Using In-bailiwick Nameservers in .ARPA

- Improving reverse DNS lookup performance -

Masato MINDA <minmin@jprs.co.jp>

Japan Registry Services Co., Ltd.

RIPE 50 Stockholm, Sweden

Topics

- Glueless issue in .JP
- The current reverse DNS situation
- Improving reverse DNS lookup performance

Definition of Terms

- In-bailiwick nameserver
  - FQDN with their domain to nameserver
  - Glue is necessary in delegation
- Out-of-bailiwick nameserver
  - FQDN with outside domain to nameserver
- C-NS
  - Abbreviation of "Caching nameserver"
- A-NS
  - Abbreviation of "Authoritative nameserver"

.JP case

Outline of a problem in .JP

- JPRS changed the way of handling glue in June 2004
  - This change is described in RFC2181 Section 6.1
- All of Out-of-bailiwick glues were deleted.
  - As a result, glueless delegations are increased.
- But this causes a problem
  - Some domains make difficulty of name resolution
    - It includes one of the most famous WEB sites in Japan
- We found a problem in BIND8(and older) C-NS behavior about glueless delegation processing

BIND 8 Caching Nameserver Behavior (1/2)

- In iterative query, C-NS starts name server hostname resolution at glueless delegation.
- But if all name servers are glueless and all IP addresses are unknown (not in cache), C-NS stops first iterative query and does not answer anything.
- After timeout (5 or 10 sec), stub resolver or application re-tries querying to C-NS.

BIND 8 Caching Nameserver Behavior (2/2)

- Before then, glueless nameserver addresses may be in cache.
- As a result, name resolution becomes slower for waiting timeouts. (5-30sec)
- At the second time (and after that) DNS query, the DNS cache works well, therefore the problem has been hided.

A worst case: BIND 8.2 Caching Nameserver Behabiour

- BIND, up to version 8.2.7 (including BIND 4)
- Name resolution fails when glueless delegations twice continuously.

```
EXAMPLE.JP  IN NS NS1.EXAMPLE.COM      glueless once
EXAMPLE.COM IN NS NS1.EXAMPLE.ORG      glueless twice
EXAMPLE.ORG IN NS NS.EXAMPLE.ORG       with glue
```

  - Old BIND can not resolve “example.jp”
- This is a problem caused by older BIND8 and BIND4. - from ISC

Live Example

- A RR of "www.good.co.dnslab.jp"
  - without any problem
  - get IP address immediately
- 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.)

- A RR of "www.bad2.co.dnslab.jp"
  - Configured to respond NS RR with gluelessness continuously twice
  - Old BIND 8 (and BIND 4) on C-NS can not resolve.
- TTL is set at 20sec
  - Experiment query should be sent at interval of 20 seconds.

Other Caching Nameservers Behavior

- The well-known implementations
  - BIND 9
  - dnscache (djbdns)
  - Windows DNS service (2000/2003 server)
- There are no issues of gluelessness

The current reverse DNS situation

Common DNS Operational and Configuration Errors (RFC1912)

- In section 2.3 "Glue A Records"
  - `You shouldn't have any A records in an in-addr.arpa zone file (unless you're using RFC 1101-style encoding of subnet masks).`
- This is assumption of usual naming, and this is technically correct.
  - Eg; ns.ripe.net, ns1.apnic.net, ns01.jprs.co.jp. etc. . .

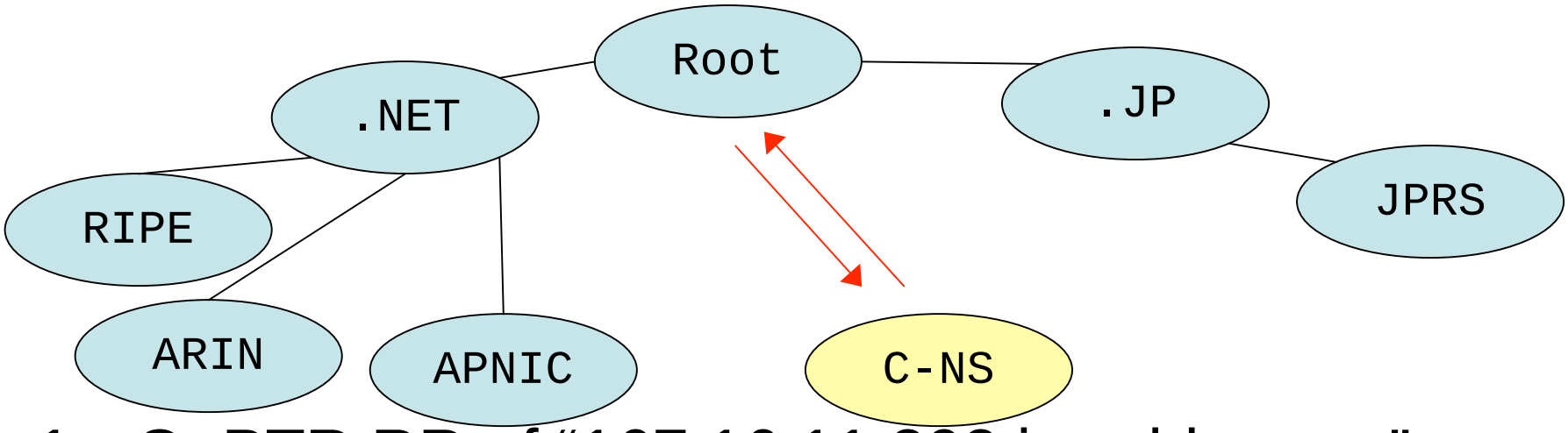
RIRs/NIRs/LIRs reverse DNS registration

- E.g.; APNIC Reverse DNS Delegation Form
  - http://ftp.apnic.net/apnic/docs/reverse-dns
  - Nserver object
    ```
    List of nameservers for a domain
    object; a minimum of two is mandatory.
    Please use fully qualified domain name
    (FQDN) of the nameserver and not the
    IP address.
    ```
- Reverse DNS registration is limited to FQDN which is outside of "in-addr.arpa" zone.
- As a result, reverse DNS lookup is always glueless.

Why is reverse DNS lookup slow?

- In many cases, reverse DNS lookup is slower than standard DNS lookup.
- The LAME delegation is thought of the most popular cause of this.
- But glueless delegation is certainly the one of the biggest cause of this slow DNS lookup.
  - Most of nameservers in ARPA zone are out-of-bailiwick names and this causes gluelessness.
  - BIND 9 C-NS can make reverse DNS lookup much faster than BIND 8 C-NS in most cases

hostname of 202.11.16.167 (1)



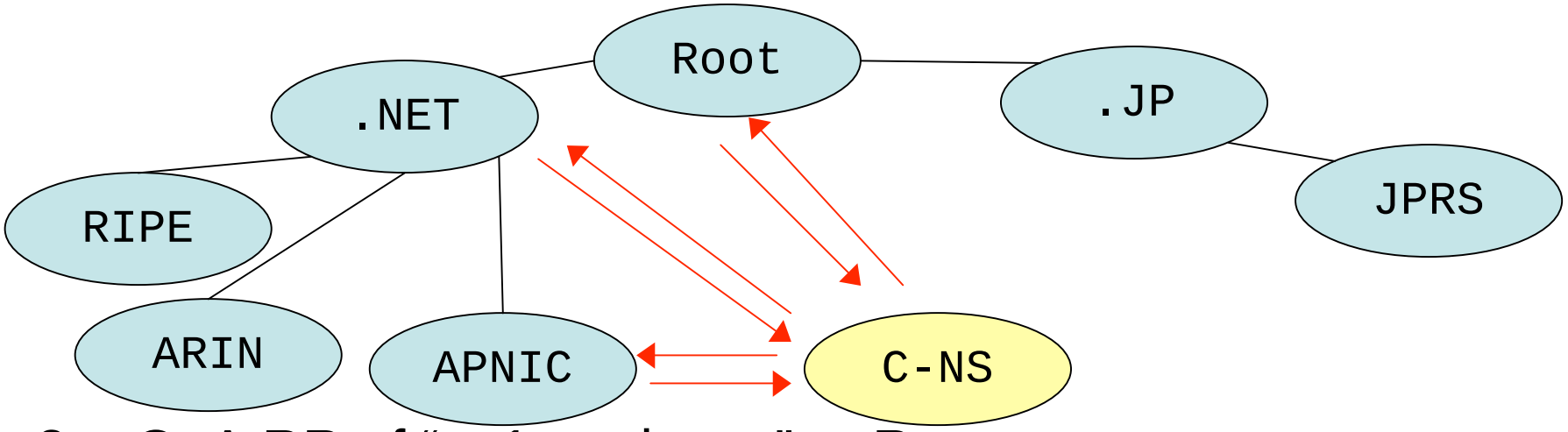
1. Q: PTR RR of “167.16.11.202.in-addr.arpa.”

   A: NS of “202.in-addr.arpa”

   ns.ripe.net. ns1.apnic.net. ns3.apnic.net. ns4.apnic.net. dns1.telstra.net. tinnie.arin.net.

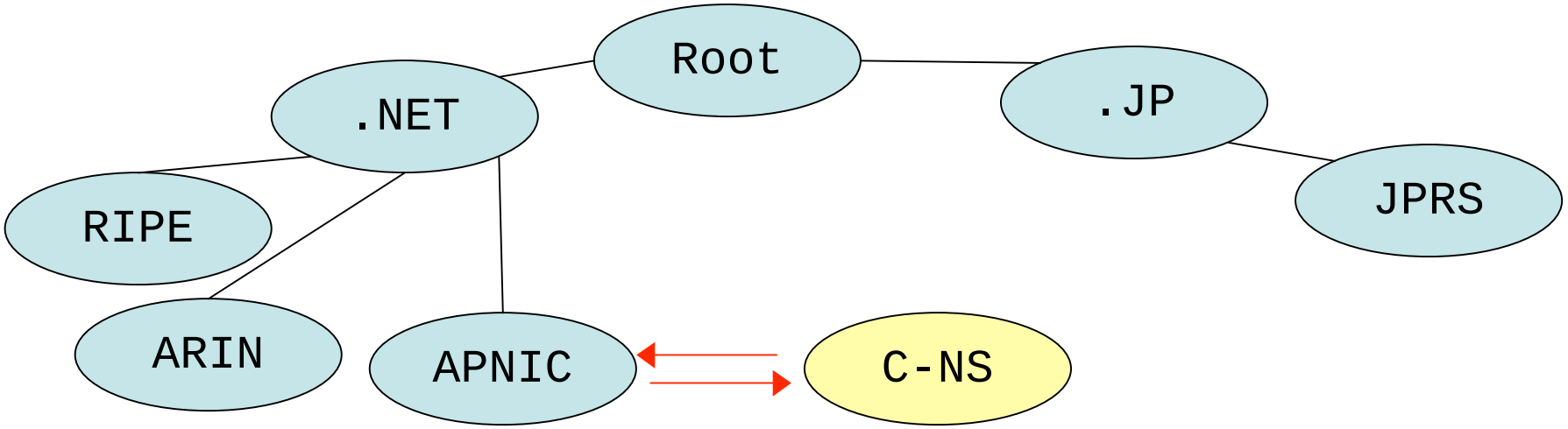
   – **Glueless!** -> **BIND 8 C-NS causes client TIMEOUT!**

hostname of 202.11.16.167 (2)



2. Q: A RR of “ns1.apnic.net” to Root
   A: .NET namesevers with glue
3. Q: A RR of “ns1.apnic.net” to .NET NS
   A: APNIC NS with glue
4. Q: A RR of “ns1.apnic.net” to APNIC NS
   A: A RR of “ns1.apnic.net”
   – Got One of “202.in-addr.arpa” NS’s address.

hostname of 202.11.16.167 (3)

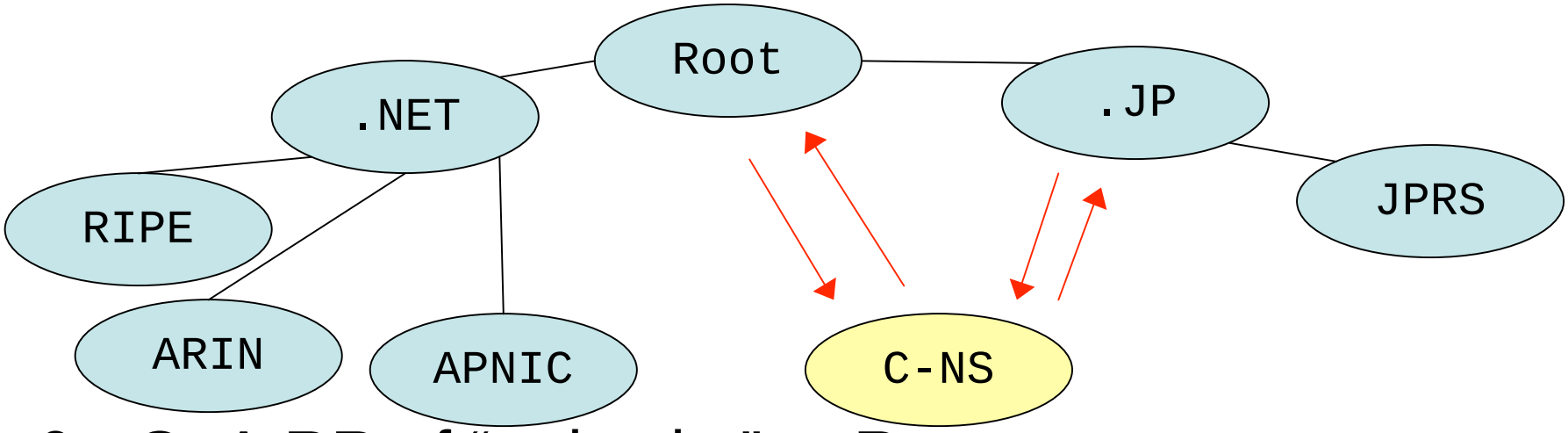


5. Q: PTR RR of “167.16.11.202.in-addr.arpa.

   A: NS of “11.202.in-addr.arpa.”.
   a.dns.jp. b.dns.jp. d.dns.jp. e.dns.jp. f.dns.jp

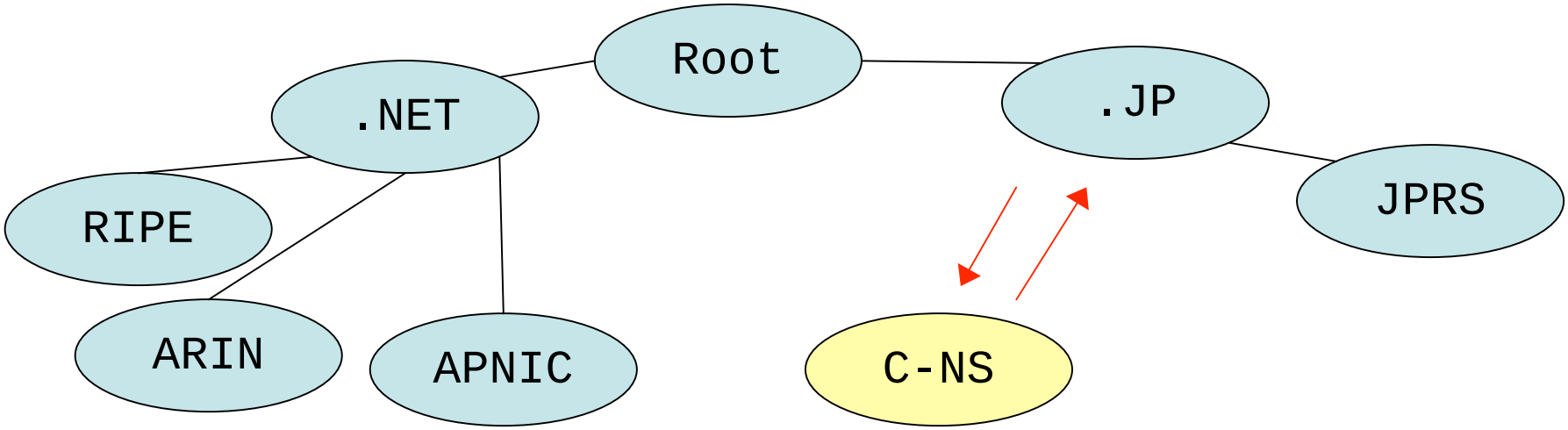
   – **Glueless! -> BIND 8 C-NS causes client TIMEOUT!**

hostname of 202.11.16.167 (4)



6. Q: A RR of “a.dns.jp.” to Root.
   A: JP nameservers with glue.
7. Q: A RR of “a.dns.jp.” to .JP NS
   A: A RR of “a.dns.jp.”
   – Got One of “11.202.in-addr.arpa” NS

hostname of 202.11.16.167 (5)

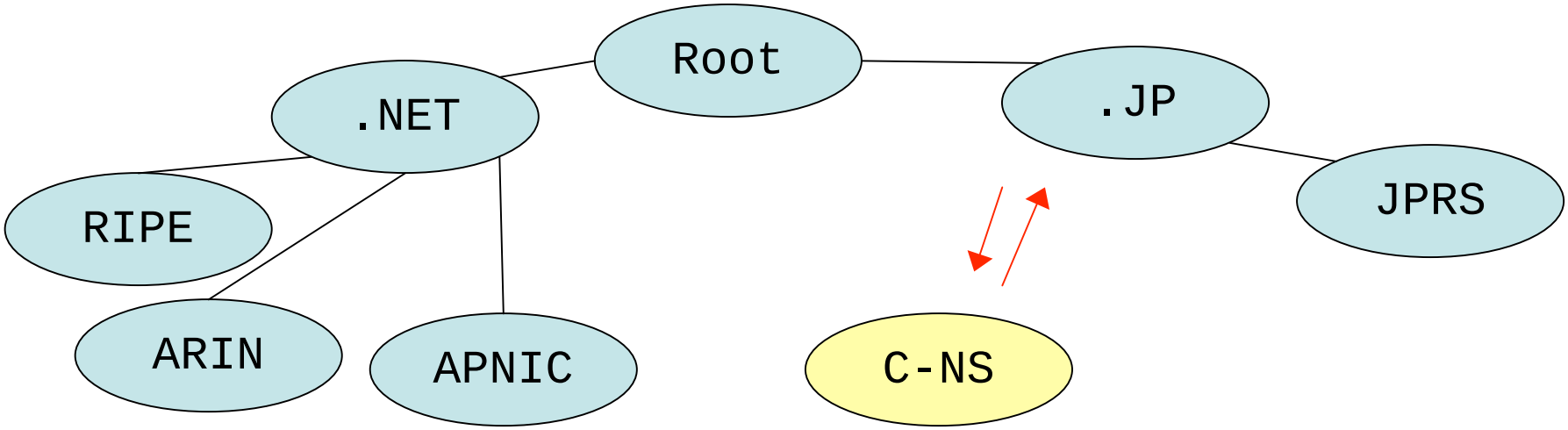


8. Q: PTR RR of “167.16.11.202.in-addr.arpa.” to .JP NS.

   A: NS of “16.11.202.in-addr.arpa.”
   ns01.jprs.co.jp ns02.jprs.co.jp.

   – **Glueless!** -> **BIND 8 C-NS causes client TIMEOUT!**

hostname of 202.11.16.167 (6)

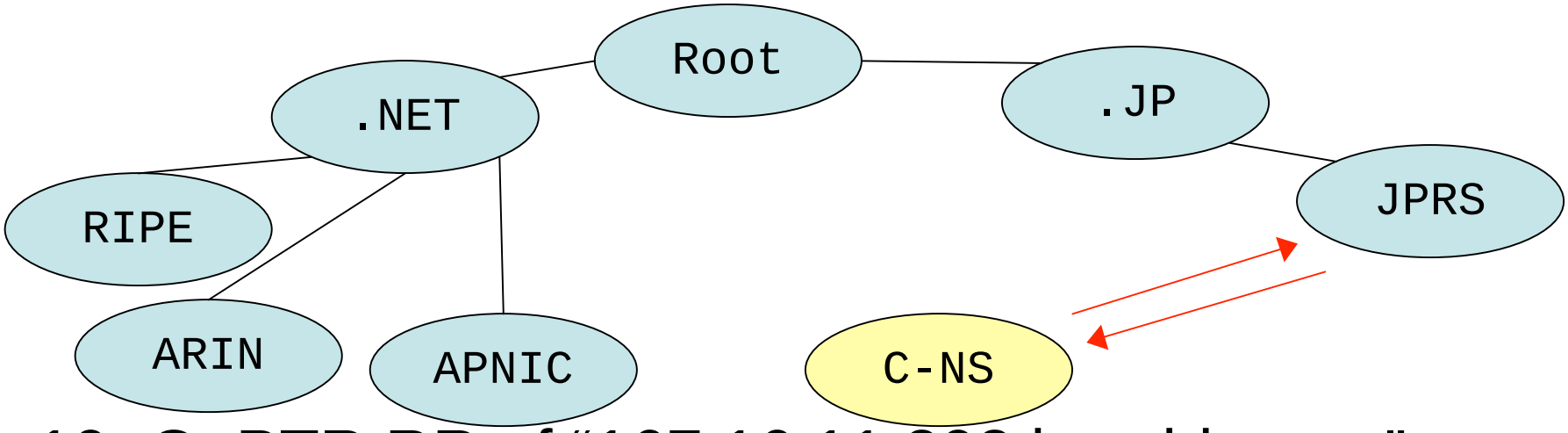


9. Q: A RR of “ns01.jprs.co..jp.” to .JP

   A: JPRS nameservers with glue

   – Got One of “16.11.202.in-addr.arpa” NS

hostname of 202.11.16.167 (7)



10. Q: PTR RR of “167.16.11.202.in-addr.arpa” to JPRS NS

    A: 167.16.11.202.in-addr.arpa. IN PTR jprs.jp.

    Result: “jprs.jp.”

hostname of 202.11.16.167 (8)

- As a result, this reverse DNS lookup requires 10 A-NS queries.
- In BIND 8 case, 3 client timeouts occur.
  - “dig” case, default timeout is 5 sec,
    5 sec * 3 = 15 sec
- CIDR delegation (especially by using CNAME) needs more queries.
- Real C-NS resolves multiple NS’s addresses.

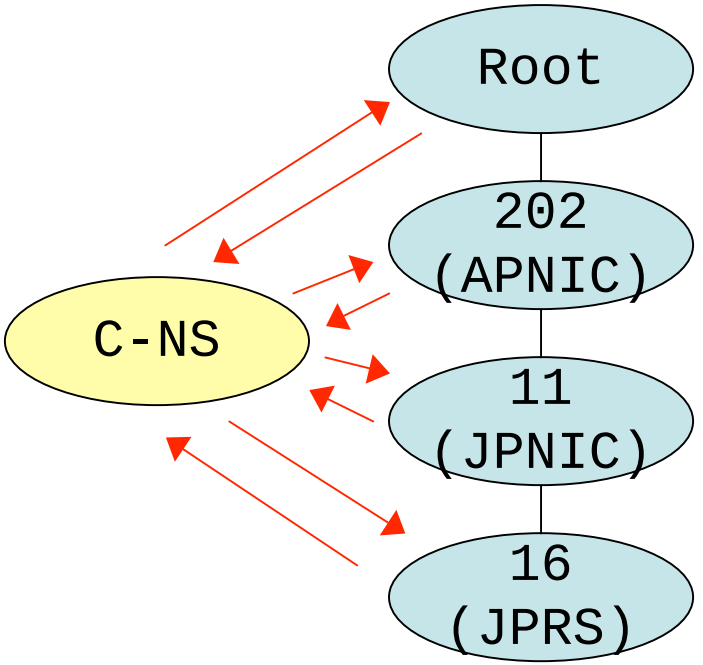
Improving reverse DNS lookup performance

Avoid glueless delegation

- My recommendations:
  - Use In-bailiwick nameservers in .ARPA
  - Add glue information to reverse DNS
- For example, 202.11.16.0/24 case
  - 16.11.202.in-addr.arpa domain's nameserver:
    A.NS.16.11.202.in-addr.arpa.
    B.NS.16.11.202.in-addr.arpa.
  - A.NS.16.11.202.in-addr.arpa glue A: 202.11.17.107
  - B.NS.16.11.202.in-addr.arpa glue A: 202.11.17.227

In-Bailiwick Nameservers in .ARPA
PTR RR of 167.16.11.202.in-addr.arpa

1. Root server answers APNIC server [202.in-addr.arpa] with glue
2. APNIC server answers JPNIC server [11.202.in-addr.arpa] with glue
3. JPNIC server answers JPRS server [16.11.202.in-addr.arpa] with glue
4. JPRS server answers 167.16.11.202.in-addr.arpa PTR.



Only 4 Times!

No BIND 8 Timeout!

In-Bailiwick Nameservers Benefits

- Decreasing resolving cost
- Decreasing resolving time
- Using In-bailiwick nameservers removes a dependency of TLD's DNS tree.
  - Only depends on root servers and .ARPA DNS tree
  - It makes easy to troubleshoot.

Another Points

- ENUM
  - Using in-bailiwick nameservers on e164.arpa zone is very useful to resolving.
- DNSSEC
  - Using in-bailiwick nameservers with DNSSEC is much reduce the cost of verify on C-NS.

Required changes

- Registration system
  - To accept In-bailiwick nameservers
  - To accept glue A/AAAA
- Reverse DNS registration policy
- User’s DNS configuration

In-Bailiwick Nameservers Disadvantage

- The RIR's nameservers have a lot of names.
  - E.g.; 193.0.0.193 (ns.ripe.net) have . . .

    ripe.58.in-addr.arpa. ripe.59.in-addr.arpa.

    ripe.60.in-addr.arpa. ripe.61.in-addr.arpa.

    ripe.124.in-addr.arpa. ripe.125.in-addr.arpa.

    etc.
- When the IP address is changed (including other changes), the more attention is needed.
  - But. . .
    Step by step changes are easily.
    Load balancing are easily.

Future Work

- Past and Current Work
  - “Using In-Bailiwick Nameservers”
    Masato Minda, JPRS, NANOG33
  - “Improving reverse DNS lookup performance”
    Kazunori Fujiwara, JPRS, APNIC19
  - RIPE50 :-)
- We need new Internet-Draft about this issue.
  - I will write it until IETF@Paris.

    may be. . .

Questions?

http://jprs.co.jp/

Acknowledgements

- This presentation is included the research activities founded by National Institute of Information and Communications Technology (NICT)