DNS Session 3: Configuration du serveur autoritaire

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

RECAPITULATION

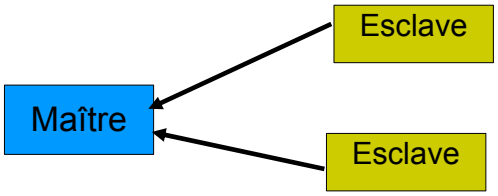
- Le DNS est une base de données distribuée
- Le resolver s'adresse au serveur Cache pour l'information
- Le cache traverse l'arborescence du DNS pour trouver le serveur autoritaire qui a l'information demandée.
- Une mauvaise configuration des serveurs autoritaires peut aboutir à la panne du domaine

REPLICATION DNS

- Pour chaque domaine, nous avons besoin de plus d'un serveur autoritaire avec la même information (RFC 2182)
- Les données sont enregistrées sur un seul serveur
- (maître) et répliquées sur les autres (les esclaves)
- Le monde extérieur ne "peut faire la différence" entre le maître et le slave
  - Les enregistrements NS sont retournés de façon aléatoire pour le partage de charge égal

Sont aussi appelés "primaire" et "secondaire"

Les esclaves se connectent au maître pour rechercher la copie des données de la zone



- Le maître n'envoie pas automatiquement les données aux esclaves

Quand est-ce que la replication a lieu ?

- L'esclave scrute le maître périodiquement– appelé ``Temps de rafraîchissement``
  - A l'origine c'était le seul mécanisme
- Une extension du protocole permet maintenant au maître d'informer les esclaves quand les données ont changé
  - Aboutit à des mises à jours plus rapides
  - DNS NOTIFY
- La notification est incertaine (ex. le réseau peut perdre le paquet); ainsi nous avons toujours besoin de contrôler l'intervalle de rafraîchissement

Le numéro de série

- Chaque fichier de zone a un numéro de série (Serial Number)
- L'esclave copiera les données uniquement quand le numéro de série a changé (AUGMENTE)
  - Les requêtes UDP périodiques pour vérifier le numéro de série
  - S'il a augmenté, transfert par TCP des données de zone
- C'est votre responsabilité d'augmenter le numéro de série après chaque changement, autrement les esclaves et le maître seront contradictoires

Format de numéro de série recommandé: AAAAMMJJNN

- AAAA = Année
- MM = mois (01-12)
- JJ = jour (01-31)
- NN=numéro de changements par jour (00-99)
  - Ex. Si vous changez le fichier le 27 septembre 2004, le numéro de série sera 2004092900. Si vous le changez encore une fois le même jour, ce sera 2004092901

Numéro de série: Danger 1

- Si jamais vous diminuez le numéro de série, les esclaves ne se mettront plus jamais à jour jusqu'à ce que le numéro de série aille au dessus de sa précédente valeur
- Au pire, vous devez entrer en contact avec tous vos esclaves et obtenir d'eux de supprimer leur copie de zone

Numéro de série: Danger 2

- Le numéro de série est un nombre de 32 bits non signé
- Choix : 0 à 4294967295
- Toute valeur plus grande que celle-ci est silencieusement tronquée
- Ex. 20040303000 (noté un chiffre extra)
  - = 4AA7EC198 (hex)
  - = AA7EC198 (32 bits)
  - = 2860433816
- Si vous faites cette erreur et que vous la corriger la, le numéro de sérié aura diminué

Configuration du Maître

- /etc/namedb/named.conf pointe sur le fichier de zone (créé manuellement)
- Choisissez un emplacement logique pour le garder
- Ex. /etc/namedb/m/example.com
- Ou /etc/namedb/m/com.example

```
Zone "example.com" {
    type master;
    file"m/example.com";
Allow-transfer { 192.188.58.126; 192.188.58.2; };
};
```

Configuration de l'esclave

- /etc/named.conf pointe vers l'adresse IP du maître et l'emplacement du fichier de zone
- Le fichier de zone sont transféré automatiquement

```
Zone "example.com" {
    type slave;
    masters { 192.188.58.126; }
    file "s/example.com";
    allow-transfer  { none;};
};
```

Maître et esclave

- Un serveur peut-être maître pour certaines zones et esclave pour d'autres au même moment
- C'est pourquoi nous recommandons de maintenir les fichiers dans des répertoires différents
  - /var/named/m
  - /var/named/s

allow-transfer { ....; }

- Les machines à distance peuvent demander le transfert du contenu entier d'une zone
- Par défaut, ceci est autorisé à n'importe qui
- Vaut mieux limiter ceci
- Vous pouvez en fixer un par défaut, et passer les autres dans la configuration de chaque zone s'il y a lieu.

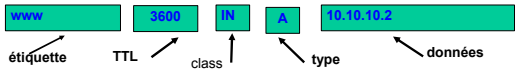
```
Options {
    allow-transfer { 127.0.0.1; };
};
```

La structure d'un fichier de zone

- Options globales
  - $TTL 1d
  - Fixe le TTL par défaut pour tous les enregistrements
- ER SOA
  - "Start Of Authority"
  - Indique les informations de gestion pour la zone
- Les ERs NS
  - Listent les serveurs du nom de domaine, le maître et les esclaves
- Les autres ERs
  - Les données réelles que souhaitez publier

Format des Enregistrements de Ressources (ERs)

- Un par ligne (excepté SOA qui peut se prolonger au-delà de plusieurs lignes)
- Si vous omettez le nom de domaine, c'est pareil que la ligne précédente
- Les raccourcis TTL : Ex. 60s, 30m, 4h, 1w2d
- Si vous omettez le TTL, il prend la valeur $TTL par défaut
- Si vous omettez la Classe (Class), il se réfère à la classe IN
- Le type et les données ne peuvent pas être omis
- Les commentaires commencent avec le Point-virgule (;)



Les raccourcis

- Si un nom ne se termine pas par un point, le nom du domaine("origin") est ajouté
- Un nom de domaine de "@" signifie l'origine elle-même
- Ex. dans le fichier de zone de example.com
  - @ signifie example.com.
  - www signifie www.example.com.

Si vous écrivez ceci

```
$TTL 1d
@           SOA ( . . . . )
            NS   ns0
            NS   ns0.as9105.net.
;Serveur de messagerie et web
www         A    212.74.12.80
            MX   10 mail
```

Il devient ceci

```
example.com 86400   IN   SOA (. . . )
example.com  86400   IN   NS   ns0.example.com.
example.com  86400   IN   NS   ns0.as91504.net.
www.example.com. 86400   IN   A    212.74.112.80
www.example.com. 86400   IN   MX   10 mail.example.com.
```

Format de l'enregistrement SOA

```
$TTL 1d
@    1h      IN  SOA    ns1.example.net.   brian.nsrc.orig. (
                2004030300     ;Serial
                8h       ;Refresh
                1h       ;Retry
                4w       ;Expire
                1h)      ;Negative
          IN   NS  ns1.example.net.
          IN   NS  ns2.example.net.
          IN   NS  ns1.othernetwork.com.
```

Format de l'enregistrement SOA

- ns1.example.net
  - Nom du serveur maître
- brian.nsrc.org.
  - L'adresse électronique de la personne responsable, avec “@” changé en point
- Numéro de série (serial Number)
- Intervalle de rafraîchissement (Refresh interval)
  - Après combien de temps, l'esclave vérifie le numéro de série sur le maître et initie un transfert de zone si nécessaire
- Intervalle de nouvelle tentative (Retry Interval)
  - Après combien de temps, l'esclave essaye de recontacter le maître pour vérifier le numéro de série

Format de l'enregistrement SOA

- Temps d'expiration (Expire time)
  - Si l'esclave ne peut entrer en contact avec le maître pour cette période, il arrêtera de servir la zone
- Negative / Minimum
  - Etait utilisé comme valeur minimum de TTL
  - Maintenant, ceci est utilisé pour la cache négatif: indique combien de temps le cache peu stocker la non-existence d'un ER
- RIPE-2003 a recommandé des valeurs
  - http://www.ripe.net/ripe/docs/dns-soa.html

Format de l'enregistrement NS

```
$TTL 1d
@   1h   IN   SOA      ns1.example.net.   Brian.nsrc.org. (
                       2004030300   ; Serial
                       8h           ; Refresh
                       1h           ; Retry
                       4w           ; Expiry
                       1h)          ; negative
          IN   NS   ns1.example.net.
          IN   NS   ns2.exaple.net.
          IN   NS   ns1.othernotwork.com.
```

- **Liste tous les serveurs de noms autoritaires de la zone (maître et les esclaves)**
- **Doit utiliser les noms complètement qualifiés**

Doit pointer vers des noms et non vers des adresses IP

Format des autres ERs

- IN A 1.2.3.4
- IN MX 10 mailhost.example.com
  - Le nombre est une "valeur de préférence". Le courrier est délivré en premier au serveur ayant le “plus petit nombre” comme préférence
  - Doit pointer vers le NOM , et non vers l'adresse IP
- IN CNAME host.example.com.
- IN PTR host.example.com
- IN TXT "tout texte que vous voulez"

Quand vous modifiez le fichier de zone

- Vérifiez le numéro de série
- Vérifiez le fichier de zone

'named-checkzone example.com /etc/namedb/m/example.com”

  - Dispositif de Bind 9
  - Il rapporte les erreurs de syntaxe; corrigez les
- rndc reload
  - ou : rndc reload exemple.com
- Tail /var/log/messages

Voir page suivante →

Ces vérifications sont essentielles

- Si vous avez une erreur dans le fichier named.conf ou dans le fichier de zone, “named” continuera de fonctionner, mais sans charger la mauvaise zone
- Vous serez "une boiteuse” pour la zone sans le savoir
  - “Lame delegation”
- Les esclaves ne pourront pas contacter le maître
- Par la suite (ex. pendant 4 semaines plus tard) les esclaves ne serviront plus la zone
- Votre domaine arrêtera de fonctionner

D'autres vérifications que vous pouvez faire

- Dig +norec @x.x.x.x example.com. Soa
- Vérifie le “flag” AA
- Vérifie le maître et tous les esclaves
- Vérifie que les numéro de séries correspondent
- Dig @x.x.x.x example.com. axfr
  - "Transfert autoritaire"
  - Demande la copie complète du fichier de zone à travers TCP, comme les esclaves le font pour le maître
  - Ceci fonctionnera uniquement à partir des adresses IP énumérés dans la section de : allow-transfer { .....}

Vous avez donc maintenant des serveurs autoritaires fonctionnels

- Mais rappelez vous qu'aucun d'eux ne pourra être consulté dans la hiérarchie du DNS, jusqu'à ce que vous ayez la délégation du domaine au-dessus de vous (parent).
- “Le parent” devra mettre des ERs NS pour votre domaine en pointant vers vos serveurs de noms.
- Ces ERs doivent correspondent avec ceux que vous avez placé dans votre fichier de zone

Les 10 erreurs principales pour les serveurs autoritaires

- Tous les opérateurs de serveurs autoritaires doivent lire le RFC 1912
  - Fonctionnement classique du DNS et les erreurs de configuration
- Lire aussi le RFC 2182
  - Sélection et fonctionnement des serveurs DNS secondaires

Les erreurs de numéro de série

- Oublier d'incrémenter le numéro de série
- Incrémenter le numéro de série puis le décrémenter
- Utiliser le numéro de série plus grand que 2^{32}
- Impact :
  - Les esclaves ne sont pas à jour
  - Le maître et les esclaves ont des données contradictoires
  - Les caches obtiennent parfois les nouvelles données et parfois les anciennes – Problème intermittent

Les commentaire dans les fichiers de zone commencent avec " ; " au lieu de " # "

- Erreur de syntaxe dans le fichier de zone
- Le maître n'est plus autoritaires pour la zone
- Les esclaves ne peuvent pas vérifier le SOA
- La zone va expirer chez les esclaves
- Utilisez 'named-checkzone"
- Utilisez "tail /var/log/messages"

Manque du point à la fin des noms

```
;    zone example.com.
@ IN MX     10 mailhost.example.com
```

devient

```
@ IN MX     10      mailhost.example.com.example.com.
```

```
;  zone          2.0.192.in-addr.arpa.
1    IN          PTR      host.example.com
```

devient

```
1    IN   PTR     host.example.com.2.0.192.in-addr.arpa.
```

ERs NS et MX pointant vers des adresses IP

- Ils doivent pointer vers les nom d'hôtes, pas vers les adresses IP
- Malheureusement, certains serveurs de messagerie acceptent des adresses IP dans les enregistrements MX, Ainsi vous n'aurez pas les mêmes problèmes avec tous les sites distants

Les esclaves ne peuvent pas transférer le fichier de zone

- Accès limité par allow-transfer { ....} et les esclaves ne sont pas listés.
- Ou les filtres IP ne sont pas bien configurés
- L'esclave sera “boiteuse” (lame server) car non autoritaire

“Délégation boiteuse –Lame delegation”

- Vous ne pouvez pas lister simplement n'importe quel serveur dans les ERs NS pour votre domaine
- Vous devez obtenir l'accord de l'administrateur du serveur de noms et il doit le configurer comme esclave pour votre zone
- Au mieux: une résolution DNS plus lente et un manque de redondance
- Au pire; échecs intermittents pour résoudre votre domaine

Aucune délégation du tout

- Vous pouvez configurer " example.com" sur vos serveurs de noms mais le monde extérieur ne leur enverra pas des requêtes jusqu'à ce que vous ayez la délégation.
- Le problème est parfois masqué si votre serveur de noms agit comme votre cache et comme un serveur autoritaire
  - Vos propres clients peuvent résoudre www.example.com, mais le reste du monde ne peut pas

Les enregistrements « glue records »

- A voir plus tard

La mauvaise gestion du TTL pendant les changement

- Ex. Si vous avez un TTL de 24 heures, et vous changez www.example.com pour le pointer vers le nouveau serveur;
  - alors il aura une longue période pendant laquelle certains utilisateurs contacteront l'ancienne machine et certains la nouvelle
- Suivre cette procédure
  - Diminuer le TTL à 10 minutes
  - Attendre au moins 24 heures
  - Faire le changement
  - Remettre le TTL à 24 heures

Questions ?

AUTRES

- **DNS inverse**
- **Comment déléguer un sous-domaine**

Comment gérer le DNS inverse

- si vous avez au moins un /24 de l'espace d'adressage, alors votre fournisseur se chargera de la délégation à votre serveur de noms.
- Ex. Votre bloc réseau est 192.0.2.0/24
- Créer la zone 2.0.192.in-addr.arpa.
- Si vous avez plus qu'un /24 (Ex. Un /22) alors chaque /24 sera une zone séparée
- Si vous avez assez de chance d'avoir /16 alors il sera une zone unique.
  - 172.16.0.0/16 est 16.172.in-addr.arpa

Exemple : 192.0.2.0/24

```
Zone "2.0.292.in-addr.arpa" {
        type master:
        file "m/192.0.2";
        allow-transfer { . . . };
};

/etc/named/m/192.0.2
@IN         SOA             .....
        IN          NS          ns0.example.com
        IN          NS          ns0.otherwork.com.

1       IN          PTR             router-e0.example.com.
2       IN          PTR             ns0.example.com.
3       IN          PTR             mailhost.example.com.
4       IN          PTR             www.example.com.
; etc...
```

Comment fonctionne le DNS inverse?

- Ex. pour 192.0.2.4, l'hôte distant consultera 4.2.0.192.in-addr.arpa. (PTR)
- La requête suit l'arborescence de la délégation comme la normale. Si tout est correct, Il atteint vos serveurs et aura la réponse.
- Les octets sont placés dans l'ordre inversé
  - Poids plus faible en premier.
- Le propriétaire du grand bloc réseau (192/8) peut déléguer le DNS inverse dans de gros morceaux de /16. le propriétaire d'un /16 peut déléguer des /24

Il n'y a rien de spécial au sujet du DNS inverse

- Vous avez toujours besoin du maître et esclaves
- Il ne fonctionnera pas à moins que vous obteniez la délégation du parent
- S'assurer que si vous avez un enregistrement PTR pour une adresse IP, le nom peut se résoudre à la même adresse IP
  - Ceci n'est pas obligatoire, mais certains sites l'utilise comme critère de filtrage.

Que faire si vous avez moins d'un /24 ?

- Le DNS inverse pour le /24 est délégué à votre fournisseur
- **Option 1:** demandez à votre fournisseur d'insérer les ERs PTR pour votre bloc dans la zone du /24.
  - **Problème**: vous devez leur demander chaque fois que vous voulez de faire un changement
- **Option 2** : Suivez la procédure décrite dans le RFC2317
  - Utilisez l'astuce avec le CNAME pour rediriger les requêtes PTR pour vos adresses IP vers vos serveurs de noms.

Ex. Vous possédez 192.0.2.64/29

```
; Dans le fichier de zone du fournisseur 2.0.192.in-addr.arpa
64    IN      CNAME    64.64/29.2.0.192.in-addr.arpa.
65    IN      CNAME    65.64/292.0.192..in-addr.arpa.
66    IN      CNAME    66.64/29.2.0.192.in-addr.arpa.
67    IN      CNAME    67.64/29.2.0.192.in-addr.arpa.
68    IN      CNAME    68.64/29.2.0.192.in-addr.arpa.
69    IN      CNAME    69.64/29.2.0.192.in-addr.arpa.
70    IN      CNAME    70.64/29.2.0.192.in-addr.arpa.
71    IN      CNAME    71.64/29.2.0.192.in-addr.arpa.
64/29    IN      NS     ns0.customer.com.
64/29    IN      NS     ns1.customer.com.

Configuration de la zone "64/29.2.0.192.in-addr.arpa" sur vos serveurs
65    IN      PTR      www.customer.com.
66    IN      PTR      mailhost.customer.com.
; etc
```

Comment déléguer un sous-domaine ?

- En principe simple : juste insérer les ERs NS pour le sous-domaine, pointant vers les serveurs autoritaires pour le sous-domaine
- Si vous faites attention, vous devriez en premier *vérifier* que les serveurs sont autoritaires pour les sous-domaines.
  - En utilisant "dig" sur tous les serveurs
  - Si le sous-domaine est mal géré, ceci n'est pas bon pour l'image du parent

Le fichier de zone pour "example.com"

```
$TTL 1d
@   1h      IN   SOA     ns1.example.net. Brian.nsrc.org. (
                         2004030300 ; Serial
                         8h         ;Refresh
                         1h         ; Retry
                         4w         ; expire
                         1h )       ; Negative
            IN   NS  ns1.example.net.
            IN   NS  ns2.example.net.
            IN   NS  ns1.othernetwork.com.
; Les données de ma propre zone
            IN   MX  10   mailhost.example.net.
www         IN   A   212.74.112.80

; Sous domaine délégué
Subdom      IN   NS  ns1.othernet.net.
Subdom      IN   NS  ns2.othernet.net.
```

Il n' y a aucun problème dans l'exemple précédent, mais.....

- Mais et si "subdom.example.com" est délégué à ns1.subdom.example.com
- Quelqu'un qui est en cours de résolution de www.subdom.example.com doit d'abord résoudre ns1.subdom.example.com
- Mais il ne peut pas résoudre ns1.subdom.example.com sans résoudre en premier subdom.example.com!!!

Dans ce cas vous avez besoin de "glue record"

- L'enregistrement "glue record" est un ER de type A ou AAAA pour les noms de serveur placés hors de leur zones autoritaires
- Exemple :

```
; zone .com
example          NS  ns.example.com.
                 NS      ns.othernet.net.
ns.example.com.     A    192.0.2.1    ; GLUE RECORD
```

Ne pas mettre les « glue record » exceptés en cas de besoin

- Dans l'exemple précédent, "ns.othernet.net" n'est dans le sous-domaine "example.com" . Par conséquent pas de glue record.
  - Les « glue record » non à jour sont de graves sources de problèmes
    - Après une renumérotation de votre serveur de noms dans un autre réseau
- " dig +norec" est votre ami pour le diagnostic.

Exemple où un « glue record » est nécessaire

```
; Les  données de ma propre zone
          IN   MX       10 mailhost.example.net.
www       IN   A        212.74.112.80

; Le sous-domaine délégué
Subdom       IN   NS  ns1.subdom       ; nécessaire
               IN     NS  ns2.othernet.net.       ; non nécessaire
Ns1.subdom  IN   A    192.0.2.4
```

Vérification des « Glue record »

- Dig +norec @a.gtld-servers.net. www.as9105.net. A
- Rechercher les enregistrements A dans la section "Additionnelle" dont le TTL ne dimunue pas.

Exemple dig +norec @a.gtld-servers.net www.psg.com a

DNS : Résumé global

- Base de données distribuée d'enregistrements de ressources (ER)
- Trois rôles : resolver, le cache(récursif) et l'autoritaire
- Le resolver est configuré avec le(s)cache les plus proches(s)
  - Ex. /etc/resolv.conf
- Les caches sont configurés avec la liste des serveurs racine
  - Le type de zone "hint", /var/named/named.ca

DNS : Résumé global (suite)

- Les serveurs de noms racine contiennent les délégations (NS et glue) vers les gTLD, les ccTLD et le .arpa...
- Les TLDs ont les délégations vers les sous-domaines
- Le cache se localise finalement sur un serveur autoritaire contenant le ER que nous recherchons
- Les erreurs dans les délégations ou dans les configurations aboutissent à “pas de réponse” ou à des “réponses contradictoires.”