An idea of IP anycast analysis using DITL dataset

Kazunori Fujiwara, JPRS
<fujiwara@jprs.co.jp>

OARC 37 Workshop

Analysis of IP anycast effectiveness

- [RFC 8499] defines "Anycast" as "The practice of making a particular service address available in multiple, discrete, autonomous locations, such that datagrams sent are routed to one of several available locations." (Quoted from [RFC4786], Section 2) See [RFC4786] for more detail on Anycast and other terms that are specific to its use.
- To analyze the effectiveness of IP anycast, need to know
  - Location of each server instance
  - Location of each client (by IP map services)
  - Latencies between each client and each root server node

DNS-OARC 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/oarc/data/ditl
  - 50 hours (or less) packet capture at (some of) root DNS servers and other DNS servers
  - Source IP addresses may be anonymized

Guessing anycast node location from DITL dataset

- DITL dataset is located on DNS-OARC's analysis servers
- Directory names may contain anycast location name
  - $BASE_DIR/DITL-20200505/CLEAN/?-root/{node}/DATETIME.pcap.gz
  - {node} patterns depend on each letter
  - Typical {node} patterns
    - A, J: label-area-[aj] → extract second word. If area =~ /^el¥
    - C: name.c.root-servers.org → extract fourth lebel label
    - K: ns?.area.k.ripe.net → extract fourth level label
    - M: MROOT-area{|-IXname} → extract second label (separator -)
  - prepend letter name → For example, atko, csin, dtojp, hhkg, jtko, kjptyo, mNRT

M-root node names

- M-root uses the airport code of the neighborhood as the instance name
  - M-Root locations: Tokyo (3), Osaka, Seoul, San Francisco (2), Paris (2), Brisbane, Hanoi (See m.root-servers.org)
- From directory name, there are 10 names in DITL-2021 dataset
  - BNE: Brisbane airport → Brisbane, Australia
  - ICN: Incheon Airport → Seoul, Korea
  - KIX: Kansai International Airport → Osaka, Japan
  - NRT-DIXIE: Narita Airport → Tokyo, Japan (DIX-IE) (→ merge as NRT)
  - NRT-JPIX: Narita Airport → Tokyo, Japan (JPIX) (→ merge as NRT)
  - NRT-JPNAP: Narita Airport → Tokyo, Japan (JPNAP) (→ merge as NRT)
  - CDG: Paris Charles de Gaulle Airport → Paris, France
  - ORY: Paris Orly Airport → Paris, France
  - SFO: San Francisco, United States
  - SJC: San Jose, United States (near San Francisco)

Query latencies

- DITL dataset contains packet sequences of complete TCP sessions

t1: Client sends SYN	to Server	(client: connect)
t2: Server sends SYN/ACK	to Client	(at/before server:accept)
t3: Client sends ACK	to Server	(writable at client)
t4: Client sends DNS query	to Server	(send at client)
t5: Server sends DNS reply	to Client	(send at server)

  - Each timestamps (t1,t2,t3,t4,t5) is captured at the server
  - If TCP protocol stack and DNS software send the next packet immediately after receiving the previous packet,
    - the RTT is calculated as the sending interval

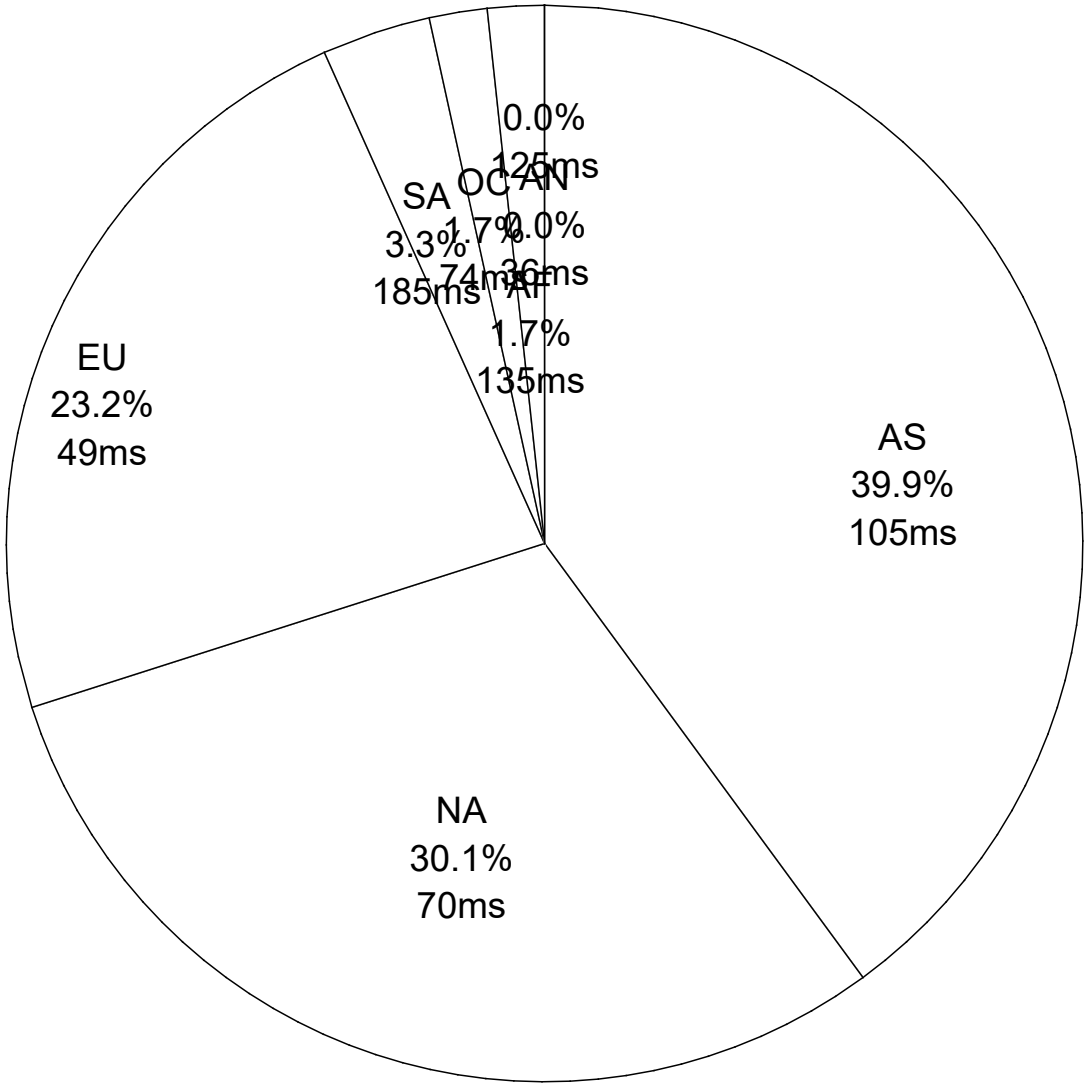
t{ACK}-t{SYN} or t{DNS query}-t{SYN}	t3-t1 or t4-t1
t{DNS reply} – t{SYN/ACK}	t5 - t2

Preliminary results

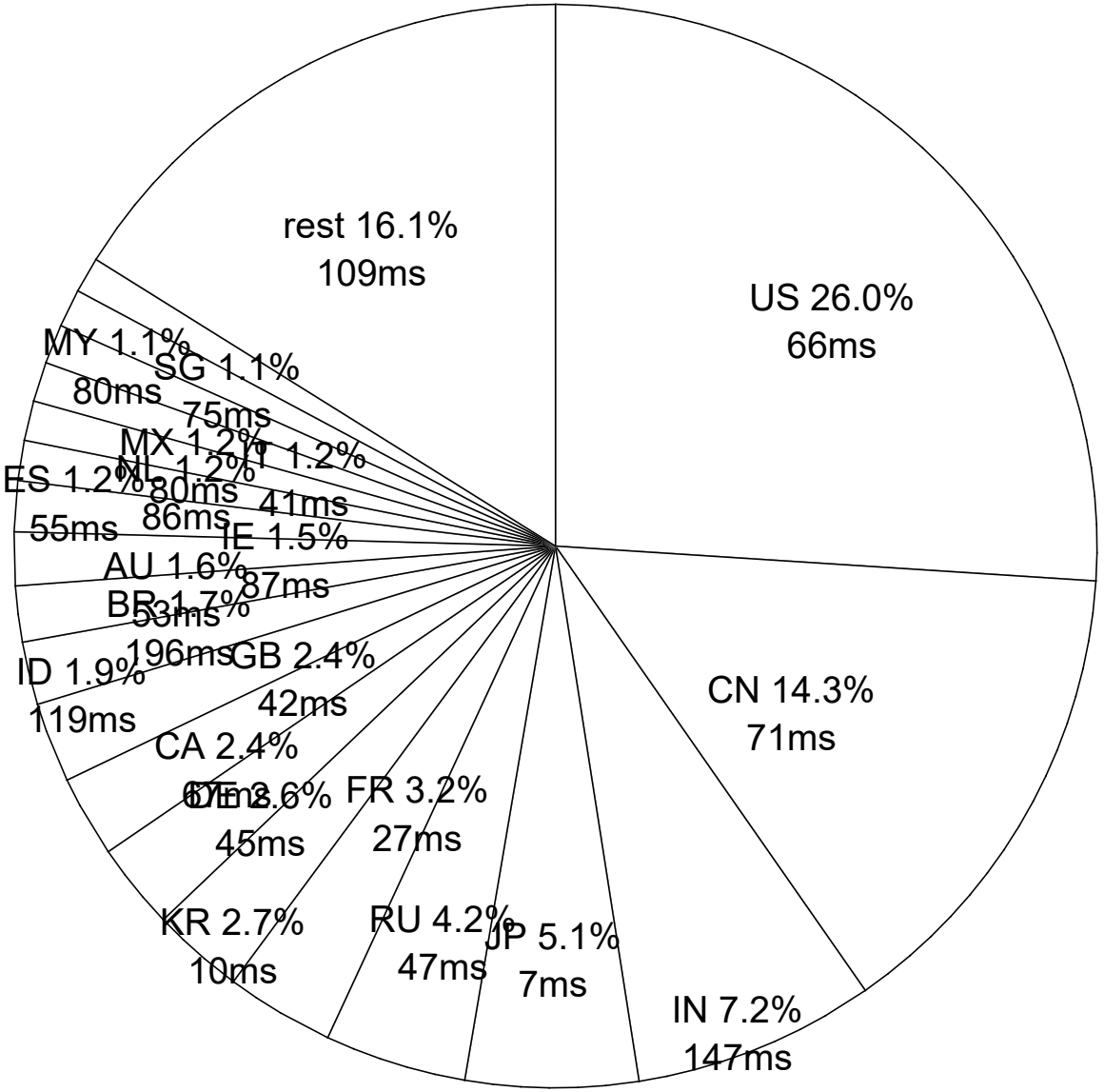
- Limitations
  - Long latencies are ignored (over 1 second)
    - Because of TCP retransmission
  - It compares average RTT between each client and each node

- M-Root 2021 : queries, average delay by continents/countries
- M-Root2021: RTT / query CDF by continent
- M-root Paris nodes: queries, average RTTs by countries
- Effectiveness of M-root Brisbane, AU node

M-Root2021 : queries, average RTT by continents/countries

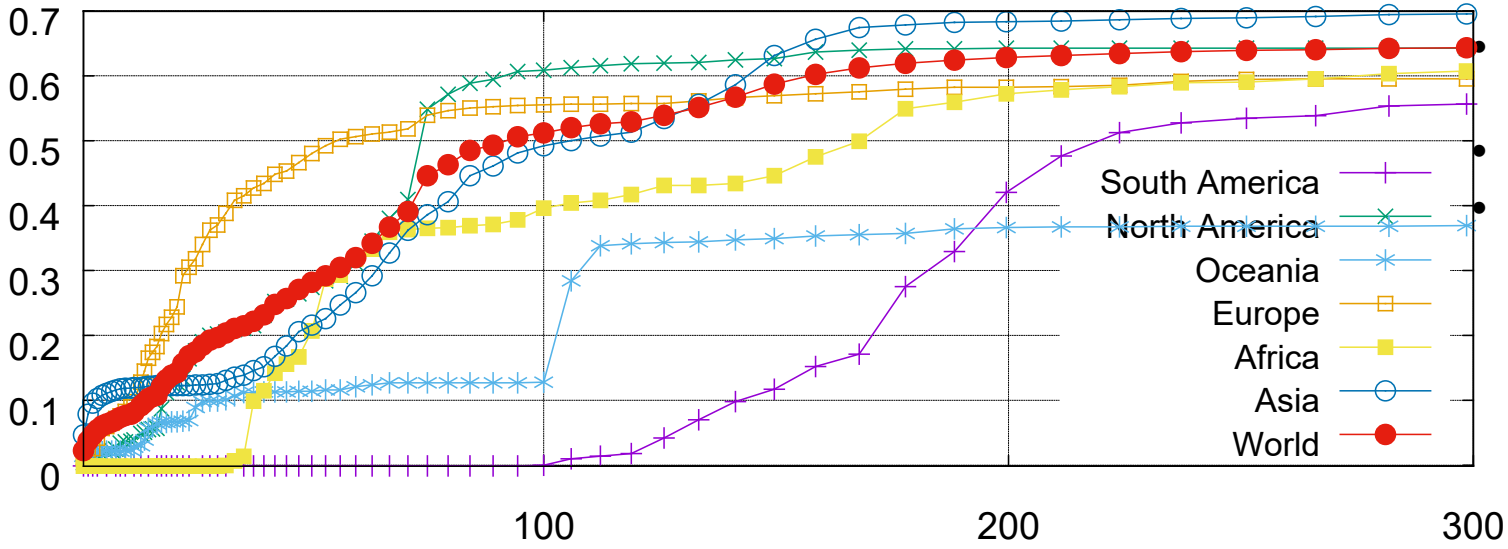


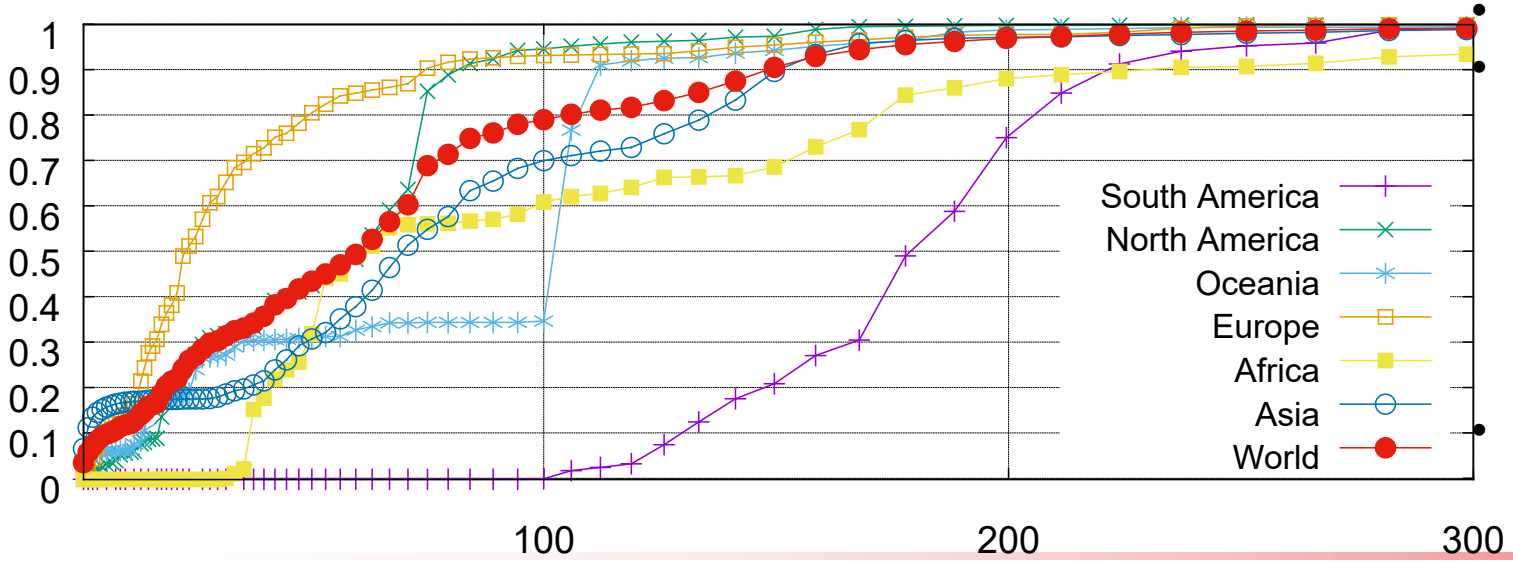
Queries/average delay by continents



Queries/average delayt by countries

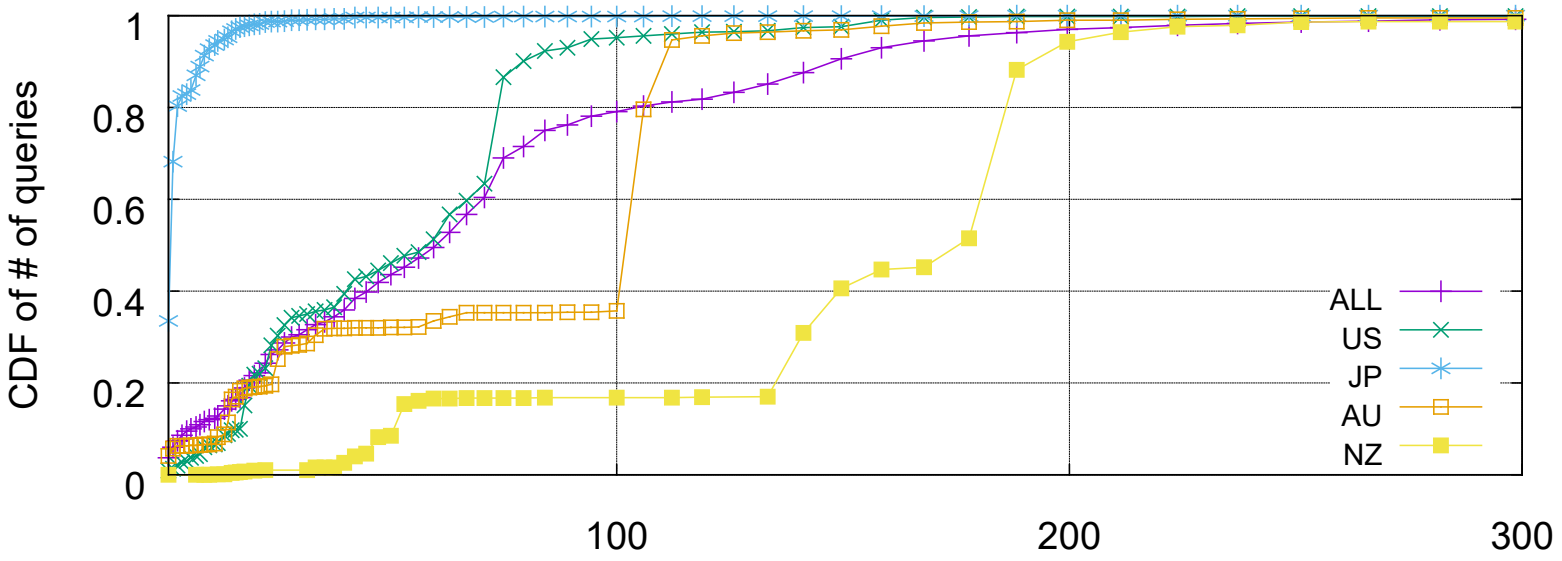
M-Root2021: RTT / query CDF by continent

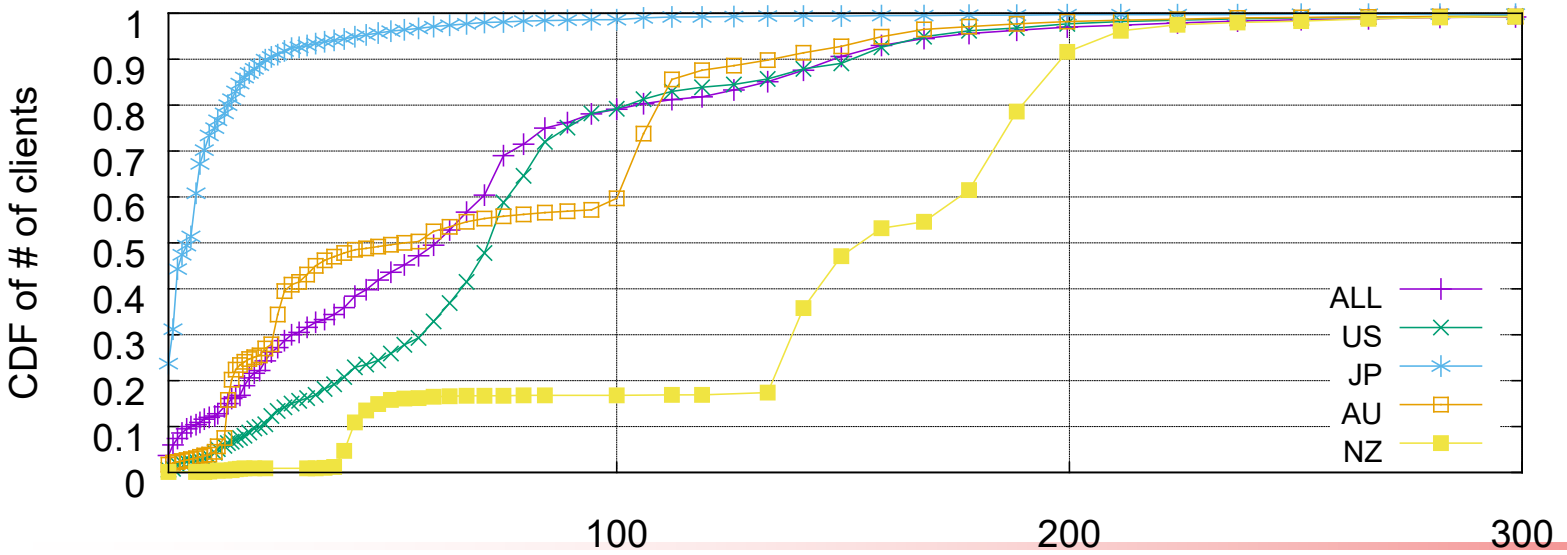




- Vertical axis: Cumulative Distribution function of number of queries
- Horizontal: RTT (milli-seconds)
- The graph above contains queries without TCP RTT data
  - Queries between each client and root node are UDP only or TCP without complete packet sequences (35% of queries)
  - For a quick comparison of RTT, the following CDF graph ignores data without RTT
- 65% of queries have RTT information
- M-Root offers short (<100ms) RTT for over 90% of queries from North America and Europe
  - Over 70% of queries from Asia
  - Over 60% of queries from Africa
    - However, Africa is large from north to south
    - over 10% of queries from Africa have over 200ms RTT
- Over 100ms RTT for South America and 35% of queries from Oceania

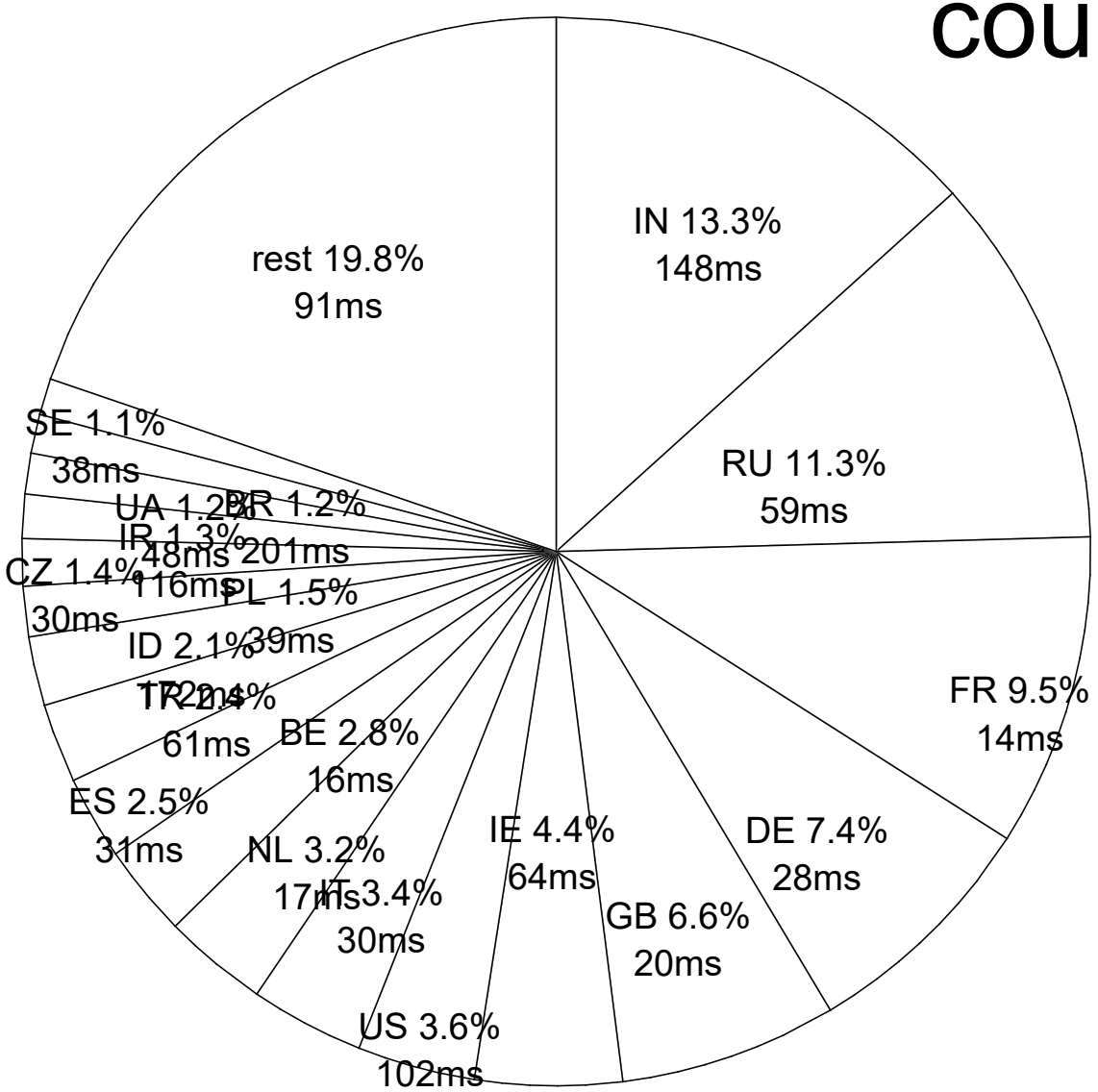
M-Root2021: RTT CDF by queries by countries





- World 79% Q, <100ms
  - 71% IP, <100ms
- JP: > 92% Q, <10ms
  - > 73% IP, < 10ms
- US: >95% Q, < 100ms
  - over 79% IP, < 100ms
- AU
  - 19%Q, 25%IP <20ms,
  - 32%Q, 49%IP <50ms
  - 35%Q, 57%IP < 100ms
  - Others ≧ 100ms
- NZ
  - 16%Q, 16%IP < 60ms
  - Others > 150ms

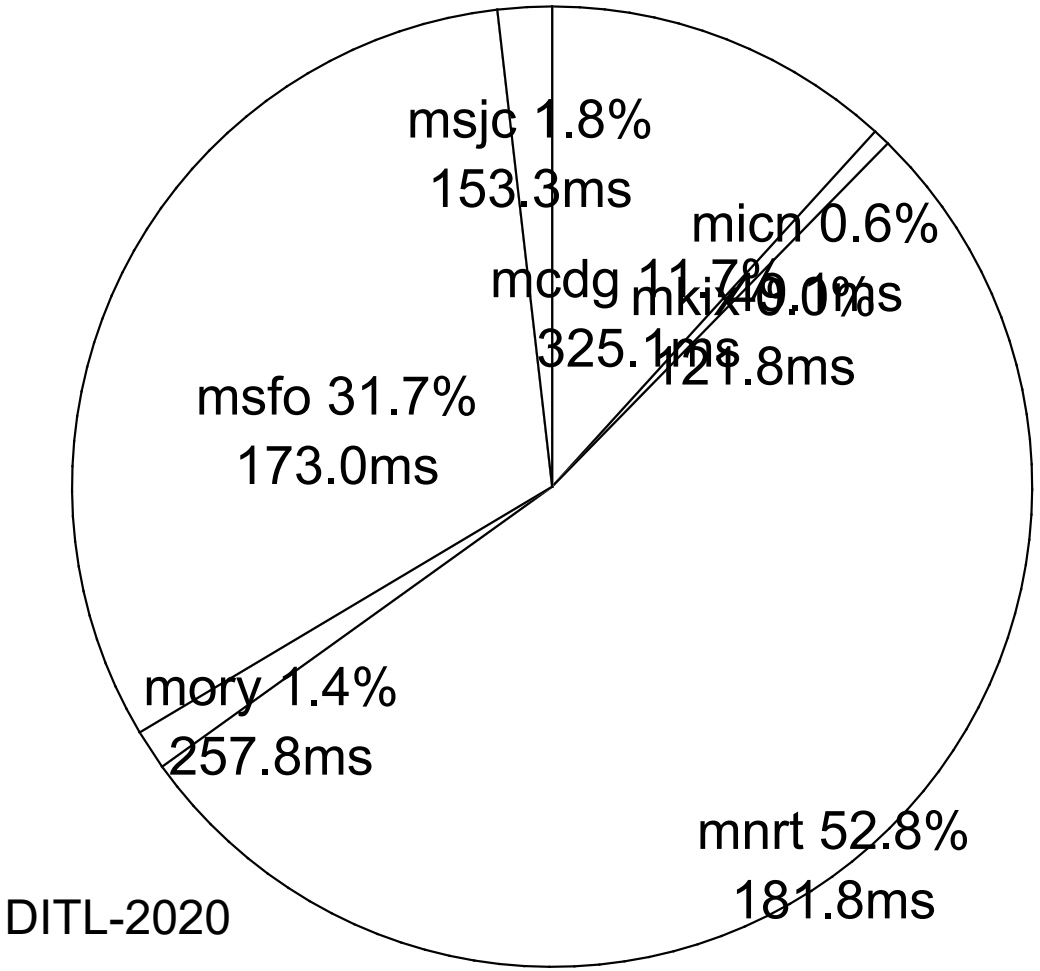
M-root Paris nodes: queries, average RTTs by countries

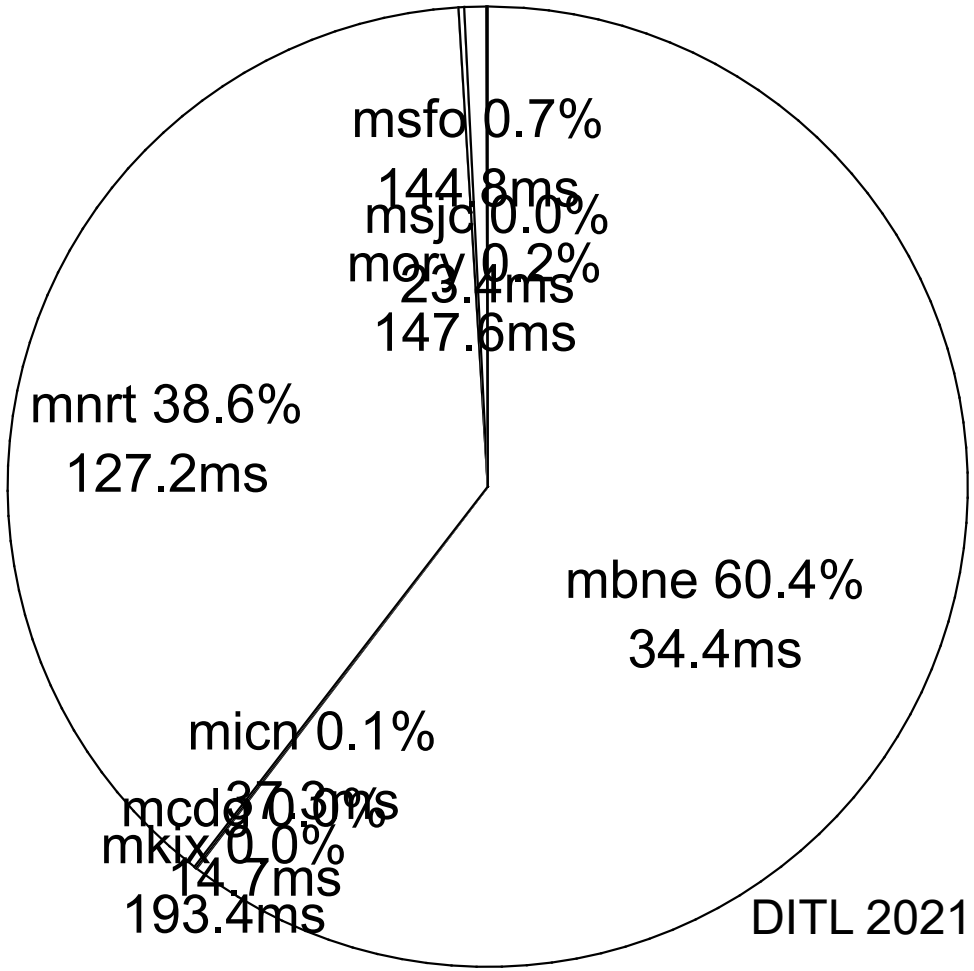


- IP anycast seems to work well
- Paris nodes provide short RTTs for Europe

Effectiveness of M-root Brisbane, AU node

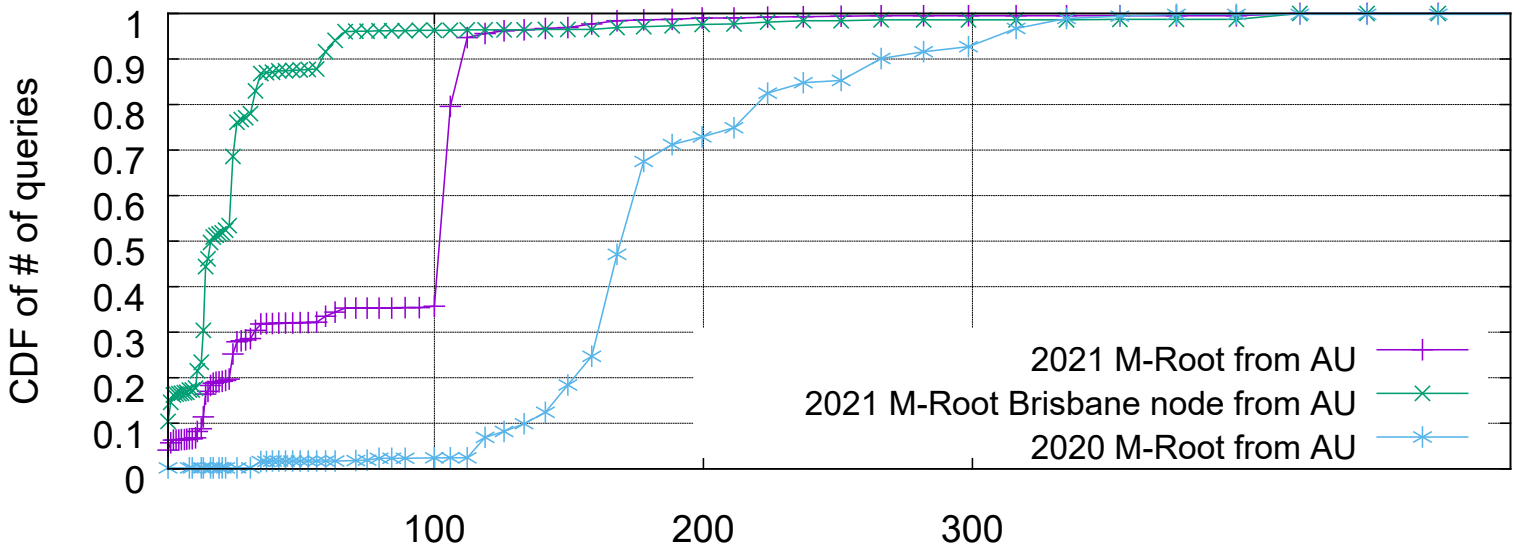
M-Root Instance Launched in Brisbane on 18, Dec, 2020

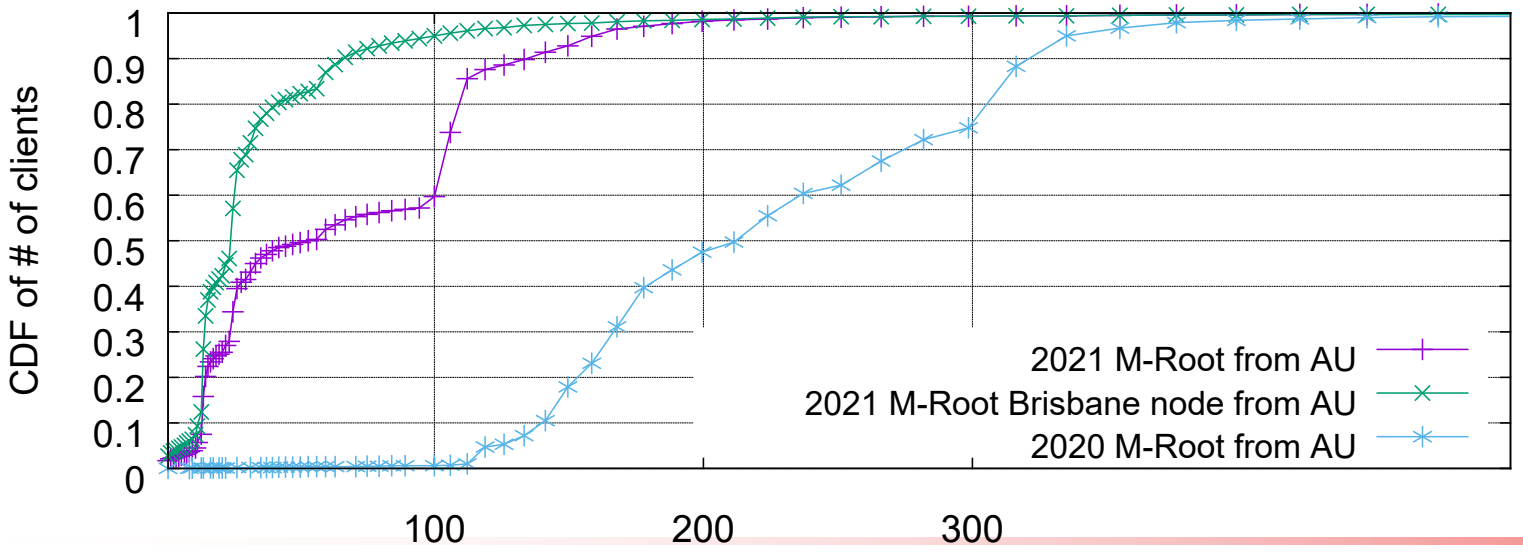




DITL-2020: Before launching Brisbane node, 52.8% queries from Australia are sent to Tokyo node (Ave RTT 181ms)
DITL-2021: 60% of queries from Australia to M-root are sent to Brisbane node, average RTT is 34.4ms

M-Root RTT CDF from Australia





- RTTs from Australia to M-Root was over 120ms in 2020
  - 8% Q, 25% IP, >300ms
- After launching Brisbane node, 28% of queries, 41% of IP addr are answered within 30ms, 95% of Q, 85% of IP are within 112ms

Summary

- Anycast node name and query latencies can be measured from DITL dataset
- M-Root anycast seems to work well in Japan, North America, Europe, and Australia
- M-Root Brisbane node provides short RTT for Australia

- IP anycast effectiveness of root servers may be analyzed in this method
- This method may find the under-served area of root servers

Acknowledgement

- DNS-OARC as the data source of Root DITL dataset
- This research uses GeoIP Data created by MaxMind