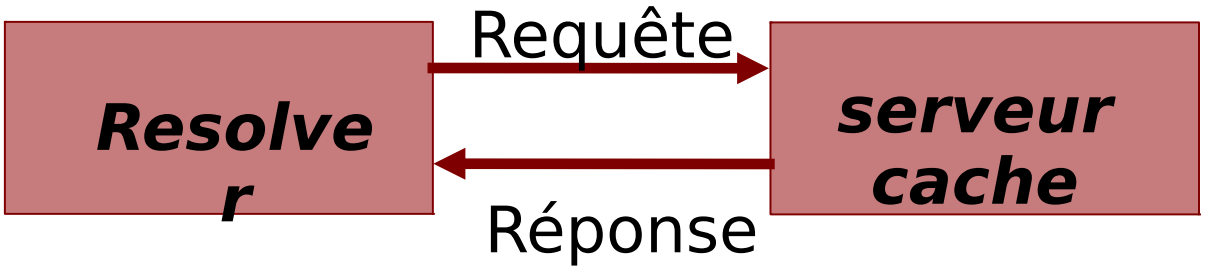
DNS Session 2: Fonctionnement du cache DNS et corrections du DNS

Arnaud A. A. AMELINA
AFNOG 2015

Comment fonctionne le serveur cache (1)

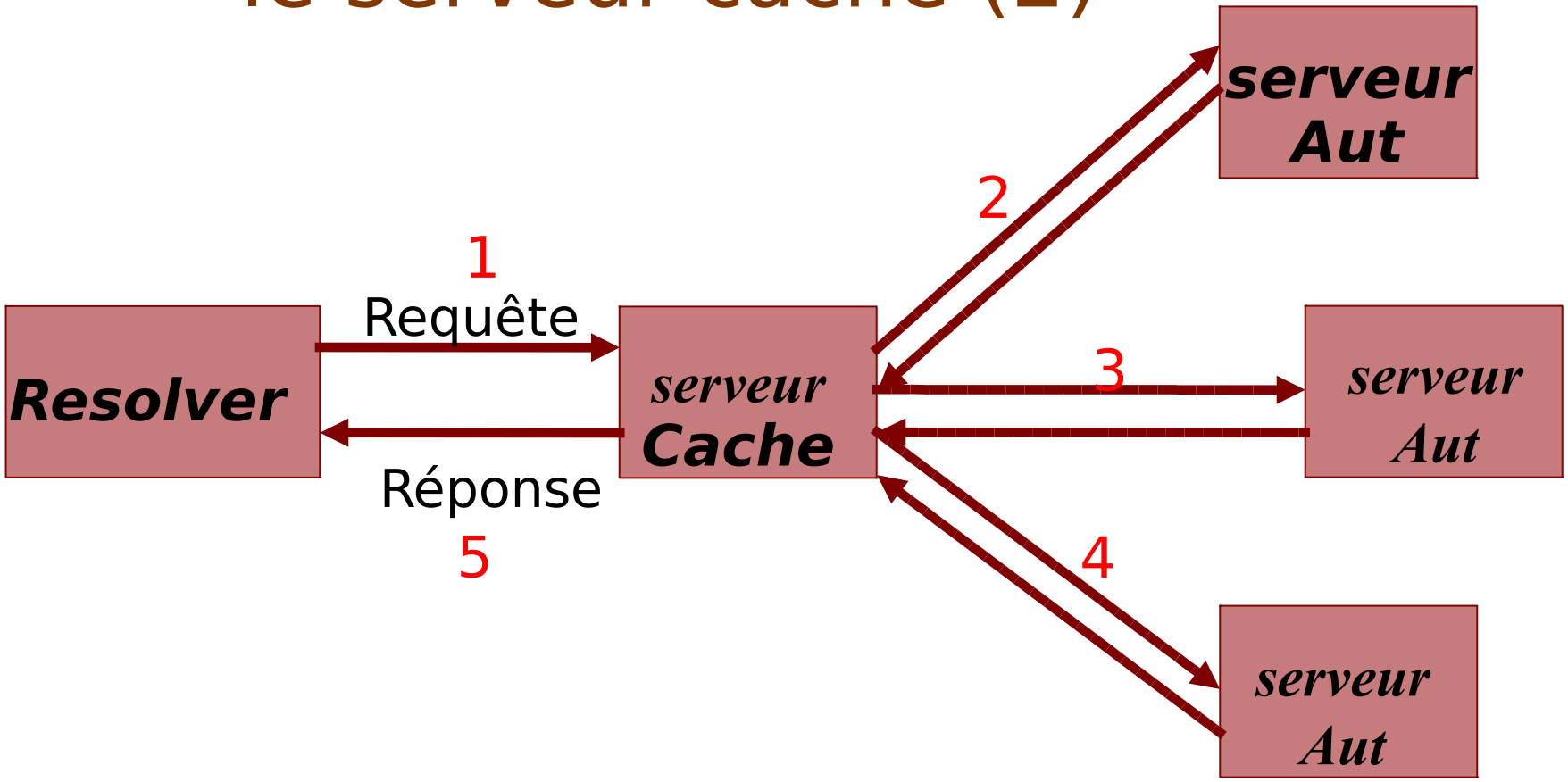


Si nous avons traité cette question récemment, la réponse est déjà dans le cache- facile!

Qu’est ce qui se passe si la réponse n’est pas dans le cache?

- DNS est une base de données distribuée : les parties de l’arborescence (appelées "zones") sont gardées sur de différents serveurs
- Ils sont appelés les “serveurs autoritaires" pour leur partie particulière de l'arborescence (zones)
- C’est la tâche d’un serveur cache de localiser le bon serveur autoritaire et de récupérer le résultat
- Il peut devoir demander à d'autres serveurs de nom de localiser celui dont il a besoin

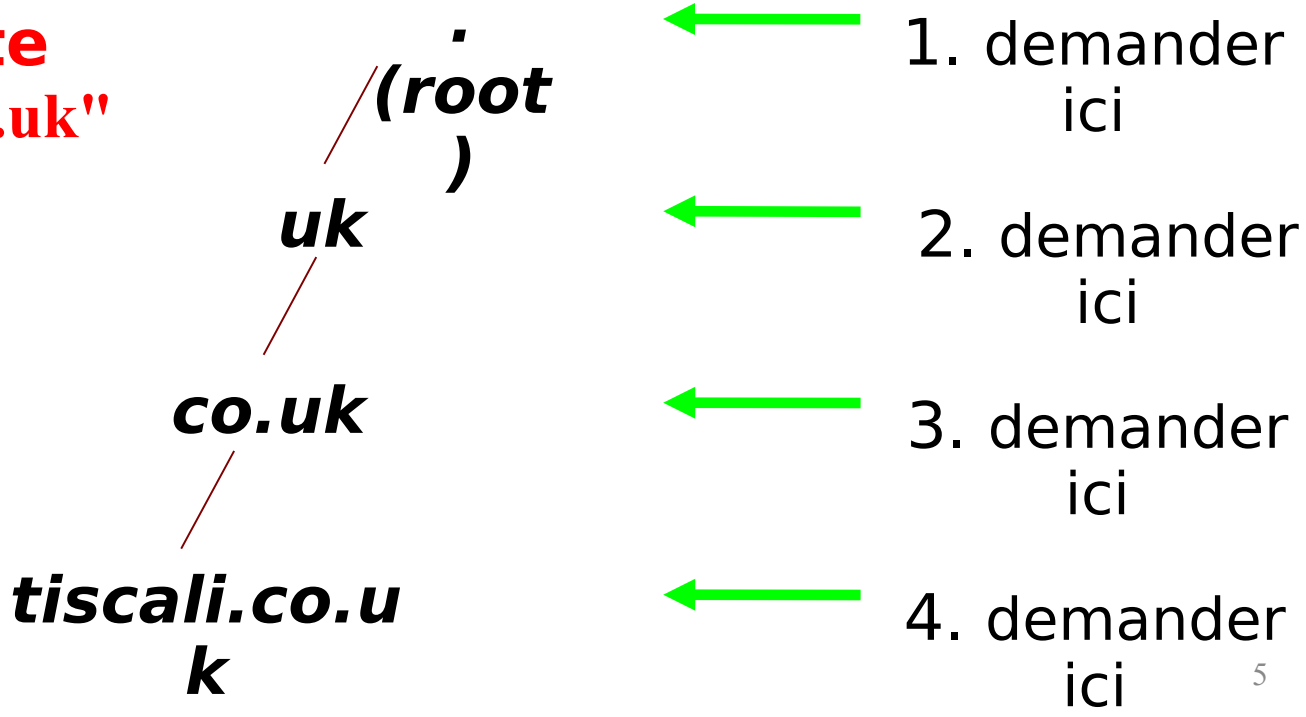
Comment fonctionne le serveur cache (2)



Comment sait-il à quel serveur autoritaire demandé?

Il suit la structure hiérarchique de l'arborescente

e.x. requête
"www.tiscali.co.uk"



Les serveurs intermédiaires retournent les enregistrements de ressources " NS "

- "Je n’ai pas la réponse, mais essayez ces autres serveurs de nom à la place "
- appelés une RÉFÉRENCE (REFERRAL)
- Déplacez-vous en bas de l'arbre par un ou plusieurs niveaux

Ce processus pourra soit :

- Trouver un serveur autoritaire qui connaît la réponse (positive ou négative)
- Ne trouvé aucun serveur de nom fonctionnel : SERVFAIL
- Terminer au serveur de noms défectueux

  -Ne peut répondre et aucune autre délégation, ou réponse fausse!

(Note: Le serveur cache peut s'avérer également être un serveur autoritaire pour les requêtes. Dans ce cas, il peut répondre immédiatement sans demander n'importe où ailleurs. Nous parlerons plus tard pourquoi c'est une bonne idée d'avoir les machines séparées pour les serveurs cache et autoritaires

Comment ce processus commence t-il ?

Chaque serveur cache est configuré avec une liste de serveurs racines

/etc/bind/named.conf.default-zones

```
zone "." {
        type hint;
        file "/etc/bind/db.root";
};
```

/etc/bind/db.root

```
;
.                       3600000  IN  NS    A.ROOT-SERVERS.NET.
A.ROOT-SERVERS.NET.     3600000      A     198.41.0.4
A.ROOT-SERVERS.NET.     3600000      AAAA  2001:503:BA3E::2:30
;
; FORMERLY NS1.ISI.EDU
;
.                       3600000      NS    B.ROOT-SERVERS.NET.
B.ROOT-SERVERS.NET.     3600000      A     192.228.79.201
;
; FORMERLY C.PSI.NET
;
…..etc
```

D'où provient le db.root ?

- ftp://ftp.internic.net/domain/named.cache
  - Intéressant de vérifier tous les 6 mois et ainsi

Démonstration

- ***dig +trace www.tiscali.co.uk.***

- Au lieu d'envoyer la requête au cache, " dig +trace " traverse l'arborescence de la racine et montre les réponses qu'elle obtient.

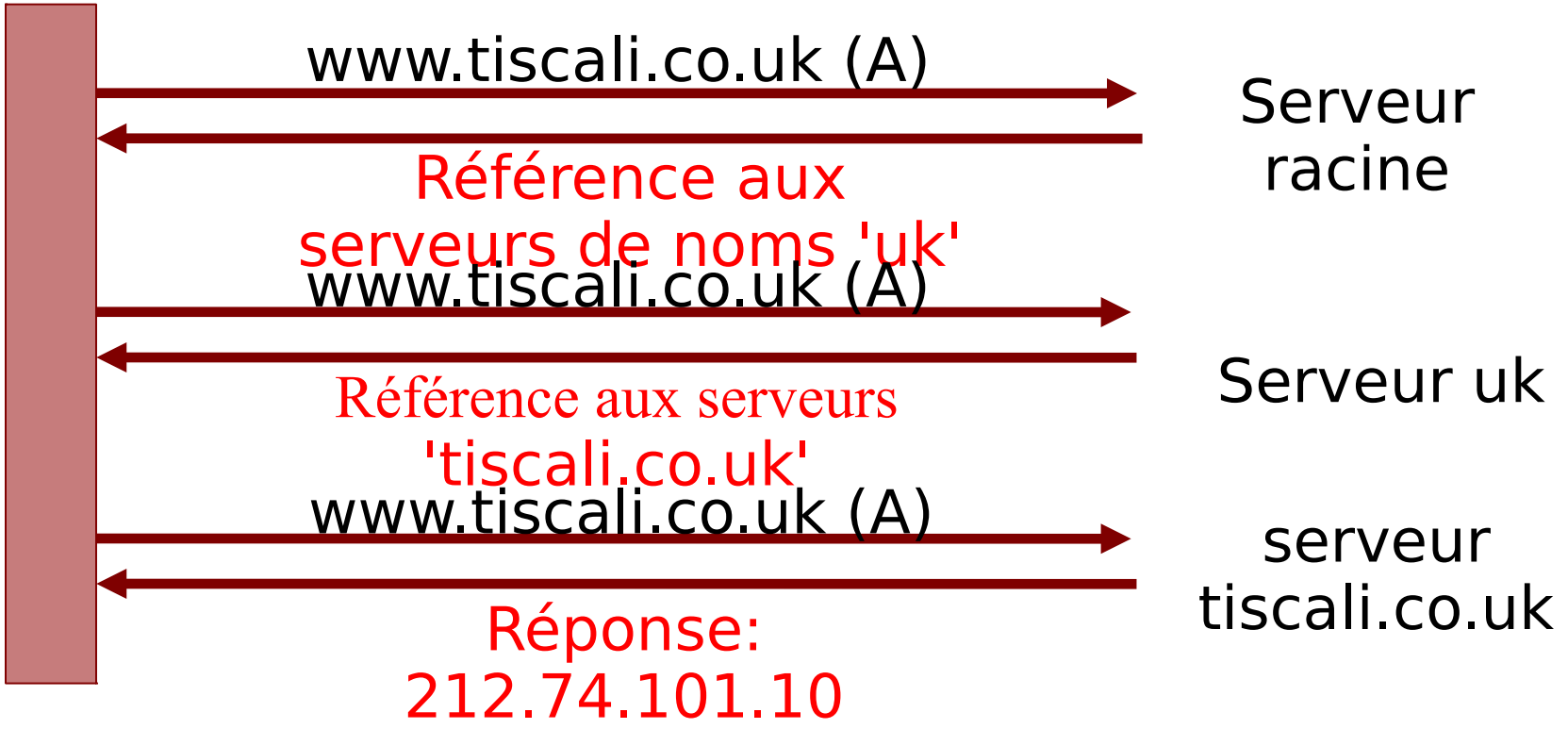
Les systèmes distribués ont beaucoup de points d'échec!

- Tellement chaque zone a deux ou plus de serveurs de noms autoritaires pour sa résilience
- Ils sont tous équivalents et peuvent être essayés dans n'importe quel ordre
- L'essai s'arrête dès que l'un d'eux donne une réponse
- Aident également le partage de charge
- Les serveurs racines sont occupés - il y a actuellement 13 d'entre eux (dont chacun est un grand groupe)

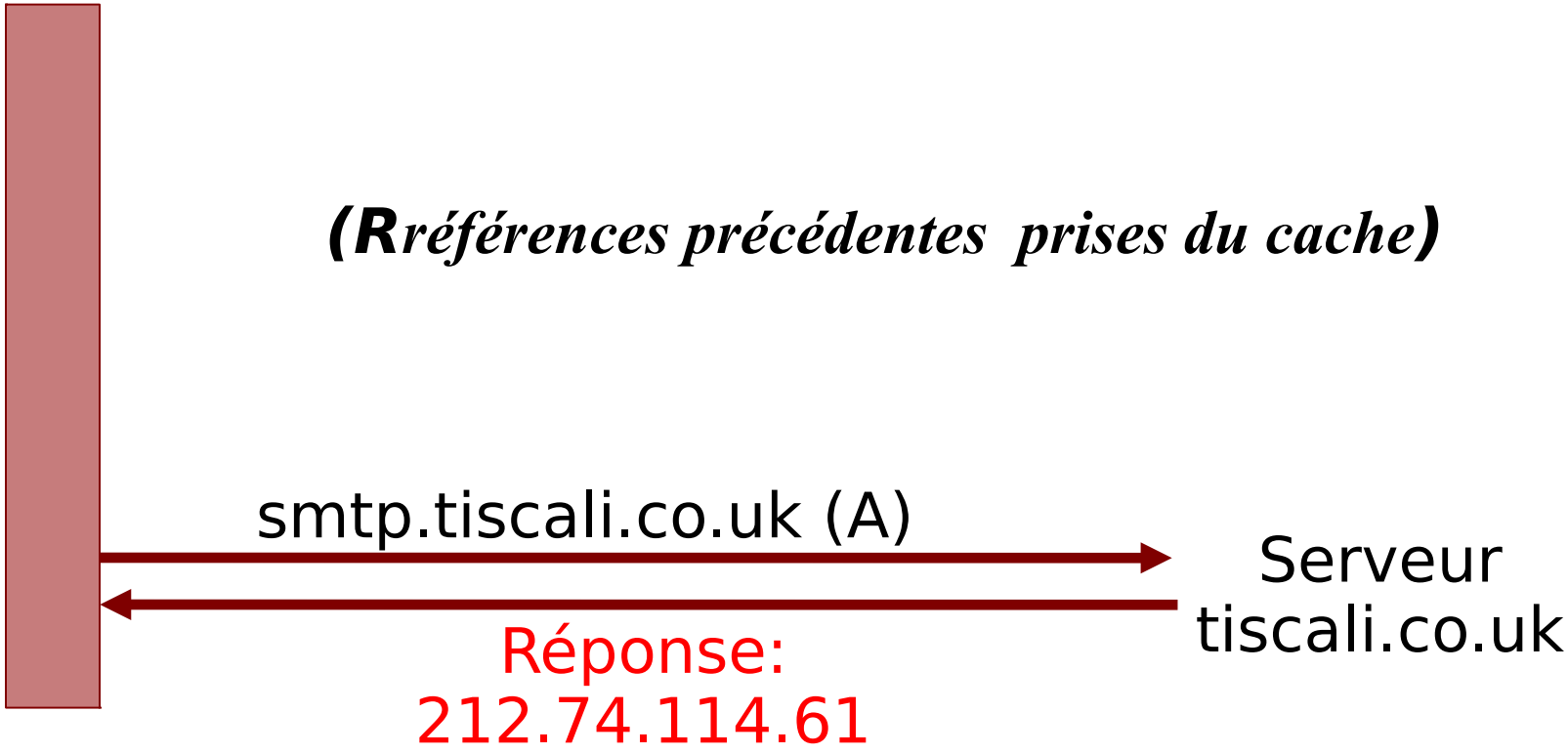
Le cache réduit la charge sur les serveurs autoritaires

- Particulièrement important au niveau plus élevé : serveurs racines, serveurs GTLD (.com, .net etc)
- Toute information intermédiaire est cachée comme la réponse finale - ainsi que les enregistrements NS des RÉFÉRENCES

Exemple 1: www.tiscali.co.uk (sur un cache vide)



Exemple 2: smtp.tiscali.co.uk (après l'exemple précédent)



Les caches peuvent être un problème si les données deviennent anciennes

- Si les serveurs caches gardent des données pendant trop longtemps, ils peuvent distribuer des réponses fausses si les données autoritaires changeaient
- Si les serveurs caches contiennent des données pendant trop peu de temps, cela signifie du travail accru pour les serveurs autoritaires

Le propriétaire d'un serveur autoritaire peut contrôler comment leurs données sont cachées

- Chaque enregistrement (RR) a un temps de vie "Time To Live" (TTL) qui indique combien de temps il peut être maintenu dans le cache

- L'enregistrement SOA indique combien de temps une réponse négative peut être cachée (c.-à-d. la non-existence d'un enregistrement)

(Le propriétaire du serveur cache n'a aucun contrôle - mais ils ne le voudraient pas de toute façon)

Une politique de compromis

- Définir le TTL assez long – 1 ou 2 jours
- Quand vous savez que vous êtes sur le point de faire un changement, réduisez le TTL à 10 minutes
- Attendre 1 ou 2 jours AVANT DE faire le changement
- Après le changement, remettre encore le TTL

Quelle sorte de problèmes pourrait se produire quand un serveur cache fonctionne?

- Se rappeler que suivre les références est en général un multiple processus pas à pas
- Se rappeler aussi du cache

(1) Un serveur autoritaire est en panne ou inaccessible

- Pas de problème : arrêter et essayer le prochain serveur de nom (se rappeler qu'il y a de serveurs autoritaires multiples pour une zone, donc la référence retourne des enregistrements NS multiples).

(2) *TOUS* les serveurs autoritaires sont en panne ou inaccessibles!

- C'est mauvais: la requête ne peut s'exécuter
- Vérifier que tous les serveurs de noms ne sont pas sur le même sous-réseau (l'échec de commutateur ou/et routeur)
- Vérifier que tous les serveurs ne sont pas dans le même bâtiment (la panne de courant)
- Vérifier que tous les serveurs ne sont pas sur le même backbone (l'échec de lien ascendant)
- Pour plus de détail lire RFC 2182

(3) La référence pointe sur un serveur qui n'est pas autoritaire pour cette zone

- Mauvaise erreur, appelée "Lame Delegation“
- La requête ne peut s’exécuter - le serveur n'a plus la bonne réponse ou la bonne délégation

Erreur Typique : l’enregistrement NS pointe sur le serveur cache qui n’a pas été configuré comme autoritaire pour cette zone

- Ou : Une erreur de syntaxe dans le fichier de zone ce qui fait que le logiciel du serveur de nom l'ignore

(4) contradictions entre les serveurs autoritaires

- Si les serveurs autoritaires n'ont pas la même information alors vous obtiendrez différentes informations selon lequel vous a sélectionné (aléatoire)
- En raison du cache, il peut être très difficile de corriger ces problèmes. Le problème est intermittent.

(5) contradictions dans les délégations

- Les enregistrements NS dans la délégation ne correspondent pas aux enregistrements NS dans le fichier de zone (nous écrirons les fichiers de zone plus tard)

- Lequel est exact?

(6) Mélange du Cache et des serveurs autoritaires

- Si le serveur cache contient un ancien fichier de zone, mais le client a transféré leur DNS quelque part d'autre

- Le serveur cache répond immédiatement avec l'ancienne information, quoique l'enregistrement NS pointe sur les serveurs autoritaires des ISP différents qui tiennent la bonne

(7) Choix inadéquat des paramètres

- Par exemple TTL est défini trop court ou trop long

Ces problèmes ne sont pas la faute du serveur cache!

- Ils proviennent tous de la mauvaise configuration des serveurs AUTORITAIRES
- Plusieurs de ces erreurs sont faciles à faire mais difficile à corriger, particulièrement à cause du cache
- Faire fonctionner un serveur cache est facile . Faire fonctionner correctement le serveur autoritaire exige une grande attention aux

Comment corriger ces problèmes?

- Nous devons dévier le cache
- Nous devons essayer tous les serveurs de noms pour une zone (un serveur cache s'arrête après un seule)
- Nous devons dévier la récursivité pour examiner toutes les références intermédiaires
- "dig +norec" est votre ami

```
dig +norec @1.2.3.4 foo.bar. a
```

Serveur de requête — Domaine — Type de requête

Comment interpréter des réponses (1)

- Rechercher "status: NOERROR"
- "flags :.... ***aa***" signifie que c'est une réponse autoritaire (c.-à-d. non caché)
- "ANSWER SECTION" donne la réponse
- Si vous récupérez juste les enregistrements NS : c'est une référence

```
;; ANSWER SECTION
foo.bar.   1H IN A   1.2.3.4
```

Nom de domaine — TTL — Réponse

Comment interpréter des réponses(2)

- **"status: NXDOMAIN"**
  - Réponse négative (le domaine n'existe pas). Vous devriez récupérer un SOA
- **"status: NOERROR" avec zéro ERs**
  - Réponse négative (le domaine existe mais aucun enregistrement du type demandé).Vous devriez récupérer un SOA
- D'autres états peuvent indiquer une erreur
- Regarder également l'expression " Connection Refused " (le serveur DNS n'est pas démarré) ou timeout (pas de réponse)

Comment corriger un domaine en utilisant " dig +norec "(1)

- Commencer avec n'importe quel serveur racine

dig +norec @a.root-servers.net. www.tiscali.co.uk. a

Rappelez vous du point à la fin des noms de domaine!

- Pour une référence noter les ERs NS retournés
- Répéter la requête pour *tous* les ERs NS
- Retourner à l'étape 2, jusqu'à ce que vous obteniez les réponses finales des requêtes

Comment corriger un domaine en utilisant '" dig +norec "(2)

- Vérifier toutes les réponses qui ont "flags: aa"
  et ces réponses d'un groupe de serveurs autoritaires qui sont conformées avec l'un et l'autre
- Noter que les enregistrements NS sont des noms et non des adresses. Donc, vérifier maintenant que chaque enregistrement NS correspond aux adresses IP qui utilise le même processus!

Comment corriger un domaine en utilisant '" dig +norec "(3)

- Pénible, exige la patience et l'exactitude, mais elle paye au loin
- Apprenez ceci premièrement avant de jouer avec des outils plus automatisés, par exemple.http://zonecheck.nic.fr/v2/

Exemples

Construction de votre propre serveur cache

- Facile!
  - Logiciel standard est le "bind" (Berkeley Internet Name Daemon) de ISC: www.isc.org
  - La plupart des système UNIX l'ont, et déjà configuré comme un cache
    - Les paquetages Red Hat: "bind" et "caching-nameserver"
- Quelle sorte de matériel choisiriez-vous au cours de la réalisation d'un DNS cache ?

Amélioration de la configuration

- Limiter les accès des clients à vos propres adresses IP seulement
  - Aucune raison pour que les autres sur l'Internet utilisent vos ressources de cache
- Faire le cache autoritaire pour les requêtes qui ne devraient pas aller sur l'Internet
  - localhost ➜ A 127.0.0.1
  - 127.0.0.1 ➜ PTR localhost.
  - RFC 1918 (10/8, 172.16/12, 192.168/16)
  - Donne une réponse plus rapide et sauve des envois de requêtes inutiles à Internet

Configuration de BIND : /etc/bind/named.conf

Bind9 sous ubuntu utilise 3 fichiers :

```
include "/etc/bind/named.conf.options";
include "/etc/bind/named.conf.local";
include "/etc/bind/named.conf.default-zones";
~
```

Configuration de la limitation d'accès : ACL

Par défaut votre serveur cache est ouvert à tous les réseaux, il faut donc limiter l'accès par une liste d'accès ACL. Compléter les lignes rouges ci-dessous au fichier :

vi /etc/bind/named.conf.options

……

```
acl mynetwork {
    127.0.0.1;
    10.10.0.0/27;
};
…..
options {
    directory "/var/cache/bind";
    recursion yes;
    allow-query { mynetwork; };
    # note: utiliser 'allow-recursion' en lieu et place si votre serveur jour les rôles
    # de cache serveur et serveur  autoritaire
};
```

"localhost":/etc/bind/named.conf.default-zones

```
zone "localhost" {
    type master;
    file "/etc/bind/db.local";
    allow-update { none; };
};
```

/etc/bind/db.local

```
$TTL    604800
@       IN      SOA     localhost. root.localhost. (
                              2         ; Serial
                         604800         ; Refresh
                          86400         ; Retry
                        2419200         ; Expire
                         604800 )       ; Negative Cache TTL
;
@       IN      NS      localhost.
@       IN      A       127.0.0.1
@       IN      AAAA    ::1
```

127.0.0.1 reverse lookups : /etc/bind/named.conf.default-zones

```
zone "127.in-addr.arpa" {
    type master;
    file "/etc/bind/db.127";
    allow-update { none; };
};
```

/etc/bind/db.127

```
@       SOA     localhost.      root.localhost. (
                2014022800      ; serial
                8h              ; refresh
                1h              ; retry
                4w              ; expire
                1h )            ; negative TTL

@        NS      localhost.
1.0.0   PTR     localhost.
; Don't forget the trailing dots!
```

RFC1918 reverse lookups :

/etc/bind/named.conf.local inclut /etc/bind/zones.rfc1918

```
zone "10.in-addr.arpa" { type master; file "/etc/bind/db.empty"; };
zone "16.172.in-addr.arpa" { type master; file "/etc/bind/db.empty"; };
zone "17.172.in-addr.arpa"  { type master; file "/etc/bind/db.empty"};
# repeat for 18.172.in-addr.arpa
#     ...to  168.192.in-addr.arpa
```

/etc/bind/db.empty

```
@       SOA     localhost.      root.localhost. (
                2014022800      ; serial
                8h              ; refresh
                1h              ; retry
                4w              ; expire
                1h )            ; negative TTL
@        NS      localhost.
```

Administration un serveur cache

- /etc/init.d/bind9 start
- rndc status
- rndc reload
  - Après les changements de configuration; cause moins de rupture qu'en redémarrant le démon
- rndc dumpdb
  - /var/named/named_dump.db
- rndc flush
  - Détruit le contenu du cache; ne pas faire sur un système en fonction

Absolument critique!

- Vous DEVEZ vérifier le fichier /var/log/syslog après n'importe quel changement du serveur de noms
- Une erreur de syntaxe peut aboutir à un serveur de noms qui fonctionne, mais pas comme de la manière dont vous avez voulu
- BIND est très tatillon au sujet de la syntaxe
  - Prenez gare aux accolades (}) et au point-virgule (;)
  - Dans un fichier de zone, les commentaires commencent par le point-