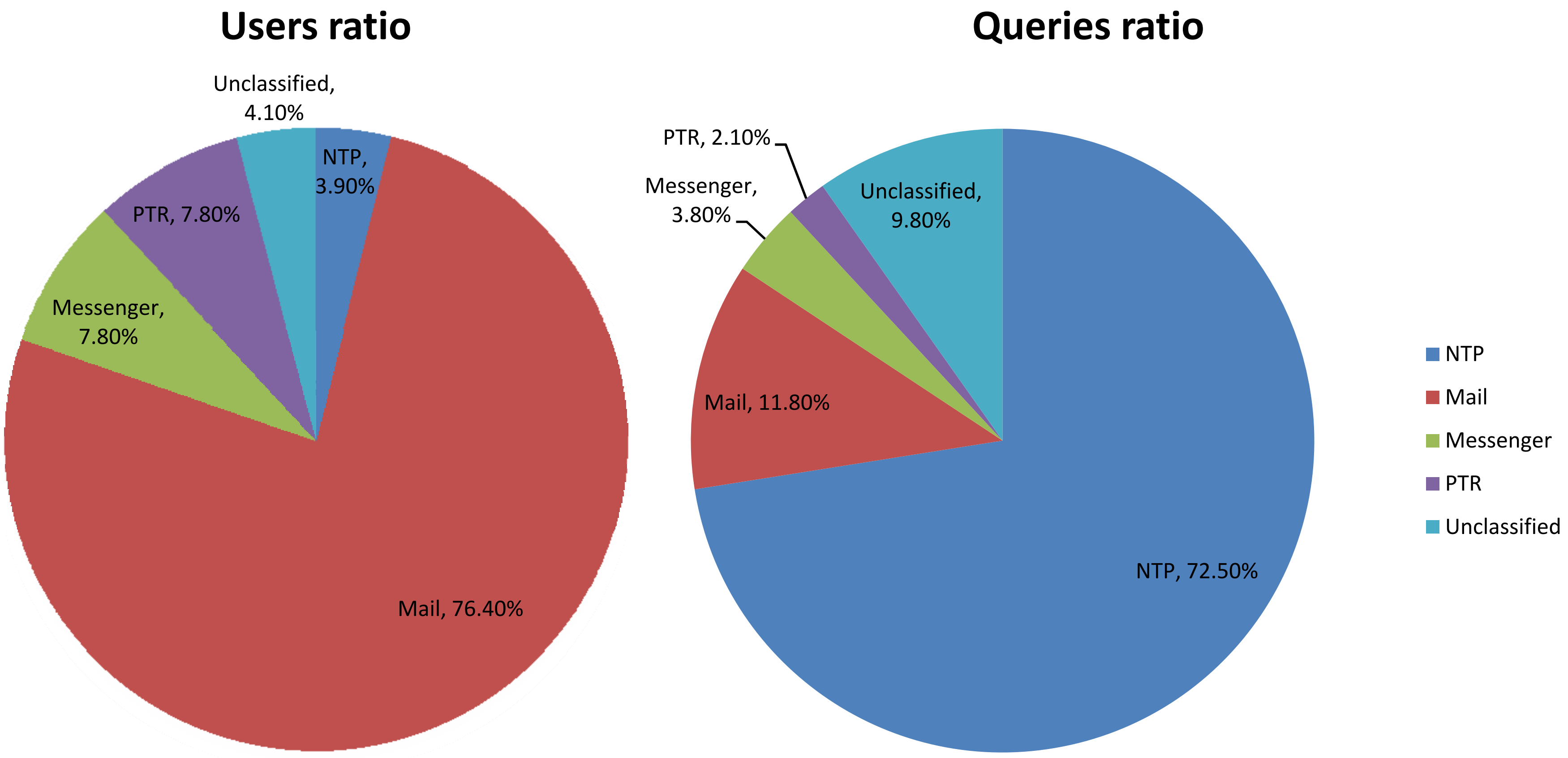
- All queries from high query rate user are bogus or unnecessary.
- We can prevent these anomalous queries easily.
  - Apply query rate limit control per user
    - In this case, 300 qps
  - The load on DNS servers will decrease.

Conclusion

- We should consider the way to exclude bogus queries.
- We hope for the development of strong BIND for caching servers.

Fin.

Analysis of details of repeat queries


Users ratio
Unclassified, 4.10%
NTP, 3.90%
PTR, 7.80%
Messenger, 7.80%
Mail, 76.40%
Queries ratio
PTR, 2.10%
Messenger, 3.80%
Unclassified, 9.80%
Mail, 11.80%
NTP, 72.50%
NTP
Mail
Messenger
PTR
Unclassified