

## M3102 : Services Réseaux

### 4. DNS (Domain Name System)

Franck Butelle Micaela Mayero

IUT Villetaneuse,  
Département Informatique, S3

09/2015

### Pourquoi DNS ?

- Quelle est l'adresse IP de `www.google.fr` ?  
209.85.129.99, 74.125.230.81, ...
- ↪ Outils pour faire de la résolution
  - Adresse symbolique → Adresse IP
  - Adresse IP → Adresse symbolique
  - Hiérarchisation (y compris gestion)
  - Décentralisation
  - Unicité mondiale
  - Services supplémentaires (serveur mail etc)
  - Redondance (serveur secondaire)
  - Efficacité (serveur cache)

### Plan

- 1 Généralités
- 2 Éléments fondamentaux
- 3 Principes d'administration
- 4 Algorithme de résolution d'adresse
- 5 Configuration d'un serveur
- 6 Cache, esclave, etc
- 7 Format des PDU

### Sans DNS (résolution statique)

Résolution : fichier texte appelé *hosts* :

- local à chaque ordinateur
- sous UNIX, dans le répertoire `/etc`
- sous WindowsXP, dans `\SYSTEMROOT\SYSTEM32\DRIVERS\ETC`
- chaque ligne correspond à une adresse IP à laquelle peuvent être associés un ou plusieurs noms de machine

Problème de maintenance :

- difficile d'ajouter de nouveaux noms
- fichier doit être dupliqué entre tous les ordinateurs du réseau

## Résolution statique

- Fichier `/etc/hosts`
- Format texte :  

```
#Syntaxe:
#
#IP-Address Fully-Qualified-Domainname aliases
#
127.0.0.1      localhost
194.254.163.24 lipn.univ-paris13.fr lipn
```

## Les composants DNS

- Serveurs de noms
- Résolveurs (les clients)
- Espace des noms de domaines et les informations associées :  
des *Resource Records* (ou RR)

## Principe de la résolution avec DNS

- par l'intermédiaire d'un **serveur** :
  - le client demande au serveur l'adresse IP de `www.google.fr`
- **Résolution inverse** : quel est le nom de machine associé à telle adresse IP ?

### Propriétés :

- Plusieurs adresses symboliques peuvent être associées à une même adresse IP (alias)
- Une adresse symbolique peut être associée à plusieurs adresses IP (tourniquet)
- Une adresse IP est "inversement résolue" en une seule adresse symbolique (canonique)

## Syntaxe d'un enregistrement (RR)

[*nom*] [TTL] [Qclass] **Type\_RR** DATA

### Type\_RR :

- **A** : associe *nom* à **adresse IPv4**
- **AAAA** : associe *nom* à **adresse IPv6**
- **CNAME** : définit un *alias* à un *nom*
- **TXT** : pour définir un texte libre (description du domaine), *nom*=" "
- **NS** : pour définir le *nom* du NameServer responsable, *nom*=" " sauf si délégation d'autorité
- **SOA** : Start Of Authority : marque le début d'une zone de responsabilité/autorité
- **MX** : déclare un serveur/relais de mail pour le domaine avec un certain niveau de préférence, *nom*=" "
- **PTR** : pour la résolution inverse
- ...

### Champs optionnels :

TTL : *Time To Live* (durée de vie de l'info en secondes).

Qclass : IN pour Internet (seule valeur possible de nos jours).

## L'espace de nommage

- Arborescence ; découpée en niveaux de domaines :
  - Racine ( . ) ( {A-M}.ROOT-SERVERS.NET)
  - TLD : Top Level Domain (génériques : com, edu, ... et nationaux : fr, de, ...)
  - SLD : Secondary Level Domain...
- A chaque nœud (ou feuille)
  - Un nom ou étiquette (max. 63 octets)
  - Si pas feuille, peut comporter un ensemble de RR (Resource Record) : une *zone*
- Le nom de domaine d'un nœud
  - Suite des étiquettes en remontant jusqu'à la racine.

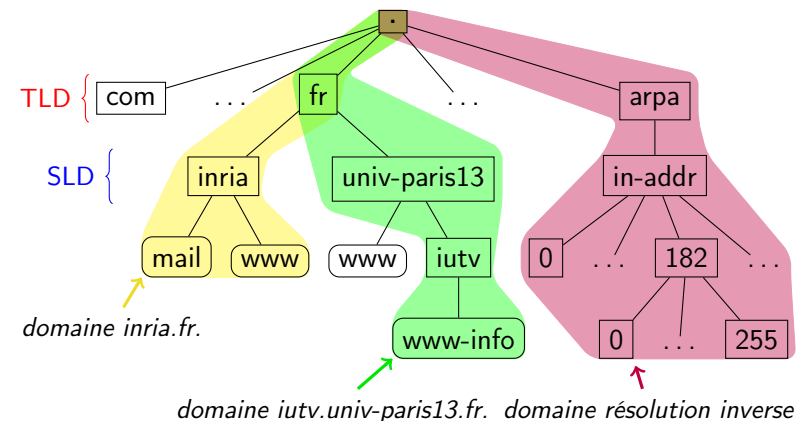
## Exemple d'adressage hiérarchique

`www.iutv.univ-paris13.fr.`

`www ⊂ iutv ⊂ univ-paris13 ⊂ fr ⊂ .`

- Une adresse symbolique :  
Fully Qualified Domain Name : max 255 octets
- Recommandation initiale : uniquement [A-Z, 0-9,-] et doit commencer par une lettre.
  - Mais il est désormais possible d'utiliser Unicode/UTF8...
- Insensible à la casse (majuscules = minuscules)
- Codage d'un FQDN : un octet pour la longueur de l'étiquette puis le code ASCII de l'étiquette (00 pour le . final)

## Extrait d'une arborescence DNS



Tout nœud interne peut contenir une zone.

## Domaine, zone

- Un domaine est identifié par un Nom de domaine : Sous-arborescence qui a pour origine ce nœud/nom.
- Un domaine inclus dans un autre est un sous-domaine
- Le gérant d'un domaine doit garantir l'unicité d'un nom dans ce domaine (pas deux univ-paris13 dans fr !)
- Une zone (de responsabilité) : partie descriptive d'un nœud, incluse dans un domaine.
- Un serveur peut faire autorité sur plusieurs zones.

## Autorité

- Un serveur fait autorité (authority) pour une zone dans un domaine.
- Il connaît et fait autorité pour toutes les correspondances de sa zone.
- Il connaît (déclare) les responsables des sous-domaines de son domaine.
- Il ne peut faire autorité que s'il se déclare comme tel et que son supérieur le déclare comme tel.

## Résolution inverse

- Adresse IP → Adresse symbolique
- Pseudo-domaine spécial : **in-addr.arpa.**, des pointeurs (**PTR**), des adresses *à l'envers*!

### Exemple

l'adresse 192.16.1.1 (classe C) est en fait

- Dans le domaine 1.16.192.in-addr.arpa.
- Lui même inclus dans 16.192.in-addr.arpa....
- La résolution inverse donne une seule adresse symbolique :
  - L'adresse canonique.
  - pas de **CNAME** dans la résolution inverse

## Administration hiérarchique

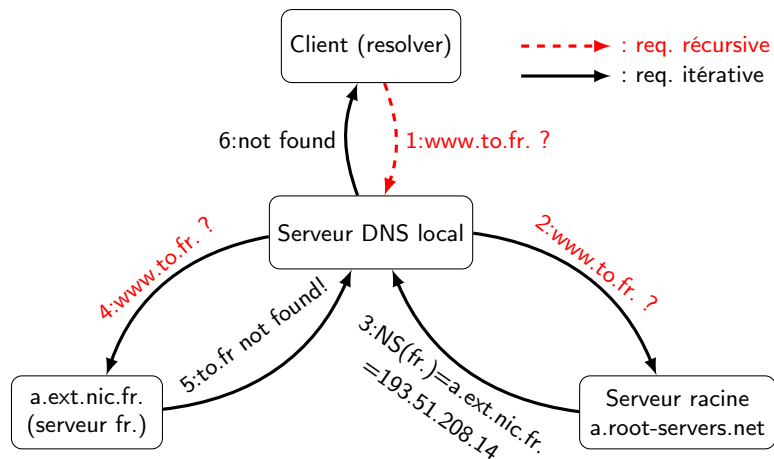
- Administration hiérarchisée :  
Coordination mondiale : le Network Information Center
- Décentralisée :
  - Le NIC a donné délégation à RIPE-NCC pour l'europe
  - RIPE-NCC à délégué à l'AFNIC la gestion de la France
  - AFNIC : Agence Française pour le Nommage Internet en Coopération ([www.afnic.fr](http://www.afnic.fr))
  - Le gérant de Y.Z est responsable de la désignation du serveur pour X.Y.Z et de la délégation éventuelle pour les sous-domaines

## Le service de nom : cas d'Unix

Plusieurs mécanismes => sélection à faire par `/etc/nsswitch.conf`  
ligne **hosts** : organise l'ordre des recherches :

- files : consulter `/etc/hosts`
- dns : consulter le serveur de noms (dont l'adresse est dans `/etc/resolv.conf`)
- nis, nisplus, db, etc.

## Exemple de résolution



## Configuration d'un serveur Bind 9

`/etc/named.conf` ou `/etc/bind/named.conf`

```
options {
    directory "/var/named"; // rép. pour les fichiers suivants
};
zone 127.in-addr.arpa. {
    type master; // maître sur la résolution inverse de 127...
    file "127.zone"; // Attention ceci n'est qu'un nom de fichier.
};
zone p13.fr. {
    type master; // maître sur la résolution de p13.fr...
    file "p13.fr.zone";
};
zone 10.150.in-addr.arpa. {
    type master; // maître sur la résolution inverse de 150.10...
    file "150.10.zone";
};
```

## Algorithme de résolution

- Le client contacte son serveur DNS local (**53/UDP**) : NS1
  - Demande aaa.bbb.ccc.fr (sinon ajout de suffixe prédef.)
  - Cette demande est en général "récursive"
- Si le serveur local fait autorité (ou en cache) : réponse immédiate.
- Sinon recherche du plus grand suffixe connu (bbb.ccc.fr ou ccc.fr ou fr ou .) avec **RR** de type **NS** : soit NS2
- Si en mode **itératif** : demande à NS2 une résolution (NS2 réponds avec un nouveau NS à contacter etc.)  
Au final NS1 peut apporter une réponse au client
- Si en mode **récursif** : demande à NS2 toute la résolution !

## Exemple de fichier de zone

⚠ Syntaxe spécifique normalisée

```
$ORIGIN p13.fr. ← important!
$TTL 3D
@ IN SOA (dns.p13.fr. hostmaster.p13.fr.
    2015092101 24H 6H 3W 1D )
NS dns.p13.fr.
TXT "Le domaine de P13"
MX 10 mail.p13.fr.
MX 20 mail2.p13.fr.
ns CNAME dns.p13.fr.
dns A 150.10.0.1
mail A 150.10.0.2
mail2 A 150.10.0.3
```

Annotations: **Version** (2015092101), **Retry** (24H), **negTTL** (6H), **Refresh** (3W), **expire** (1D), **optionnel, vu \$ORIGIN** (dns.p13.fr.)

## SOA : Start Of Authority

- @ : remplace le nom domaine défini par \$ORIGIN
- IN : classe Internet
- Nom du serveur DNS
- Adresse mail admin mais avec @ remplacé par un point.
- n° version (serial number) : à changer si RRs modifiées
  - Sert à savoir s'il faut mettre à jour caches, secondaires, etc.
  - Format très recommandé : YYYYMMDDxx
- Refresh : délai d'interrogation du numéro de version
- Retry : délai à attendre avant de refaire un Refresh raté
- Expire : si aucun Refresh n'a pu avoir lieu, durée de vie ultime des données. Recommandé : Retry « Refresh « Expire
- negTTL : Durée de vie des réponses négatives ou valeur par défaut !

## Types de serveurs

- Serveur relais (*forwarder* ou *cache*)  
se contente de relayer toute demande vers un autre :
  - Mais garde en cache mémoire
  - Permet de limiter la charge du primaire.
- Serveur secondaire (ou *esclave*) :
  - Prendre le relais en cas de panne du primaire
  - Répartition de charge
  - On peut être primaire pour une zone et secondaire pour une autre
  - Mis à jour par 53/tcp par PUSH ou PULL :  
"Transfert de zone"
- Délégation d'un sous-domaine à un serveur :  
Par le Resource Record NS.

## Exemple de fichier de zone inverse

```
$ORIGIN 10.150.in-addr.arpa. ;domaine par défaut
$TTL 3D                      ;durée vie d'une info par défaut
@ IN SOA (dns.p13.fr. hostmaster.p13.fr.
          2015092101 24H 6H 3W 1D)
      NS      dns.p13.fr.
      TXT     "domaine de P13."
1.0 PTR      dns.p13.fr.
2.0 PTR      mail.p13.fr.
3.0 PTR      mail2.p13.fr. ← obligatoire, vu $ORIGIN
```

Donc p.ex. 2.0.10.150.in-addr.arpa. pointe sur l'adresse canonique mail.p13.fr. qui est donc à l'adresse IPv4 150.10.0.2...

## Configuration de Bind 9 : serveur « cache » ou « relais »

*Pas de stockage sur disque*

/etc/named.conf ou /etc/bind/named.conf

```
options {
    directory "/var/named";
    forwarders {150.10.0.1; }; // faire suivre à dns1 si la résolution échoue
    forward only; // ne faire QUE de la mémorisation temporaire
};
```

## Config. de Bind 9 : serveur « secondaire » ou « esclave »

*Stockage sur disque*`/etc/named.conf` ou `/etc/bind/named.conf`

```
options {
    directory "/var/named/slave";
    forwarders {150.10.0.1;}; // faire suivre à dns1 si résolution échoue
};
zone p13.fr. {
    type slave; // secondaire pour la zone p13.fr.
    file "p13.fr.zone"; // sauve la zone dans ce fichier
    masters {150.10.0.1;}; // dns1 a autorité sur cette zone
};
zone 10.150.in-addr.arpa. {
    type slave;
    file "150.10.zone";
    masters {150.10.0.1;};
};
```

## Délégation d'autorité (fichier de zone)

```
$ORIGIN p13.fr. ;domaine par défaut
$TTL 3D ;TTL de 3j par défaut
@ IN SOA (dns.p13.fr. hostmaster.p13.fr.
          2015092101 24H 6H 3W 1D)
      NS dns.p13.fr.
      TXT "domaine de P13"
dns A 150.10.0.1
iutv NS dns.iutv ;délégation du domaine iutv.p13.fr. à dns.iutv
dns.iutv A 150.10.2.1 ;il faut donner une IP à dns.iutv
```

## Délégation d'autorité (zone inverse)

```
$ORIGIN 10.150.in-addr.arpa.
$TTL 3D
@ IN SOA (dns.p13.fr. hostmaster.p13.fr.
          2015092101 24H 6H 3W 1D)
      NS dns.p13.fr.
      TXT "domaine de P13."
1.0 PTR dns.p13.fr.
2.0 PTR mail.p13.fr.
2 NS dns.iutv.p13.fr.
1.2 PTR dns.iutv.p13.fr.
```

## Format des échanges

⚠ sur TCP ajout longueur du message DNS sur 2 octets avant.

| Trans ID | Flags | Nb req (#S1) | Nb réponses (#S2) | Nb (#S3) | Nb (#S4) | S1   | S2   | S3   | S4   |
|----------|-------|--------------|-------------------|----------|----------|------|------|------|------|
| 2 o.     | 2 o.  | 2 o.         | 2 o.              | 2 o.     | 2 o.     | var. | var. | var. | var. |

**Format de S1 :**

|              |             |              |
|--------------|-------------|--------------|
| Qname (var.) | Qtype (2o.) | Qclass (2o.) |
|--------------|-------------|--------------|

**Qname** : FQDN (codage voir diapo 11)**Qtype** : 0001=A, 0002=NS, 0005=CNAME, 0006=SoA, 000C=PTR, 000F=MX, 0010=TXT, 001C=AAAA, ...;**Qclass** : 0001=IN (INternet, seule valeur possible actuellement)**S1** : Questions : information (RR) demandés avec leur type**S2** : Réponses : RRs répondant à S1**S3** : RRs donnant les NS autorité correspondants aux réponses dans S2**S4** : RRs donnant des infos supplémentaires

## Format de S2, S3 et S4

| Name           | Qtype                                                                                     | Qclass | TTL  | Long.Data | Data |
|----------------|-------------------------------------------------------------------------------------------|--------|------|-----------|------|
| FQDN compressé | 2 o.                                                                                      | 2 o.   | 4 o. | 2 o.      |      |
| Qtype          | Data                                                                                      |        |      |           |      |
| A              | adresse IPv4 sur 4 o.                                                                     |        |      |           |      |
| NS             | Name                                                                                      |        |      |           |      |
| MX             | Preference sur 2o, Name                                                                   |        |      |           |      |
| CNAME          | Name                                                                                      |        |      |           |      |
| SoA            | Name, MailAdmin, N° série sur 4o., Refresh sur 4o., Retry sur 4o., Expire 4o., negTTL 4o. |        |      |           |      |
| PTR            | Name                                                                                      |        |      |           |      |
| TXT            | Ascii                                                                                     |        |      |           |      |
| AAAA           | adresse IPv6 sur 16 o.                                                                    |        |      |           |      |

## Compression des messages

Idée : éliminer les répétitions de noms de domaine dans une réponse.

- le nom de domaine entier est représenté une première fois (QNAME=FQDN) : 03www04cnam02fr00
- Puis on fait référence à un bout de domaine déjà cité dans le message par un pointeur (offset à partir du TransID)
  - Codé sur 2 octets : 0xC000+ Offset (sur 14 bits)

## Exemple

04gersC010

et à 0x10 du début on trouve 04cnam02fr00

## Drapeaux

| QR | Opcode | AA | TC | RD | RA | 000 | RCODE |
|----|--------|----|----|----|----|-----|-------|
| 1b | 4bits  | 1b | 1b | 1b | 1b | 3b  | 4bits |

- QR : Question (0)/Réponse (1)
- OPCODE : 0=Std Query, 4=notify, 5=update
- AA : Authoritative Answer (réponse d'un NS)
- TC : si =1 message tronqué car trop long
- RD : Récursion demandée.
- RA : si=1 Récursion disponible (Available)
- RCODE : Response code (0=pas d'erreur)

## Exemple compression DNS

Réponse à une requête de type "adresse de www.to.fr ?"

| TransID | Flags | nb S1 | nb S2 | nb S3 | nb S4 |
|---------|-------|-------|-------|-------|-------|
| AB CD   | 81 35 | 00 01 | 00 01 | 00 01 | 00 00 |
| 00 01   | 02 03 | 04 05 | 06 07 | 08 09 | 0A 0B |

| S1    |    |    |    | S2 ...                |
|-------|----|----|----|-----------------------|
| Qname |    |    |    | Qtype Qclass Name ... |
| 03    | 77 | 77 | 77 | 02 74 0F 02 66 72 00  |
| 0C    | 0D | 0E | 0F | ...                   |

C00C - C000 = 000C (→ 12e octet)

Gain: 2 octets (C00C) au lieu de 11...

Pour aller plus loin

- RFC : 1034, 1035, 2136, 2181, 4033, 6895
- <https://kb.isc.org/article/AA-01031> : docs spécifique sur bind 9
- ...