Analyse technique du résolveur DNS public chinois 114dns

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 29 décembre 2019. Dernière mise à jour le 6 janvier 2020

https://www.bortzmeyer.org/resolveur-dns-public-chinois.html

Aujourd'hui, nous allons regarder de près un résolveur DNS public, le 114.114.114.114, qui présente la particularité d'être géré en Chine. Il y a de nombreuses choses étranges sur ce service, qui distrairont les techniciens Internet.

Des résolveurs DNS publics, il y en a plein : Google Public DNS <https://www.bortzmeyer.org/google-dns.html>, Quad9 <https://www.bortzmeyer.org/quad9.html>, et plusieurs autres. Je ne discuterai pas ici des inconvénients qu'ils présentent <https://www.bortzmeyer.org/dns-resolveurs-publics.html>, je veux juste faire quelques expériences avec un de ces résolveurs, et montrer par la même occasion quelques outils utiles pour explorer les entrailles de l'infrastructure de l'Internet.

Donc, depuis pas mal d'années, il existe un résolveur public chinois. Il a l'adresse IPv4 114.114.114.114 et il a un site Web <http://www.114dns.com/> (uniquement en chinois, il semble). Je l'ai dit, il n'est pas nouveau, mais il a fait l'objet d'un renouveau d'intérêt fin 2019 car il serait le résolveur DNS par défaut dans certains objets connectés fabriqués en Chine. Testons d'abord s'il existe et répond. Nous allons utiliser divers outils en ligne de commande qu'on trouve sur Unix. Commençons par le client DNS dig :

```
% dig +dnssec @114.114.114.114 cyberstructure.fr AAAA

; <<>> DiG 9.11.5-P4-5.1-Debian <<>> @114.114.114.114 cyberstructure.fr AAAA
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 33037
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 0

;; QUESTION SECTION:
;cyberstructure.fr. IN AAAA

;; ANSWER SECTION:
```

```
cyberstructure.fr. 86399 IN AAAA 2001:4b98:dc0:41:216:3eff:fe27:3d3f

;; Query time: 3113 msec
;; SERVER: 114.114.114.114#53(114.114.114.114)
;; WHEN: Sun Dec 29 16:31:54 CET 2019
;; MSG SIZE  rcvd: 63
```

On note d'abord qu'il est lent (depuis la France) et qu'on a souvent des délais de garde dépassés. Mais il marche. Notez que l'utilisation de ping montre que c'est bien le résolveur qui est lent, pas le réseau : les paquets envoyés par ping ont une réponse bien plus rapide :

```
% ping -c 10 114.114.114.114
...
10 packets transmitted, 10 received, 0% packet loss, time 24ms
rtt min/avg/max/mdev = 95.508/96.107/98.871/1.031 ms
```

Le RTT est bien inférieur à ce qu'on obtient quand on va en Chine, indiquant que ce résolveur a des instances en d'autres endroits, probablemement grâce à l'anycast. C'est confirmé par un ping depuis une machine aux États-Unis, qui montre des RTT qui sont encore moins compatibles avec la durée d'un aller-retour en Chine :

```
% ping -c 10 114.114.114.114
...
--- 114.114.114.114 ping statistics ---
10 packets transmitted, 10 received, +1 duplicates, 0% packet loss, time 20ms
rtt min/avg/max/mdev = 19.926/20.054/20.594/0.226 ms
```

Qui gère ce résolveur ? whois nous dit que c'est bien en Chine :

```
% whois 114.114.114.114
% [whois.apnic.net]
...
inetnum:        114.114.0.0 - 114.114.255.255
netname:        XFInfo
descr:          NanJing XinFeng Information Technologies, Inc.
descr:          Room 207, Building 53, XiongMao Group, No.168 LongPanZhong Road
descr:          Xuanwu District, Nanjing, Jiangsu, China
country:        CN
...
source:         APNIC
```

Ce service n'est apparemment pas un résolveur menteur <https://www.bortzmeyer.org/censure-franca
html>. Les noms censurés en Chine, comme `facebook.com` fonctionnent. Si on le souhaite, la documentation semble indiquer (je n'ai pas testé) que d'autres adresses abritent des résolveurs menteurs, par exemple `114.114.114.119` filtre le logiciel malveillant et `114.114.114.110` filtre le porno.

Quelles sont les caractéristiques techniques du service ? D'abord, notons qu'il ne valide **pas** avec DNSSEC, ce qui est dommage. On le voit car il n'y a pas le "*flag*" AD ("*Authentic Data*") dans la réponse

au dig plus haut (alors que le domaine visé est signé avec DNSSEC). Une autre façon de voir qu'il n'a pas DNSSEC est de demander des noms **délibérement** invalides comme `servfail.nl.114.114.114.114` donne une réponse alors qu'il ne devrait pas.

Pendant qu'on est sur DNSSEC, notons une bizarrerie du résolveur : il renvoie parfois une section EDNS (RFC 6891 [1]) et parfois pas. Dans l'exemple dig ci-dessus, il n'y en avait pas, mais parfois si :

```
% dig +dnssec @114.114.114.114 afnic.fr

; <<>> DiG 9.11.5-P4-5.1-Debian <<>> +dnssec @114.114.114.114 afnic.fr
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 59581
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
; COOKIE: c46f830f4a12bfab (echoed)
;; QUESTION SECTION:
;afnic.fr. IN A

;; ANSWER SECTION:
afnic.fr. 579 IN A 192.134.5.25

;; Query time: 20 msec
;; SERVER: 114.114.114.114#53(114.114.114.114)
;; WHEN: Sun Dec 29 17:00:19 CET 2019
;; MSG SIZE  rcvd: 65
```

Ce comportement est tout à fait anormal, et il se peut qu'il y ait plusieurs machines derrière un répartiteur de charge, avec des configurations différentes.

Autre chose qui manque : il n'a pas de NSID ("*Name Server IDentifier*", RFC 5001), cette indication du nom du serveur, très pratique lorsqu'on analyse un service anycasté, c'est-à-dire composé de plusieurs instances. (NSID peut également indiquer si une machine qui répond est la vraie, au cas, fréquent, où un détourneur n'ait pas eu l'idée de faire un faux NSID.)

```
% dig +dnssec +nsid @114.114.114.114 framagit.org

; <<>> DiG 9.11.5-P4-5.1-Debian <<>> +dnssec +nsid @114.114.114.114 framagit.org
; (1 server found)
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 745
;; flags: qr rd ra; QUERY: 1, ANSWER: 1, AUTHORITY: 0, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
; NSID
; COOKIE: 7f99ae3cb5e63ad1 (echoed)
```

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc6891.txt`

```
;; QUESTION SECTION:
;framagit.org. IN A

;; ANSWER SECTION:
framagit.org. 10779 IN A 144.76.206.42

;; Query time: 20 msec
;; SERVER: 114.114.114.114#53(114.114.114.114)
;; WHEN: Sun Dec 29 17:05:40 CET 2019
;; MSG SIZE  rcvd: 73
```

Le NSID est vide, ce qui est bizarre. Normalement, quand un serveur DNS ne veut pas répondre à la question NSID, il ne renvoie pas l'option EDNS dans la "*OPT pseudo-section*". Ici, ce zèbre étrange renvoie l'option, mais vide.

Le service n'a apparemment pas d'adresse IPv6, ce qui est étonnant en 2019. En tout cas, il n'y a pas d'enregistrement de type AAAA pour le nom :

```
% dig +dnssec public1.114dns.com. AAAA

; <<>> DiG 9.11.5-P4-5.1-Debian <<>> public1.114dns.com. AAAA
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 4421
;; flags: qr rd ra; QUERY: 1, ANSWER: 0, AUTHORITY: 1, ADDITIONAL: 1

;; OPT PSEUDOSECTION:
; EDNS: version: 0, flags: do; udp: 4096
; COOKIE: 0d4b4b8dd9fc8613c831aa085e08cf5e9b2e8073e1d6741d (good)
;; QUESTION SECTION:
;public1.114dns.com. IN AAAA

;; AUTHORITY SECTION:
114dns.com. 600 IN SOA ns1000.114dns.com. dnsadmin.114dns.com. (
20120510   ; serial
3600       ; refresh (1 hour)
300        ; retry (5 minutes)
604800     ; expire (1 week)
300        ; minimum (5 minutes)
)

;; Query time: 114 msec
;; SERVER: 127.0.0.1#53(127.0.0.1)
;; WHEN: Sun Dec 29 17:07:58 CET 2019
;; MSG SIZE  rcvd: 127
```

(Au fait, j'ai trouvé ce nom `public1.114dns.com` en faisant une résolution « inverse », avec `dig -x 114.114.114.114`.)

Enfin, ce résolveur n'a que le DNS traditionnel, sur UDP. Pas de TCP, pourtant normalisé dès les débuts du DNS, au même titre que UDP, et surtout pas de chiffrement pour sécuriser la communication, ce résolveur n'a pas DoT (DNS sur TLS, RFC 7858) ou DoH (DNS sur HTTPS, RFC 8484). On peut tester ces deux protocoles avec Homer <`https://framagit.org/bortzmeyer/homer`> :

```
% homer https://114.114.114.114/ cyberstructure.fr
...
pycurl.error: (7, 'Failed to connect to 114.114.114.114 port 443: Connection timed out')

% homer --dot 114.114.114.114 cyberstructure.fr
timeout
```

Tester depuis un seul point de mesure, ou même deux, est évidemment insuffisant pour un résolveur "*anycasté*", donc nous allons utiliser les sondes RIPE Atlas <https://atlas.ripe.net/> via l'outil Blaeu <https://labs.ripe.net/Members/stephane_bortzmeyer/creating-ripe-atlas-one-off-measur Ici, on demande à mille sondes RIPE Atlas, réparties sur toute la planète, d'interroger le résolveur public :

```
% blaeu-resolve -r 1000 --nameserver 114.114.114.114  --displayrtt --dnssec --nsid --displayvalidation write.as
Nameserver 114.114.114.114
[2001:4800:7812:514:500b:b07c:ff05:694d NSID: b''] : 909 occurrences Average RTT 129 ms
[2001:4800:7812:514:500b:b07c:ff05:694d] : 31 occurrences Average RTT 430 ms
[TIMEOUT] : 42 occurrences Average RTT 0 ms
Test #23727649 done at 2019-12-29T15:54:15Z
```

On voit le nombre relativement élevé de "*timeout*" (comme je l'ai noté plus haut, ce service est lent), et le fait que parfois, on n'a pas de réponse EDNS (et donc pas de NSID, même vide). Plus drôle, en répétant le test, les sondes Atlas trouvent plusieurs détournements de ce résolveur public. Une des grosses faiblesses d'un résolveur public sans authentification (ce qui est le cas de tous ceux qui n'offrent ni DoT, ni DoH) est qu'on ne peut pas les authentifier. Il est donc trivial de faire un faux serveur, en jouant avec le routage, comme le font souvent les censeurs <https://www.bortzmeyer.org/turquie-dns-frnog.html>. J'ai ainsi vu une sonde Atlas détecter un NSID « CleanBrowsing v1.6a - dns-edge-usa-east-dc-c », manifestement un outil de filtrage qui se faisait passer pour 114.114.114.114 et qui l'assumait, en affichant un NSID à lui.

Puisqu'on a parlé de routage, c'est l'occasion de tester traceroute. Un traceroute ordinaire ne donne rien, ses paquets étant manifestement jetés dans le réseau du résolveur. On va donc utiliser traceroute avec TCP vers le port 53, celui du DNS, depuis la machine états-unienne de tout à l'heure :

```
% tcptraceroute 114.114.114.114 53
traceroute to 114.114.114.114 (114.114.114.114), 30 hops max, 60 byte packets
 1  88.214.240.4 (88.214.240.4)  2.518 ms  2.613 ms  2.699 ms
 2  be5893.rcr51.ewr04.atlas.cogentco.com (38.122.116.9)  0.485 ms  1.009 ms  1.014 ms
 3  be3657.rcr21.ewr03.atlas.cogentco.com (154.24.33.169)  1.021 ms  0.986 ms  1.091 ms
 4  be2495.rcr24.jfk01.atlas.cogentco.com (154.54.80.193)  1.350 ms be2390.rcr23.jfk01.atlas.cogentco.com (154.5
 5  be2897.ccr42.jfk02.atlas.cogentco.com (154.54.84.213)  1.631 ms  1.669 ms be2896.ccr41.jfk02.atlas.cogentco.
 6  be2890.ccr22.cle04.atlas.cogentco.com (154.54.82.245)  13.427 ms  13.581 ms be2889.ccr21.cle04.atlas.cogentc
 7  be2718.ccr42.ord01.atlas.cogentco.com (154.54.7.129)  20.108 ms  21.048 ms  20.062 ms
 8  be3802.rcr21.ord07.atlas.cogentco.com (154.54.47.38)  20.992 ms 154.54.89.2 (154.54.89.2)  20.794 ms be3802.
 9  * * *
10  23.237.126.230 (23.237.126.230)  22.188 ms  22.502 ms  22.331 ms
11  23.237.136.242 (23.237.136.242)  24.471 ms  23.732 ms  24.089 ms
12  * * *
13  public1.114dns.com (114.114.114.114) <syn,ack>  19.963 ms  20.631 ms  20.658 ms
```

---

https://www.bortzmeyer.org/resolveur-dns-public-chinois.html

On y voit que la machine répond bien en TCP (bien qu'elle refuse d'assurer un service DNS en TCP, encore une bizarrerie). On note aussi que le RTT confirme celui de ping. (Ce qui ne va pas de soi : le réseau peut traiter différemment différents protocoles.)

Il y a une dernière chose très étrange avec ce service. Son adresse fait partie d'un préfixe IP annoncé en BGP, `114.114.112.0/21`. Vous pouvez le trouver auprès d'un service comme RIPE Stat <`https://stat.ripe.net`>, ou bien en utilisant curl pour interroger une API publique :

```
%  curl https://bgp.bortzmeyer.org/114.114.114.114
114.114.112.0/21 174
```

Après le préfixe (`114.114.112.0/21`), on voit le numéro d'AS, 174. Or, cet AS est l'opérateur étatsunien Cogent :

```
% whois AS174
...
# Copyright 1997-2019, American Registry for Internet Numbers, Ltd.

ASNumber:       174
ASName:         COGENT-174
ASHandle:       AS174
Ref:            https://rdap.arin.net/registry/autnum/174

OrgName:        Cogent Communications
OrgId:          COGC
Address:        2450 N Street NW
City:           Washington
StateProv:      DC
PostalCode:     20037
Country:        US
Ref:            https://rdap.arin.net/registry/entity/COGC
```

A priori, pourquoi pas, une organisation chinoise a bien le droit de faire appel à un opérateur étatsunien pour annoncer son préfixe. Mais il y a des choses étranges. D'abord, la route enregistrée dans l'IRR d'APNIC ne mentionne pas Cogent, mais l'AS 4837, China Unicom :

```
%  whois 114.114.114.114
...
route:          114.114.112.0/21
descr:          China Unicom Shandong Province network
descr:          Addresses from CNNIC
country:        CN
origin:         AS4837
mnt-by:         MAINT-CNCGROUP-RR
last-modified:  2011-04-12T07:52:02Z
source:         APNIC
```

Il n'y a pas de ROA ("*Route Origin Authorizations*", RFC 6482) pour ce préfixe :

```
% whois -h whois.bgpmon.net 114.114.114.114
% This is the BGPmon.net whois Service
...
RPKI status:         No ROA found
...
```

(On pourrait avoir le même résultat en consultant RIPE stat <https://stat.ripe.net/widget/prefix-routing-consistency#w.resource=114.114.114.114>.) La seule information quant à la validité de l'annonce BGP est donc l'objet `route`, qui exclut Cogent. Notez qu'en fouillant un peu plus, on trouve des annonces du préfixe `114.114.112.0/21` par d'autres AS, 4837 (ce qui est conforme à l'objet `route` ci-dessous), et 4134 (ChinaNet). Mais la plupart des routeurs BGP de la planète ne voient que l'annonce de Cogent.

Que faut-il déduire de ce dernier cafouillage ? La situation est certes anormale (l'IRR n'annonce qu'un seul AS d'origine possible, mais c'est un autre qu'on voit presque partout), mais est-ce le résultat d'une attaque délibérée, comme les détournements BGP dont on parle souvent (RFC 7464), en général en les attribuant aux…Chinois ? Probablement pas : la situation générale de la sécurité du routage est médiocre, les IRR ne sont pas maintenus à jour, l'opérateur Cogent n'est pas connu pour sa rigueur (il ne devrait pas annoncer des préfixes sans qu'ils soient marqués comme tel dans un IRR) et il s'agit sans doute plus de négligence que de malveillance.

Et pour finir la partie sur le routage, une copie d'écran de RIPE stat montrant l'incohérence des annonces :

Revenons un peu à ping. Manuel Ponce a fort justement observé qu'il y avait un problème intéressant dans les réponses à ping :

```
% ping -c 10 114.114.114.114

PING 114.114.114.114 (114.114.114.114) 56(84) bytes of data.
64 bytes from 114.114.114.114: icmp_seq=1 ttl=63 time=94.0 ms
64 bytes from 114.114.114.114: icmp_seq=2 ttl=56 time=93.8 ms
64 bytes from 114.114.114.114: icmp_seq=3 ttl=81 time=93.9 ms
64 bytes from 114.114.114.114: icmp_seq=4 ttl=84 time=94.1 ms
64 bytes from 114.114.114.114: icmp_seq=5 ttl=62 time=94.2 ms
64 bytes from 114.114.114.114: icmp_seq=6 ttl=66 time=94.0 ms
64 bytes from 114.114.114.114: icmp_seq=7 ttl=80 time=93.9 ms
64 bytes from 114.114.114.114: icmp_seq=8 ttl=69 time=93.9 ms
64 bytes from 114.114.114.114: icmp_seq=9 ttl=72 time=94.2 ms
64 bytes from 114.114.114.114: icmp_seq=10 ttl=59 time=94.1 ms

--- 114.114.114.114 ping statistics ---
10 packets transmitted, 10 received, 0% packet loss, time 9011ms
rtt min/avg/max/mdev = 93.801/94.045/94.285/0.342 ms
```

Prenez cinq minutes pour regarder et voyez si vous trouvez le problème…

C'est bon, vous avez trouvé ? Moi, je n'avais pas vu : le TTL change à chaque paquet. Normalement, il devrait être à peu près constant, sauf si le routage se modifie subitement pendant le test (ce qui est rare). Ici, c'est un mystère de plus de ce curieux service. Personne n'a trouvé d'explication certaine de ce phénomène mais Baptiste Jonglez suggère cette commande Netfilter si vous voulez faire pareil sur vos serveurs (**Non testé ! Ne l'essayez pas en production comme ça !**) :

```
for ttl in {0..31}
    do
      iptables -t mangle -A OUTPUT -p icmp --icmp-type echo-reply -m statistic --mode nth --every 32 --packet $t
    done
```

Cela donne des TTL dans l'ordre. Pour les avoir aléatoires, Paul Rolland suggère de remplacer `--mode nth --every 32 --packet $ttl` par `--mode random --probability 0.03125`.

---

https://www.bortzmeyer.org/resolveur-dns-public-chinois.html