# Using In-bailiwick Nameservers

Masato MINDA

Japan Registry Services Co., Ltd.

NANOG 33

# Introduction

# JPRS and JP DNS

- Japan Registry Services (JPRS)
  - operating .JP registry system.
  - managing and administrating .JP TLD.
- JP DNS
  - authoritative nameservers of .JP
  - managed by JPRS
  - operated by JPRS and other organization.
  - "a.dns.jp" - "f.dns.jp" (part uses anycast.)

# Construction of Zones for JP DNS

- JP DNS consists of  64 zones.
  - in-addr.arpa zone is not included
- JP.
  - Top Level JP Domain for General-Use
- AC.JP. AD.JP. CO.JP. ED.JP. GO.JP. GR.JP. LG.JP. NE.JP. OR.JP.
  - Organizational Type of JP Domain Name
- TOKYO.JP. OSAKA.JP. KYOTO.JP. etc..
  - Geographic Type of JP Domain Name

# Definition of Terms

- In-bailiwick nameserver
  - FQDN with their domain to nameserver
- Out-of-bailiwick nameserver
  - FQDN with outside domain to nameserver
- C-NS
  - Abbreviation of "Caching nameserver"
- A-NS
  - Abbreviation of "Authoritative nameserver"

# Main Subject

# JP DNS changed the way of handling glue in June 2004

- All of Out-of-bailiwick glues were deleted from JP DNS
- Before June 2004
  - JP DNS responds NS RR with its glue automatically in any occasion
- After this change made in June 2004
  - JP DNS responds NS RR with its glue only when the host name of NS RR is accompanied with the parent domain.

# Change of JP DNS

- Change in June 2004 was technically correct
  - See. Section 6.1 in RFC2181
- However, the problem occurs
  - some domains encountered difficulties in name resolution,
    Can not connect to “http://www.example.jp/"
  - One of the most famous WEB sites in Japan
- Issue is affected with this change.

# Glue Handling in JP DNS (1/3)

- Case A: NS RR is internal hostname accompanied with its domain which belongs to the same zone (zone. "example.co.jp")

```
example.co.jp.     IN NS  ns.example.co.jp.
ns.example.co.jp. IN A   10.10.10.10
```

- JP DNS can provide “ns.example.jp” as NS RR with its glue, 10.10.10.10 when the C-NS sends the query of A RR for “www.example.co.jp” to JP DNS.
- "example.co.jp" uses in-bailiwick names for its nameserver

# Glue Handling in JP DNS (2/3)

- Case B-1: NS RR is hostname belongs to different zone

  `example.CO.jp. IN NS  ns.example.NE.jp.`

- JP DNS can provide only NS RR
  - JP DNS can not provide the glue of NS RR
  - A-NS on BIND 8 sometimes responds glue on this case according to the condition.

# Glue Handling in JP DNS (3/3)

- Case B-2: One of NS RR is hostname belong to the same zone and has different

```
example.co.jp.          IN NS xx.example-xx.co.jp.
example-xx.co.jp.       IN NS ns.example-xx.co.jp.
ns.example-xx.co.jp. IN A  10.12.34.56
```

- JP DNS can provide only NS RR
  - JP DNS can not provide the glue of NS RR

# Issue of Caching Nameserver with BIND 8

- C-NS with BIND 8 up to version 8.2.7
  - C-NS fails in name resolution when it receives the result of query without glue twice continuously.
- C-NS with BIND 8 version 8.3.0 or later
  - C-NS with BIND 8.3.0 or later can resolve domain name, however it depends on the re-try query from the client.
  - Name resolution becomes slower.(5-10sec)
  - The effective cache makes it be unaware.

# Other Caching Nameservers

- Well-known implementation
  - BIND 9
  - dnscache (djbdns)
  - DNS Service on Windows 2000 Server
- No issue of gluelessness
- Another issue gluelessness causes
  - This will be explained afterwards.

# Before and After at June 2004 .JP

| Count* | BIND 8 JP DNS | | BIND 9 JP DNS | |
|---|---|---|---|---|
| | Before | After | Before | After |
| 1 | 98.206% | 75.378% | 60.864% | 53.292% |
| 2 | 1.794% | 24.539% | 34.316% | 39.828% |
| 3 | 0.000% | **0.082%** | **4.778%** | **6.821%** |
| 4 | 0.000% | **0.001%** | **0.040%** | **0.057%** |

* Frequency of query C-NS send to JP DNS until obtaining glue

- BIND 8 is predominant
  - most of JP DNS use BIND 8

# After at June 2004 .JP

- After June 2004, 0.083% fails in name resolution from old BIND 8 on C-NS.
  - Although 0.083% is quite small rate, it might be a big problem if the domains frequently accessed are included in this minority rate.
- Old BIND 8 on C-NS still prevails.

# Common example at ISP's setting in .JP

- Many ISPs own more than two domains.
  - example.**NE**.jp for service use to customers
  - example.**AD**.jp for administration use
- Customer’s Domain is “example.**CO**.jp”

# ISP's setting in .JP (cont.)

- NS RR set to customer’s domain is covered with the domain for service use

```
example.CO.jp.      IN NS   ns.example.NE.jp.
```

- NS RR set to domain for  ISP service is covered with domain for administration use.

```
example.NE.jp.      IN NS   ns.example.AD.jp.
```

- NS RR set to administrative domain is covered with the domain for administration use.

```
example.AD.jp.      IN NS   ns.example.AD.jp.
ns.example.AD.jp.   IN A    10.10.10.10
```

- Old BIND 8 on C-NS can not resolve!

# Approach of JPRS

- Propagation and activity to recommend ISP configuration of NS RR with parent domain
  - BAD

    `example.NE.jp IN NS example.AD.jp`
  - GOOD

    `example.NE.jp IN NS example.NE.jp`
- Looking for another solution as JP DNS can not be updated to BIND 9 at this stage.

# Another Issue

# In-bailiwick Nameservers

- Query of A RR for "www.example.jp" when "ns.example.jp" is configured to NS RR for "example.jp"
  1. Refer to root DNS servers IP address of "www.example.jp" and receive hostname and glue for JP DNS.
  2. Refer to JP DNS IP address of "www.example.jp" and receive the hostname("ns.example.jp") of NS RR of "example.jp" and its glue.
  3. Refer IP address of "www.example.jp" to "ns.example.jp" and receive it.
- 3 times of references to A-NS are required.

# Out-of-bailiwick Nameservers(1/2)

- Query of A RR for "www.example.jp" when "jp.example.net" is configured to NS RR of "example.jp".
  1. Refer IP address of "www.example.jp" to root DNS servers and receive hostname of JP DNS and its glue.
  2. Refer IP address of "www.example.jp" to JP DNS and receive only hostname of NS RR for "example.jp" without IP address. - **Gluelessness**
  3. Refer IP address of "jp.example.net" to root DNS servers and receive hostname of NS RR of net ([a-m].name-servers.net) and its glue.

# Out-of-bailiwick Nameservers(2/2)

4. Refer NS RR of "ns.example.net" to NET DNS and receive hostname (ns.example.net) of NS RR of "example.net" and its glue.
5. Refer IP address of "jp.example.net" to "ns.example.net" and receive it.
6. Refer IP address of "www.example.jp" to "jp.example.net" and receive it.

- 6 times of references to A-NS are required.

# Live Example

- Query A RR of "www.good.co.dnslab.jp"
  - without any problem
  - get IP address  immediately
- Query A RR of "www.bad1.co.dnslab.jp"
  - Configured to respond NS RR with gluelessness once.
  - BIND 8 on C-NS can resolve, it takes time.

# Live Example (cont.)

- Query of A RR of "www.bad2.co.dnslab.jp"
  - Configured to respond NS RR with gluelessness continuously twice
  - Old BIND 8 C-NS can not resolve.
- TTL is set at 20sec
  - Experiment query should be sent  at interval of 20 seconds.

# Summary of Gluelessness Issue

- When NS RR has no glue, C-NS need to access to many A-NS.
  - Name resolution becomes slow
  - Increase of traffic for the DNS query
  - Decline of stability, reliability and maintenancibility in a trouble.
- When DNS query ends in gluelessness for NS RR **continuously twice**, BIND 8.2.x(and older) C-NS fails in name resolution.

Appendix

# *.ARPA issue

# *.ARPA issue

- Reverse DNS lookup of IN-ADDR.ARPA zone is slow!
  - Supposed to be the case is issue of LAME delegation.
- Actual case is gluelessness issue given to BIND 8 on C-NS
  - Most of nameservers in ARPA zone are out-of-bailiwick names and this causes gluelessness.
  - BIND 9 can make reverse DNS lookup much faster than BIND 8 in most cases on C-NS.

# *.ARPA issue (cont.)

- Glue is not recommended on the point of DNS operation.
  - If "e164.arpa" is used practically in the near future, it'll become a big problem.
  - Using "ns.1.8.e164.arpa" is much better than "a.dns.jp".
- This might have to be discussed at IETF, not at NANOG?

# ISC Response

- This is a problem caused by older BIND8 and BIND4
- All known workarounds involve providing more glue
  - In-bailiwick glue is an appropriate workaround
  - Additional "sibling" glue might also help.
    Eg, if a delegation from .CO.JP includes a NS in FOO.BAR.NE.JP, include glue for FOO.BAR.NE.JP in the .CO.JP zone.
- We regret this inconvenience

# Thank You and Q & A


http://jprs.co.jp/