DNS Session 2: Fonctionnement du cache DNS

Présenté par

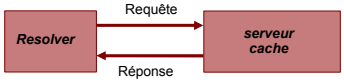
Alain Patrick AINA
Roger YERBANGA

RALL 2007
22 - 26 Novembre 2007
Rabat, Maroc

Historique du support de cours

- Création du support en septembre 2004
- Traduction du cours DNS AFNOG 2004 de
  - Alain AINA
  - Ayitey Bulley
  - Brian Candler
- Site Web des ateliers AFNOG
  http://www.ws.afnog.org

Comment fonctionne le serveur cache (1)

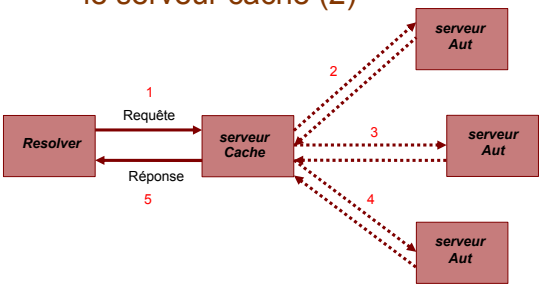


Si nous avons traité cette question récemment, la réponse est déjà dans le cache- facile!

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

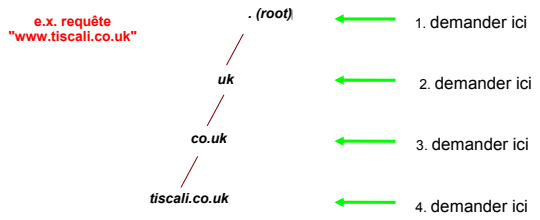
- DNS est une base de données distribuée : les parties de l'arborescence (appelées "zones") sont gardées sur de différents serveurs
- Ils sont appelés les "serveurs autoritaires" pour leur partie particulière de l'arborescence (zones)
- C'est la tâche d'un serveur cache de localiser le bon serveur autoritaire 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'arborescence



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 trouver aucun serveur de nom fonctionnel : SERVFAIL
- Terminer sur des serveurs 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/named.conf

```
zone "." {
        type hint;
        file "named.ca";
};
```

/var/named/named.ca

```
.                       3600000      NS    A.ROOT-SERVERS.NET.
A.ROOT-SERVERS.NET.     3600000      A     198.41.0.4

.                       3600000      NS    B.ROOT-SERVERS.NET.
B.ROOT-SERVERS.NET.     3600000      A     128.9.0.107

.                       3600000      NS    C.ROOT-SERVERS.NET.
C.ROOT-SERVERS.NET.     3600000      A     192.33.4.12

... etc
```

D'où provient le named.ca ?

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

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.

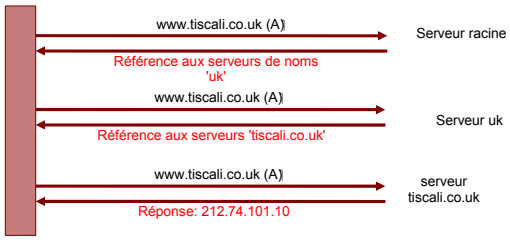
Les systèmes distribués ont beaucoup de points d'échec!

- Chaque zone a deux ou plusieurs serveurs de noms autoritaires pour la redondance
- 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 avec le partage de charge
- Les serveurs racines sont occupés
- Il y a actuellement 13 IPv4 de serveurs racine
  - Plus de 130 serveurs racine de par le monde avec le "anycast dns"

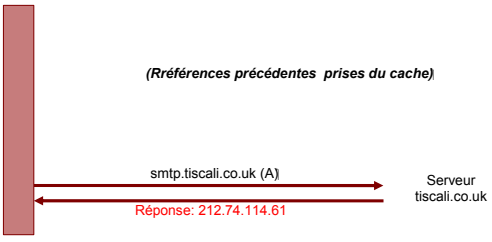
Le cache réduit la charge sur les serveurs autoritaires

- Particulièrement important au niveau plus élevé : serveurs racines, serveurs TLD(ccTLD,gTLD,arpa)
- Toute information intermédiaire est mise en cache 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 restent trop longtemps

- Si les serveurs caches gardent des données pendant trop longtemps, ils peuvent distribuer des réponses fausses si les données autoritaires changent
- Si les serveurs caches contiennent des données pendant trop peu de temps, cela signifie du travail accru pour les serveurs autoritaires
- Un compromis s'impose !!!

L'administrateur d'une zone peut contrôler la durée de vie des données dans les caches

- Chaque enregistrement (ER) 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 gardée en cache (i.e. la non-existence d'un enregistrement)

(L'administrateur du serveur cache n'a aucun contrôle)

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 à son niveau initial

Quels sont les problèmes qui peuvent se produire avec un serveur cache ?

- Se rappeler que suivre les références est en général un processus à plusieurs étapes
- Se rappeler 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 plusieurs serveurs autoritaires 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 aboutir
- S'assurer que tous les serveurs de noms ne sont pas sur le même sous-réseau (l'échec de commutateur ou/et routeur)
- S'assurer que tous les serveurs ne sont pas dans le même bâtiment (la panne de courant)
- S'assurer que tous les serveurs ne sont même pas sur le même backbone (l'échec de lien ascendant)
- Pour plus de détails, 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 aboutir - le serveur n'a non plus la bonne réponse ou la bonne référence
- Erreur Typique : l'enregistrement NS pointe sur le serveur cache qui n'a été pas configuré comme autoritaire pour cette zone
- Ou: une erreur de syntaxe dans le fichier de zone a contraint le serveur à ignorer la zone

(4) contradictions entre les serveurs autoritaires

- Si les serveurs autoritaires n'ont pas la même information alors vous obtiendrez différente information selon celui que vous avez sélectionné (aléatoire)
- En raison du cache, il peut être très difficile de détecter et de corriger ces problèmes. Souvent intermittent.

(5) contradictions dans les délégations

- Les enregistrements NS dans la délégation ne correspondent pas aux enregistrements NS dans le fichier de zone enfant (nous écrirons les fichiers de zone plus tard)
- Lequel est exact?
- Ceux de la zone enfant sont plus autoritaires

(6) Mélange du Cache et des serveurs autoritaires

- Si le serveur cache contient un ancien fichier de zone, mais le client a transféré sa zone 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 information!

(7) Choix inadéquat des paramètres

- Par exemple TTL 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 au détail

Comment corriger ces problèmes?

- Nous devons éviter le cache
- Nous devons essayer tous les serveurs de noms pour une zone (un serveur cache s'arrête après un seule)
- Nous devons éviter 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 ces réponses (1)

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

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 les expressions " Connection Refused " (le serveur rejecte votre requête) ou timeout (pas de réponse)

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

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!

2.1 Pour une référence noter les NS retournés

2.2 Répéter la requête pour *tous* les 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 vérifier si les réponses des NS autoritaires sont conformées les uns avec les autres

Noter que les enregistrements NS sont des noms et non des adresses. Vérifier donc que chaque enregistrement NS correspond aux adresses IP données en utilisant le même processus!

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

- Pénible, exige la patience et un travail de fourmi, 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"
- Quel type 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 votre serveur cache
- Faire du cache, un 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 BIND dans le fichier : /etc/named.conf

```
acl mynetwork {
    127.0.0.1;
    192.188.58.64/26;
};

options {
    directory "/var/named";
    recursion yes;   # this is the default
    allow-query { mynetwork; };
    # note: use 'allow-recursion' instead if your
    # nameserver is both caching and authoritative
};

controls {
    inet 127.0.0.1 allow { localhost; } keys { rndckey; };
};
zone "." {
    type hint;
    file "named.ca";
};
```

"localhost"

```
zone "localhost" {
    type master;
    file "localhost.zone";
    allow-update { none; };
};
```

/var/named/localhost.zone

```
@       SOA     localhost.    root.localhost. (
                2004022800    ; serial
                8h            ; refresh
                1h            ; retry
                4w            ; expire
                1h )          ; negative TTL

        NS      localhost.
        A      127.0.0.1
```

127.0.0.1 reverse lookups

```
zone "0.0.127.in-addr.arpa" {
    type master;
    file "named.local";
    allow-update { none; };
};
```

/var/named/named.local

```
@       SOA     localhost.    root.localhost. (
                2004022800    ; serial
                8h            ; refresh
                1h            ; retry
                4w            ; expire
                1h )          ; negative TTL

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

RFC1918 reverse lookups

```
zone "168.192.in-addr.arpa" {
    type master;
    file "null.zone";
};
zone "10.in-addr.arpa" {
    type master;
    file "null.zone";
};
# repeat for 16.172.in-addr.arpa
#      ...to  31.172.in-addr.arpa
```

/var/named/null.zone

```
@       SOA     localhost.    root.localhost. (
                2004022800    ; serial
                8h            ; refresh
                1h            ; retry
                4w            ; expire
                1h )          ; negative TTL

        NS      localhost.
```

Administration un serveur cache

- /etc/rc.d/init.d/named 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 opérationnel

Absolument critique!

- Vous DEVEZ vérifier le fichier /var/log/messages après n'importe quel changement du serveur de noms
- Une erreur de syntaxe peut aboutir à un serveur de noms qui fonctionne, mais pas de la manière voulue
- 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-virgule (;) NON dièse (#)