DNSの逆引きはなぜ遅いのか

民田雅人 <minmin@jprs.co.jp>
株式会社日本レジストリサービス
2005/7/29 JANOG16@**福岡**

こんな経験は?

- ssh(telnet)経由でログインしようとしたら、かなりの時間待たされたのちにログインできた
  - telnetでSMTPへの接続等も同様
- ftpサーバにanonymousでログインしようとしたら、1回目はだめで2回目にログインできた
- ご存知の通りDNSの逆引きの影響
- DNSの逆引きに一体何が起きているのか?

なぜ逆引きに時間がかかるのか?

- LAMEサーバ(不良な設定)の影響
  - 権威DNSサーバ(ゾーン情報を保持するコンテンツサーバ)のはずが、実はゾーンを持っていないとか、誤っているとか…
  - DNSサーバが動いていない
  - マシンが動いていない
  - etc...
- LAMEの影響だと考えている人は多い
  - 少なからず影響はあるが、それだけでは無い

逆引きのおさらい

- IPアドレスが10.20.30.40のホスト名
  - DNS的には 40.30.20.10.in-addr.arpa. の PTR RR を検索する
  - 通常は逆引きの結果を得たあと、正引きを検索して一致の確認を行う
- 2回のDNS検索を行うため、単純な正引きに比べて余計な時間がかかる
- しかし実は逆引きそのものが遅い

なぜ逆引きが遅いのか

- DNSサーバのホスト名にns.example.jpのような普通の名前が使われているため
- なぜ普通の名前だと逆引きが遅い?
  - DNS検索のためのキャッシュサーバから権威DNSサーバへのアクセス回数が劇的に増える
  - キャッシュサーバとして利用したBIND 8(BIND 4を含む)の不具合に中る

BIND 8(含BIND 4)の不具合(1/2)
キャッシュサーバとして使った場合

- 名前の検索中にグルーが無い場合
  - www.example.jpのAをJP DNSに問い合わせる

```
$ORIGIN JP.
EXAMPLE.JP.    IN NS   NS.EXAMPLE.NET.
```

- ネームサーバ名が得られグルーが得られない場合(まだns.example.netはキャッシュに無い)、クライアントへの返答を止め ns.example.netのIPアドレスの検索に専念

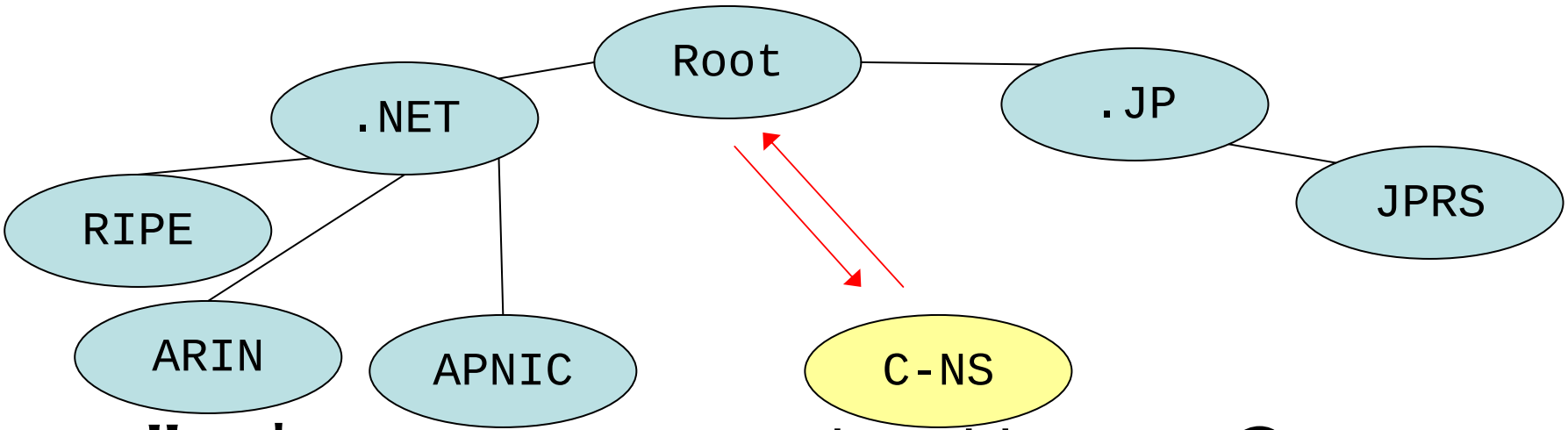
BIND 8(含BIND 4)の不具合(2/2)
キャッシュサーバとして使った場合

- クライアントは返答が無いため、タイムアウトの後に問い合わせを再送する
  - 5〜10秒あるいはそれ以上(クライアントによる)
  - キャッシュサーバのIPアドレスの検索はその間に終了
- 結果として、検索に時間がかかる
- 次の検索ではキャッシュから直ちに答える
  - 通常の運用では、ネームサーバの情報がキャッシュにあることも多く、この問題に気づきにくい

BIND 8.2.7までのBINDの不具合 (BIND 4を含む)

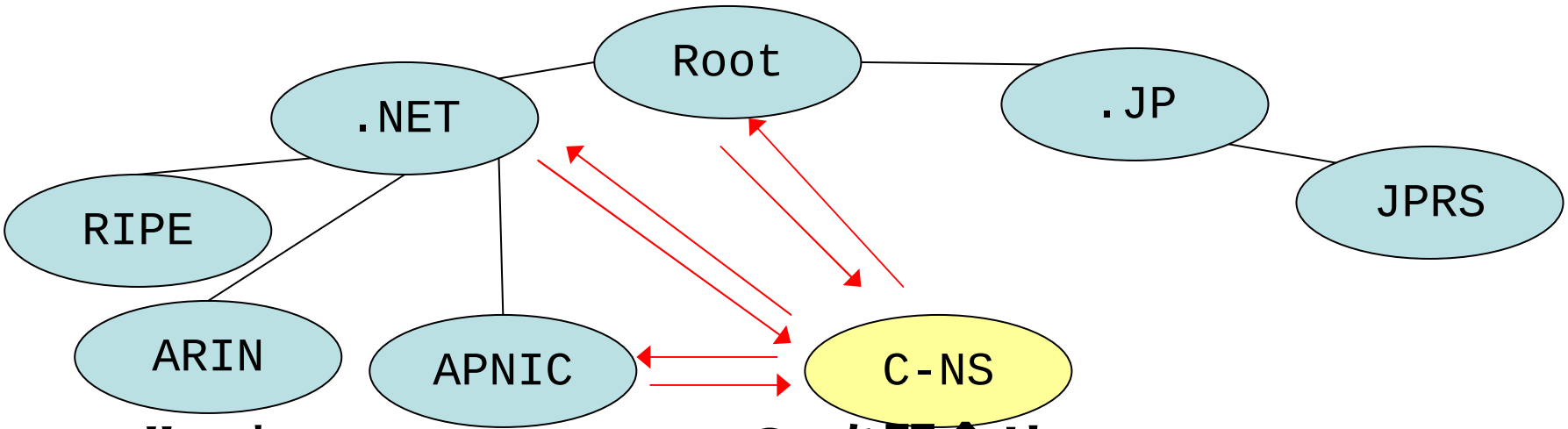
- あるドメイン名の検索中に、グルーが2段連続して得られない場合、BIND 8.2までのキャッシュサーバでは当該ドメインの検索が不能となる
  - BIND 4なども同じ

202.11.16.167のホスト名 (1/8)



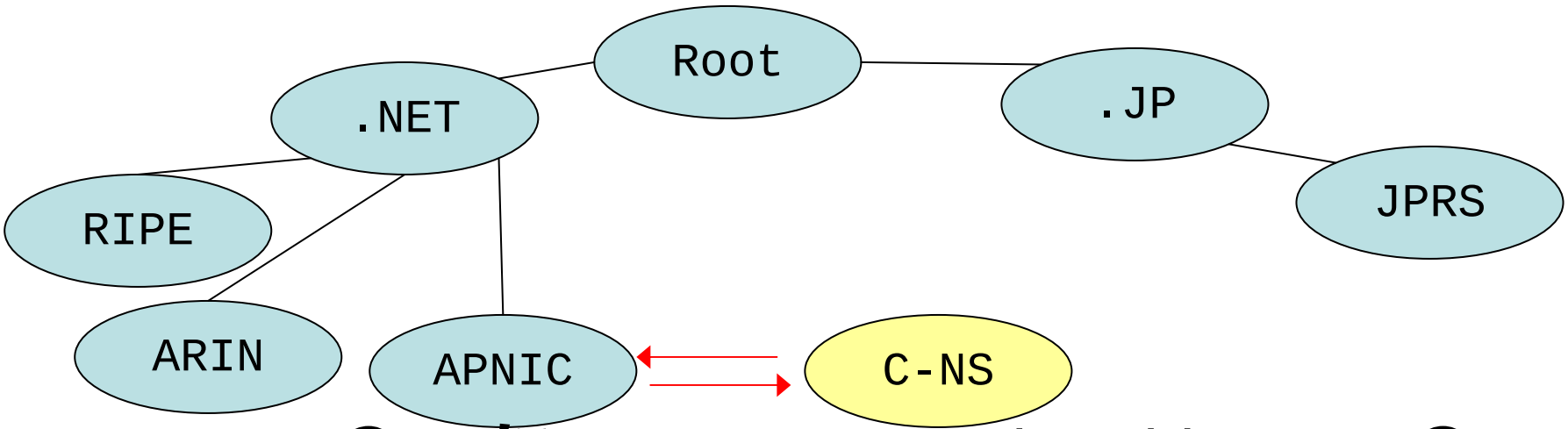
1. ルートへ“167.16.11.202.in-addr.arpa.”のPTR “202.in-addr.arpa”のNSを得る
   ns.ripe.net. ns1.apnic.net. ns3.apnic.net. ns4.apnic.net. dns1.telstra.net. tinnie.arin.net・.
   - グルー無し ⇒ BIND 8 のキャッシュサーバはタイムアウトを起こす

202.11.16.167のホスト名 (2/8)



2. ルートへ “ns1.apnic.net” のAを問合せ
   .NET のネームサーバとグルーを得る
3. .NETのネームサーバへ “ns1.apnic.net” のAを問合せる
   APNICの NS とグルーを得る
4. APNICのNSに “ns1.apnic.net” のAを問合せる
   “ns1.apnic.net”のAを得る
   – “202.in-addr.arpa” のNSのアドレスを得る

202.11.16.167のホスト名 (3/8)



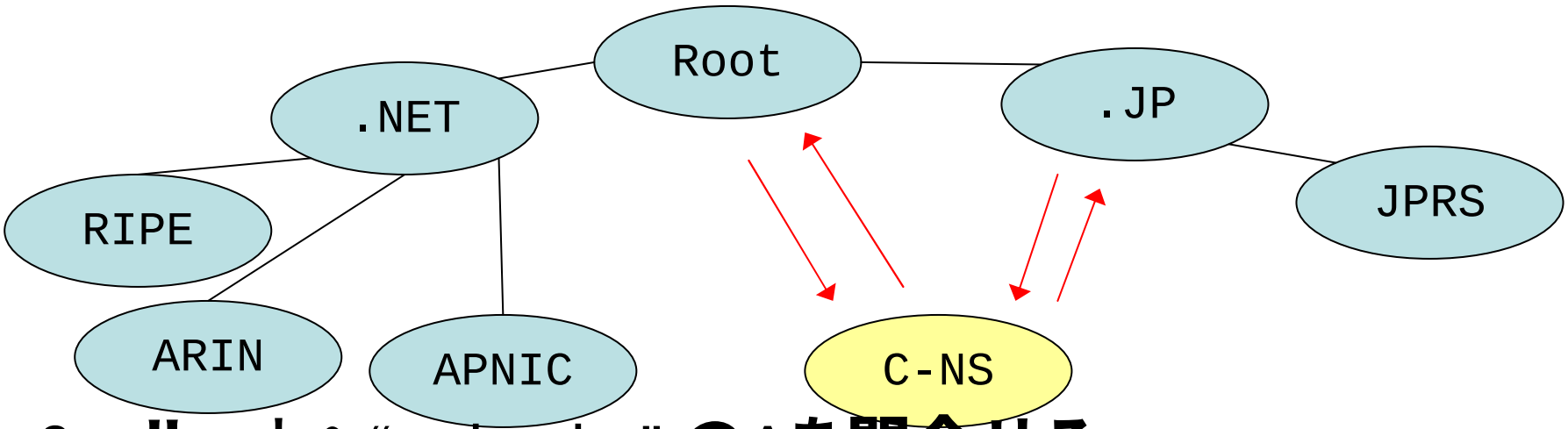
5. APNICのNSに“167.16.11.202.in-addr.arpa."のAを問合せる

   “11.202.in-addr.arpa.”.のNSを得る

   a.dns.jp. b.dns.jp. d.dns.jp. e.dns.jp. f.dns.jp

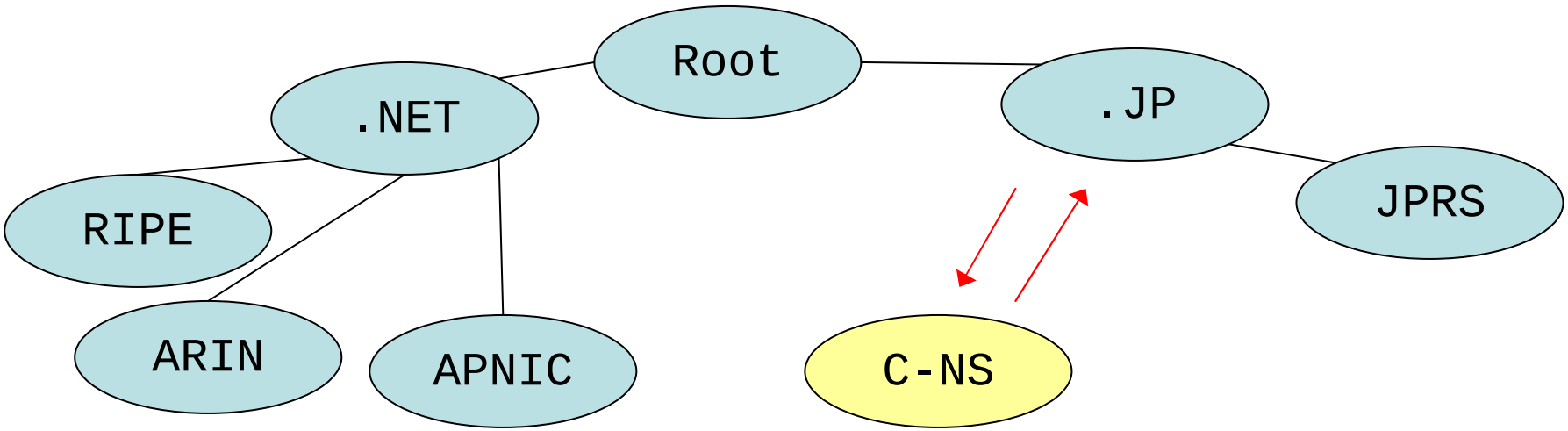
   – グルー無し ⇒ BIND 8 のキャッシュサーバはタイムアウトを起こす

202.11.16.167のホスト名 (4/8)



6. ルートへ“a.dns.jp.” のAを問合せる
   JP のネームサーバとグルーを得る
7. “a.dns.jp.”のAをJP DNSに問合せる
   “a.dns.jp.”のAを得る
   - “11.202.in-addr.arpa” のネームサーバのIPアドレスを得る

202.11.16.167のホスト名 (5/8)

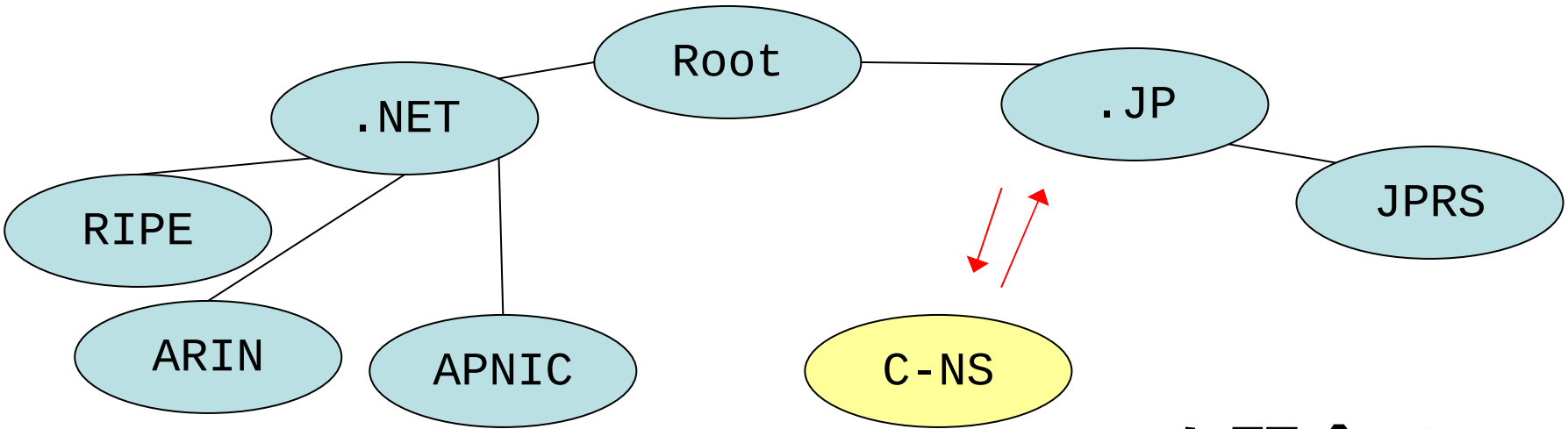


8. JP DNSに “167.16.11.202.in-addr.arpa.”のPTRを問合せる

   “16.11.202.in-addr.arpa.”のNSを得る
   ns01.jprs.co.jp ns02.jprs.co.jp.

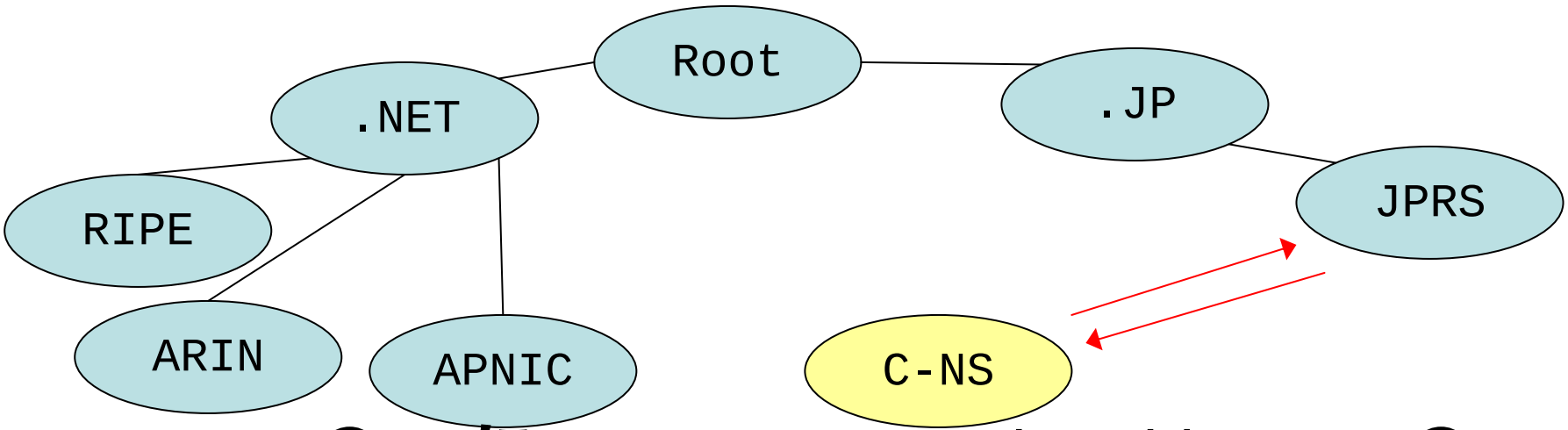
   – グルー無し ⇒ BIND 8 のキャッシュサーバはタイムアウトを起こす

202.11.16.167のホスト名 (6/8)



9. JP DNSに“ns01.jprs.co.jp.”のAを問合せる
   JPRS のネームサーバとグルーを得る
   – “16.11.202.in-addr.arpa”のNSを得る

202.11.16.167のホスト名 (7/8)



10. JPRSのNSに “167.16.11.202.in-addr.arpa” のPTRを問合せる

    以下を得て、

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

    結果として"jprs.jp."と判明

202.11.16.167のホスト名 (8/8)

- この例では10回もの権威DNSサーバへの問合せが必要となる
  - ネームサーバのたらい回し
- キャッシュサーバがBIND 8の場合、クライアントへのタイムアウトが3回発生する
  - “dig” の場合デフォルトのタイムアウトが5秒
    5 秒 × 3 = 15 秒
- /24より小さい割り当てでは、CNAMEが利用され、さらに問い合わせ回数が増える

DNSの逆引きが遅い理由

- 逆引きのネームサーバには一般的なホスト名が使われているため、グルーが得られない場合が多く、DNSツリーの辿り直しが多く発生し、検索そのものに時間がかかる
- BIND 8のキャッシュサーバを使っていて、グルーが得られない場合、タイムアウトを引き起こし、多くの時間がかかる

キャッシュサーバとしての他の実装

- BIND 9
- djbdns(dnscache)
- Windows DNS Service
  - BIND 8**のような問題は無く、逆引き検索は早い**
  - DNS**ツリーの辿り直しが多い問題は解決しない**
- BIND 8(**含む**BIND 4)**のキャッシュサーバはまったくもってお勧めできません**
  - **他にも問題山盛り**
    ⇒機会があれば話したい

改善のためのアイデア

- **一般的な現在の設定**

```
10.in-addr.arpa.     IN NS  ns.example.jp.
```

- **改善案**

```
10.in-addr.arpa.     IN NS  ns.10.in-addr.arpa.
ns.10.in-addr.arpa.     A   10.30.40.50
```

- Internet-Draft**書きました**

  draft-minda-dnsop-using-in-bailiwick-nameservers-01

グルー無しの設定と逆引きのデモンストレーション

- **グルー無しの設定**
  - www.good.co.dnslab.jp **問題無し**
  - www.bad1.co.dnslab.jp **グルー無し1回**
  - www.bad2.co.dnslab.jp 2**回連続でグルー無し**
  - TTL**は**20**秒で設定**
- **逆引き**
  - 210.150.17.51
  - 203.178.129.3

Questions?

http://jprs.co.jp/