Noms de domaine, DNS et vie privée

Stéphane Bortzmeyer stephane+root66@bortzmeyer.org

Root66, 13 janvier 2018

Plan

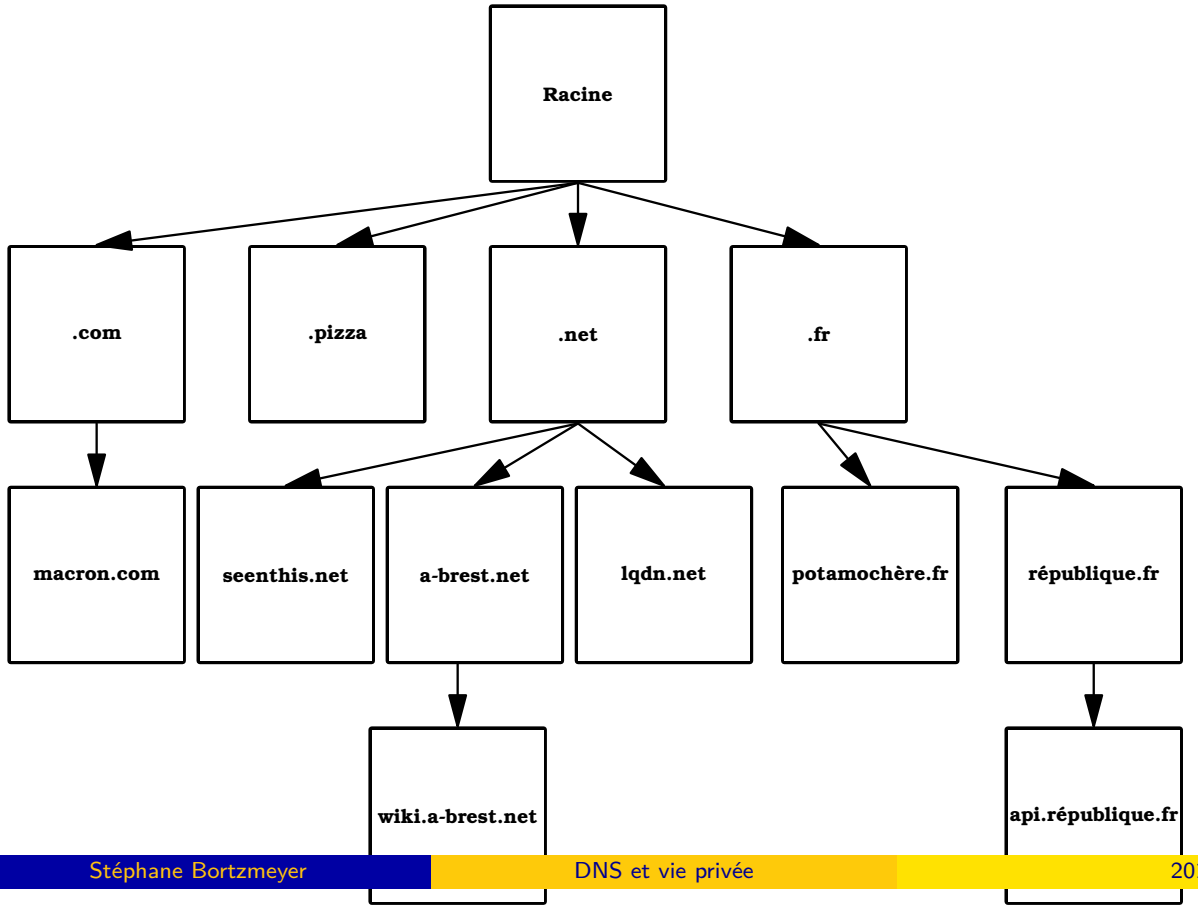
Généralités

- Des noms uniques et mémorisables,
- Un vecteur d'identité,
- Un nommage arborescent : racine, puis TLD puis domaine de deuxième niveau, de troisième niveau et ainsi de suite,
- Le nombre de composants dans un nom est quelconque (1, 2, 3, 4...)

Les noms

- Exemples de noms de domaines : `www.root66.net`, `foire-aux-greniers.fontenay-le-fleury.org`, `www.phy.cam.ac.uk`, `fr.wikipedia.org`, `www.potamochère.fr`, `gmail.com`, `www.st-cyr.terre.defense.gouv.fr`, `re`, `_sipfederationtls._tcp.en-marche.fr`, `mamot.fr`...
- `nca.x.gsi.gov.uk` a cinq composants. Le nom le plus général, le **TLD** (*Top-Level Domain*, ici `uk`) est à la fin.

L'arbre du DNS



Délégation

Des noms peuvent être **délégués** et on change alors d'organisme responsable. Par exemple `uk.com` est délégué depuis `com` et délègue à son tour.

Rien dans le nom n'indique où est la frontière de délégation : il faut utiliser le DNS.

Le DNS

DNS = *Domain Name System*

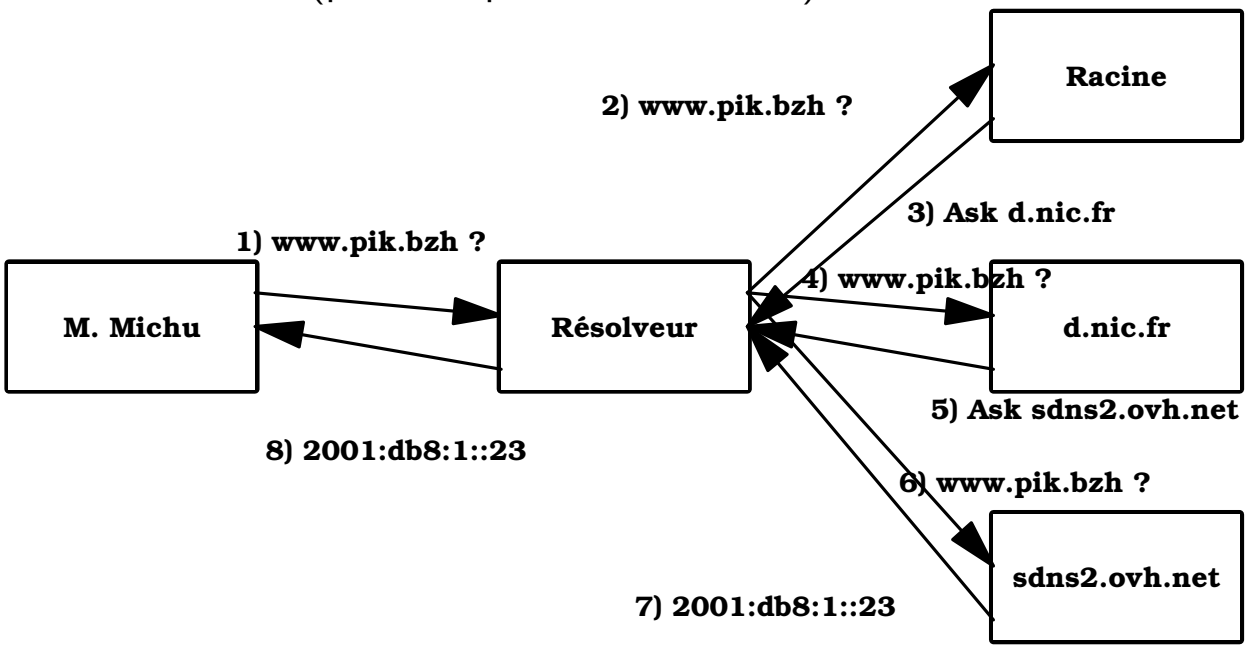
- Les adresses IP ne sont pas stables (et ont d'autres limites),
- On utilise donc plutôt des **noms** qui, eux, sont stables,
- Le DNS est une base de données et un protocole qui associent à ces noms des informations (comme les adresses IP),
- C'est une technologie d'**infrastructure** comme l'eau ou l'électricité : tant qu'elle marche, personne ne la voit. Le DNS reste donc peu connu et peu discuté. Ses failles « vie privée » ont donc été peu étudiées.

Vocabulaire important

- **Résolveur** (ou serveur récursif) : serveur DNS qui ne connait rien mais pose des questions aux serveurs faisant autorité et mémorise les réponses. Chez le FAI, ou sur le réseau local ou chez Google.
- **Serveur faisant autorité** : serveur DNS qui connait le contenu d'un domaine. Exemple : les serveurs de l'AFNIC qui connaissent ce qu'il y a dans `.fr` et peuvent répondre. Ou les serveurs de `gouvernement.fr` chez Gandi.

Résolution de noms, ou le protocole DNS en action

Résolution : demander aux serveurs DNS les informations associées à un nom de domaine (par exemple les adresses IP)



Avec le client DNS dig

```
% dig AAAA www.bortzmeyer.org
; <<>> DiG 9.10.3-P4-Debian <<>> AAAA www.bortzmeyer.org
;; global options: +cmd
;; Got answer:
;; ->>HEADER<<- opcode: QUERY, status: NOERROR, id: 39813
;; flags: qr rd ra ad; QUERY: 1, ANSWER: 3, AUTHORITY: 7, ADDITIONAL: 21
...
;; ANSWER SECTION:
www.bortzmeyer.org. 64232 IN AAAA 2605:4500:2:245b::42
...
;; Query time: 0 msec
;; SERVER: 127.0.0.1#53(127.0.0.1)
;; WHEN: Thu Jan 11 21:20:25 CET 2018
;; MSG SIZE  rcvd: 2310
```

Types de données

- AAAA : adresses IP
- NS : serveurs de noms faisant autorité
- A : adresses de l'ancien protocole IP (IPv4)
- SRV : serveurs du domaine pour une application donnée (par exemple XMPP)

Les informations dans une requête DNS

```
20:22:31.203363 IP 192.0.2.1.43850 > 195.154.55.57.53: 29630 [1au] \
        A? www.bresil-guide.com. (49)
```

1. Adresse IP source
2. QNAME (*Query Name*, le nom demandé),

Des clients BitTorrent demandent `_bittorrent-tracker._tcp.domain.example...`

Le numérique laisse des traces

- Tout peut être enregistré,
- Et, en pratique, l'est souvent,
- Et peut être traité automatiquement.

Révélations Snowden

- Juin 2013 : la NSA espionne tout le monde et de manière illégale,
- C'était déjà bien connu des spécialistes. Mais, après Snowden, plus moyen de faire semblant, de dire « mais arrête d'être parano ».

Données personnelles

- Contrairement à ce qu'on lit souvent, une « donnée personnelle » n'est pas uniquement une donnée où apparait un nom,
- Un numéro « de Sécu », un numéro de téléphone, sont des données personnelles,
- L'adresse IP aussi.

RGPD

- N° 2016/679, Règlement général sur la protection des données,
- Rentre en application le 25 mai 2018.
- Parmi les grands principes : ne pas stocker des données juste parce qu'on peut, tout stockage doit avoir une raison.
- Une grande partie de ce qu'on présente comme « obligations RGPD » était déjà dans la loi Informatique & Libertés en janvier 1978 !

Qui peut lire vos questions ?

- Le gérant du résolveur,
- Le gérant du serveur faisant autorité,
- Tout ceux qui capturent le trafic entre vous et ces deux serveurs.

Le résolveur

1. Le résolveur peut enregistrer vos requêtes (vous avez consulté `www.djihad.sa`, `pornhub.com` et `alcooliques-anonymes.fr`), quel que soit le TLD,
2. Il voit l'adresse IP de votre machine personnelle,
3. Il n'est pas gêné par les caches,
4. Le résolveur en sait beaucoup, mais est une machine que vous avez choisie (même si vous ne le savez pas).

Comment je trouve mon résolveur DNS ?

1. Par défaut, indiqué à votre machine par le serveur DHCP,
2. Ce serveur DHCP peut être contrôlé par votre FAI / votre service informatique, ou bien cela peut être un engin à vous (100 % libre avec OpenWRT ; cf. Turris Omnia),
3. On peut souvent surmonter cette décision et indiquer le résolveur qu'on veut.

Le serveur faisant autorité

1. Le serveur faisant autorité peut enregistrer vos requêtes,
2. Il voit l'adresse IP du résolveur (mais attention à ECS, RFC 7871, qui lui permet de voir celle des clients),
3. Les caches font qu'il ne voit pas toutes les requêtes,
4. Ils ne sont pas forcément sur le chemin entre vous et votre pair (vous vous connectez à `example.com` en France, vos requêtes DNS vont quand même aux USA),
5. Ils sont nombreux et vous ne les avez pas choisis.

Les écoutants

1. Ils écoutent (*sniffing*) le trafic qui passe,
2. Un écoutant situé avant le résolveur (en amont) voit tout ce que voit le résolveur.
3. L'écoutant a davantage de travail que les serveurs pour reconstituer les requêtes (fragmentation, TCP...). Pas un problème pour la NSA mais peut gêner l'amateur.
4. Il ne suffit donc pas d'avoir confiance dans le serveur !

Et en pratique ?

- L'espionnage DNS est réel (programme MoreCowBell de la NSA),
- Il y a aussi beaucoup de captures et d'analyses du trafic faites pour le bien (?) : *passive DNS*, analyse comportementale de logiciels malveillants, genre zombies...

Politique « vie privée » de serveurs DNS

- Google Public DNS « *In the permanent logs, we don't keep personally identifiable information or IP information. We do keep some location information (at the city/metro level)* »
- Quad9 « *In the permanent logs, we do not keep personally identifiable information (PII) or IP information. We do keep some location information (at the city/metro level)* »
- Je ne connais aucun gérant de serveur faisant autorité qui ait une politique « vie privée ».

Projet « DNS privacy »

- Démarrage en novembre 2013,
- Groupes de travail dnsop (existant) et dprive (créé pour l'occasion),
- RFC 7626, documentation du problème, en août 2015,
- RFC 7816, minimisation des données, en mars 2016,
- RFC 7858, chiffrement du DNS, en mai 2016,
- RFC sur l'authentification du résolveur, d'un moment à l'autre,
- En cours, l'authentification du serveur faisant autorité.

Deux classes de solutions

- Chiffrer,
- Minimiser les données envoyées.

Chiffrer

- Protège contre les écoutants **mais pas contre les serveurs**,
- « Le chiffrement vous permet d'avoir une liaison sécurisée avec un ennemi »
- DNS-sur-TLS, RFC 7858, sur le port 853,
- Ou bien dnscrypt et dnscurve (non standards).

Configuration Unbound, chiffrement vers l'aval

```
server:
  ...
  ssl-upstream: yes # Should be TLS, not SSL!

forward-zone:
  name: "."
  # https://dns-resolver.yeti.eu.org/
  forward-addr: 2001:4b98:dc2:43:216:3eff:fea9:41a@853
  forward-first: no
```


Configuration Unbound, chiffrement vers l'amont

```
server:
  interface: 2001:db8:1::dead:beef@853
  ssl-service-key: "/etc/unbound/privatekeyfile.key"
  ssl-service-pem: "/etc/unbound/publiccertfile.pem"
  ssl-port: 853
```

Quad9 utilise Unbound pour fournir ce service DNS-sur-TLS.

C'est bien du vrai TLS

Résolveur public de LDN :

```
% openssl s_client -connect \[2001:913::8\]:853  -showcerts
...
Server certificate
subject=/C=FR/ST=Some-State/O=LDN/CN=80.67.188.188
issuer=/C=FR/ST=Some-State/O=LDN/CN=80.67.188.188
```


DNS sur TLS, vu par Wireshark

```
9.542061163 2001:db8::1 → 2001:4b98:dc2:43:216:3eff:fea9:41a SSL 262
        Client Hello
9.559873308 2001:4b98:dc2:43:216:3eff:fea9:41a → 2001:db8::1 TLSv1.2 2576
        Server Hello, Certificate, Server Key Exchange, Server Hello Done
9.561219238 2001:db8::1 → 2001:4b98:dc2:43:216:3eff:fea9:41a TLSv1.2 171
        Client Key Exchange, Change Cipher Spec, Encrypted Handshake Message
...
9.579324930 2001:db8::1 → 2001:4b98:dc2:43:216:3eff:fea9:41a TLSv1.2 109
        Application Data
```

Autres mises en œuvre

- Dans Android depuis juillet 2017 (pas encore dans les versions publiées),
- Dans l'excellente bibliothèque C getdns,
- Le résolveur (incomplet) Stubby, bâti sur getdns,
- Dans l'excellente bibliothèque Go go-dns,
- En 2018, vient d'arriver dans le résolveur Knot.

Configuration Stubby

```
listen_addresses:
  - 0::1@8053

dns_transport_list:
  - GETDNS_TRANSPORT_TLS

upstream_recursive_servers:
# Quad9
# Other resolvers, see
# https://dnsprivacy.org/wiki/display/DP/DNS+Privacy+Test+Servers
   - address_data: 9.9.9.9
     tls_auth_name: "dns.quad9.net"
   - address_data: 2620:fe::fe
     tls_auth_name: "dns.quad9.net"
```

Minimiser les données

- *QNAME minimisation*, RFC 7816,
- N'envoyer au serveur faisant autorité que ce qu'il a besoin de savoir (*Need To Know*),
- Mis en œuvre uniquement dans le résolveur.

Effet de la QNAME minimisation, vu par tcpdump

```
% ping foo.toto.so
```

Sans :

```
> racine: 27934% [1au] A? foo.toto.so. (40)
> TLD: 63293% [1au] A? foo.toto.so. (40)
> SLD: 63293% [1au] A? foo.toto.so. (40)
```

Avec :

```
> racine: 58136% [1au] A? so. (31)
> TLD: 45447% [1au] A? toto.so. (36)
> SLD: 46928% [1au] A? foo.toto.so. (40)
```

Déploiement QNAME minimisation

```
% atlas-resolve -r 1000 -t TXT qnamemintest.internet.nl
[TIMEOUT(S)] : 7 occurrences
[ERROR: SERVFAIL] : 5 occurrences
["hooray - qname minimisation is enabled"] : 23 occurrences
["no - qname minimisation is not enabled"] : 960 occurrences
Test #10854359 done at 2018-01-11T20:48:07Z
```


Activer la QNAME minimisation

- Par défaut dans Knot,
- Dans Unbound, `qname-minimisation: yes`.

Remplissage

- TLS ne dissimule pas la taille des données,
- Le DNS étant public, on peut connaitre cette taille et en déduire la requête,
- Remplissage (*padding*, RFC 7830).

Choix du résolveur

- Résolveur du FAI / du service informatique : écoute difficile en amont (mais, en aval, peu font de la *QNAME minimisation*).
- Résolveur public : que font-ils des données ? Rarement chiffré (sauf Cisco OpenDNS, LDN et Quad9) donc vulnérable à l'écoute en amont. En aval, on est « protégé » par le nombre de requêtes.
- Son propre résolveur : aucune capture des données (si logiciel libre). Bonne protection des requêtes amont mais très mauvais pour les requêtes aval (pensez à la *QNAME minimisation*).