

BGP Anycast Node Requirements for Authoritative Name Servers

March 2, 2006
DNS Operations SIG @ APNIC 21

Yoshiro YONEYA <yone@jprs.co.jp>

On behalf of

Yasuhiro Orange Morishita <yasuhiro@jprs.co.jp>

Shinta Sato <shinta@jprs.co.jp>

Takayasu Matsuura <matuura@jprs.co.jp>

Background

- IP anycast technology is now being deployed in authoritative name servers
 - root servers
 - C, F, I, J, K, M
 - TLD servers
 - Over 50 TLDs DNS servers are now in IP Anycast mesh (really!)
- Some DNS hosting service providers are now using IP anycast technology
 - ISC, PCH, UltraDNS, etc.
 - These serve some TLDs

Motivation

- Increasing physical server nodes without changing DNS protocol
 - Not to exceed UDP packet size limitation
- Increasing tolerance against hardware troubles and/or cyber attacks
 - Damages are localized
- Improvement of query response time from anywhere in the world
- Distribution of load between nodes

Principle

- Technique to increase reliability
- Single contact, consistent policy
 - Never deploy disordered sites
- Use well known, matured techniques
 - Common guidelines (BCP) is required
 - RFC 3258
 - draft-ietf-grow-anycast-03.txt
 - RFC 2182
 - RFC 2870

Our intention

- To have common IP anycast guideline for TLD DNS servers
 - draft-morishita-dnsop-anycast-node-requirements-02.txt
 - Based on JP's experiences
 - Work in progress
- Why TLDs?
 - Due to different features from Root servers:
 - Larger zone size
 - Higher zone update frequency

Targets and main focuses of our document

- Targets
 - BGP anycast
 - Global node
- Main focuses
 - Selection of the Internet service provider (ISP)
 - Selection of the IP anycast node location
 - Evaluation of proper cost
 - Evaluation of proper measurement and monitoring methods

Selection of the ISP

- Requirement
 - To have **geographical and network topological diversity**
- Evaluation points
 - Reliability of the ISP backbone network
 - Connectivity of ISP outside area
 - ISP's peering status
 - Connectivity for DNS service
 - address block and AS number
 - Connectivity for administration
 - Connectivity for IPv6

Selection of the IP anycast node location

- Requirement
 - To conform to requirements of RFC 2182 and 2870
- Evaluation points
 - Security level
 - Redundancy of electric power supply
 - Tolerance against disasters
 - Diversity of locations

Evaluation of proper cost

- Initial (construction) cost and Running (maintenance) cost
 - Equipments (routers, switches, servers, etc.)
 - Facilities (data center, connectivity, etc.)
 - Human resources (daily, emergency, etc.)
 - Serious in remote site

Evaluation of proper measurement and monitoring methods

- ICANN's "CNNP test" is one of useful guideline for validating IP anycast node
- Continuous measurement
 - Routing stability, Reachability
 - Round trip time
- It is hard to make the worldwide measuring points
 - RIPE DNSMON is one of possible solution

Our findings through oversea site

- Running cost is dominant
 - Facilities
 - Human resources and traveling expense for troubleshooting and recovery
- Difference of business practices
 - Based on different commercial law
 - Some data center requires insurance contract
 - Hard to have contract with foreign customer
- Others to remind
 - Overheat due to rack placement
 - Communication with remote hand
 - Shipping for hardware replace

Future Works / Discussions

- Update the guideline and add other effective portions in it
 - Selection of node locations
 - In some places, there are two or more root (and TLD) servers at the same locations / places...
 - Is it really redundant?
 - Selection of server hardware
 - Selection of server software
 - Selection of remote maintenance tool / hardware
 - How to do effective (and reasonable) remote maintenance
 - How to do effective (and reasonable) measurement

Future Works / Discussions (cont.)

- Any other missing points?
- Any other remarkable experiences?
- How can TLDs / *IRs cooperate on this topics?
 - No need?
 - Buy existing services?
 - Information sharing?
 - Operational experiences
 - More concrete formation?

Questions and comments?

- Any comments are welcome
 - E-mail: dnstech-info@jprs.jp

Acknowledgements

- This memo is included in the results of the research activities funded by National Institute of Information and Communications Technology (NICT).