M-Root Deployment

Akira Kato

Keio Univ./WIDE Project
kato@wide.ad.jp

Review of Root DNS Servers

- **First entry point of DNS resolution process**
  - This is true
- **But DNS highly depends on "cache"**
  - If you ask example.com
    - info for com servers remains in the cache
    - no need to send query to root for any .com names
    - until TTL expires (172800sec = 2 days) for TLDs
- **Half of the queries are for "wrong" names**
  - Root servers return "NXDOMAIN"
  - "example.comm", for example
  - Negative cache : limited duration and exact names
    - Negative cache doesn’t work for "examples.comm" query

Root Servers Brief History

- ☆ **1987:**
  - 7 Root, 10 Addr, all US
- ☆ **1990:**
  - 8 Root, 11 Addr, NIC.NORDU.NET
- ☆ **1995:**
  - 9 Root, 9 Addr, "Root-Servers.net"
  - Discussion for exansion, to 13
- ☆ **1997:**
  - 11 Root, 11 Addr, "J/K at InterNIC"
  - 13 Root, 13 Addr, "L/M at ISI (IANA)"
  - May?: K Root moved to LINK operated by RIPE NCC
  - Aug: M Root moved to Tokyo operated by WIDE

M-Root History

☆ **1997:**
- Jan: servers (2 PentiumPro 200MHz) ready
- APNIC assigned "202.12.27.0/24"
- Aug: Jon Postel announced: "M-Root moved to Tokyo"
  - root-servers.net and root.cache modified
  - The first Root server in Asia-Pacific Region
  - 600+ query / second

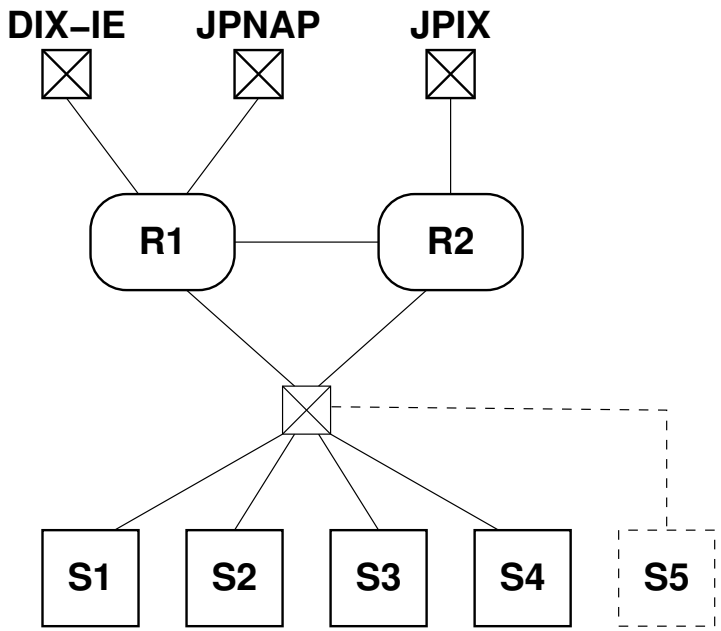
☆ **2002:**
- M-Root backup server in Osaka (not used, just backup)

☆ **2002:**
- M-Root started "Anycast in a rack"

Anycast in a rack

☆ Configuration in 2002:

- 2 Routers, 5 Servers (one spare), 3 IXes

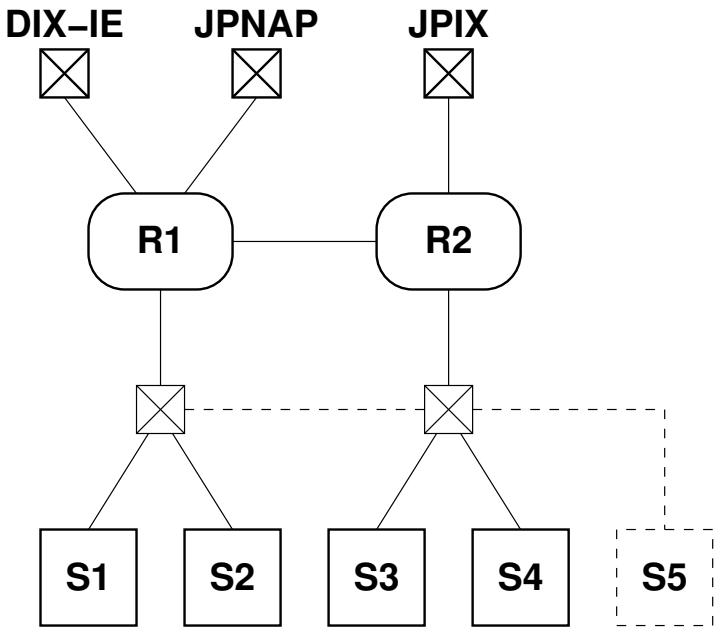


R1: to S1, S2, S3, S4
R2: to S1, S2, S3, S4
S1, S2, S3, S4: to R1
R1 relayed to R2 when necessary

Anycast in a rack

☆ Change the traffic direction in 2002

- Add extra switch



Anycast

- ☆ **Documented in RFC3258 (Apr 2002)**
- ☆ **M-Root "Anycast in a rack"**
  - One of the first cases
- ☆ **Anycast solves many issues**
  - Geographically distributed
    - Reduce RTT (in many cases, if not all)
  - Multiple sites operational in parallel
    - Increases "Capacity" drastrically
  - Answer to "Why no Root DNS in my country?"
- ☆ **But Anycast is NOT a good friend of TCP**
  - TCP can fail if
    - Topology change (short life TCP may be okay)
    - Per-packet load balancing is the enemy

M-Root History (cont)

☆ **2004:**

- Global anycast deployment started
  - Tokyo (3), Osaka, Seoul, San Francisco, Paris
  - Equipment: by WIDE Project
- Some (not all) locations :
  - Colo space, IX ports, transit offered by industry
  - Thank you very much!

☆ **2005:**

- Joint operation with JPRS (.jp registry) started

IPv6 on Root

- **IPv6 glue in Root Zone: Jul 21, 2004**
  - IPv6 glue record (AAAA) of .JP and others registered
  - These TLDs can be accessible via IPv6
  - But IPv4 ware required to get access Root
- **IPv6 addresses in Root Servers**
  - There were loooong discussion
    - If adding AAAA to Root could break DNS
  - SSAC submitted a document in Jan 2007
  - ICANN Board approved by Dec 31st, 2007
  - First 4 Root Servers are IPv6 ready since Feb 4th, 2008
    - F/H/K/M
  - Other Root Servers turned on IPv6 gradually
  - Now all 13 letters support IPv6
    - Only IPv4 is available in some anycast locations

M-Root Traditional Operation Policy

☆ **Equipments**
- Procured by JPRS

☆ **Colo-space**
- Procured by JPRS, or offered by DataCenter/IX

☆ **Access to IX**
- Mostly offered by IXes
- Cross connects are on us

☆ **Connectivity**
- Procured by JPRS, or offered by ISPs

☆ **Limitation**
- Budget offered by JPRS
- Enhanincing footprint was not easy

M-Root History (cont)

☆ **2017:**
- Start discussion with APNIC for possible support
  - Enhance M-Root footprint in (mainly) APNIC region
- "small" anycast equipment procured by APNIC budget
  - One Dell R340/R350 server
  - One switch (Catalyst9200L or similar)
  - One commercial software router license
- Colo, IX, admin access, remote hand provided by "host"

☆ **2020:**
- MoU signed with APNIC
- Brisbane (AU) operational hosted by APNIC

☆ **2021:**
- Hanoi (VN) operational hosted by VNNIC

☆ **2022:**
- Guam (GU) operational hosted by MARIIX/Univ. of Guam
- Kuala Lumpur (MY) operational hosted by MyIX

Position of M-Root in Root DNS Servers

- ☆ **M-Root:**
  - 9 locations (11th letter, [306, 6])
  - 13 sites (9th letter, [306, 6])
    - 9 sites are global, 4 sites are local
  - All sites support IPv4 and IPv6
- ☆ **Number of unique sources seen in 24hr**
  - 2015 Nov: 2.6M IPv4, 0.18M IPv6 (0.13M unique /64)
  - 2022 Jun: 4.9M IPv4, 5.7M IPv6 (1.7M unique /64)
  - Obviously IPv6 is used extensively (at least in DNS)
    - Or due to randomization of lower 64bit? No!
- ☆ **Number of queries seen in 24hr**
  - 2015 Nov: 2.4G queries, 8.4% in IPv6, 0.84% in TCP
  - 2022 Jun: 4.7G queries, 29% in IPv6, 2.9% in TCP
- ☆ **Position of M-Root**
  - 2017 Mar: 4.9% of total queries (except G)
  - 2022 Jun: 5.6% of total queries
- ☆ **Note: numbers based on rssac002 data**

Queries to each of Root DNS Server

- Source: RSSAC002 Jun 2022, one day

L	K	A	J	I	F	E	D	C	B	M	H	G

☆ **Each letter receives 11.3% to 3.2%**

☆ **Note: Modern DNS server implementations**

- Send queries mostly to the nearest servers
  - For better response time
- Often send queries to other servers
  - RTT may change over time

☆ **Only M-Root IP addresses assigned by APNIC**

- ROA diversity

Future plans of M-Root Deployment

- ☆ **In setup**
  - Jakarta (ID), Quezon (PH)
- ☆ **MoU signed, wait for equipment**
  - Bangkok (TH), Dhaka (BD), Kaohsiung (TW)
  - Kathmandu (NP), Ulaanbaartar (MH), Singapore (SG)
- ☆ **In discussion**
  - Mumbai (IN), Kolkata (IN), Beijing (CN), Lahore (PK)
- ☆ **Possible locations in AP Region**
  - Auckland (NZ), Hong Kong (HK), **...**
- ☆ **Possible locations outside of AP Region**
  - Latten America? Africa? Middle-East?

Extra notes

☆ **If you have a concern in DNS Privacy**
- QNAME minimisation (RFC7816) may help
  - Without DoT/DoH
  - At least to Root DNS Servers

☆ **If your AS doesn't have a RIPE Atlas Probe**
- Hosting a probe may generate valuable data
- Contact RIPE/NCC or even APNIC

☆ **If you are interesting to host M-Root**
- write to m-root-eoi@jprs.co.jp