Faut-il être l'esclave de la racine ?

Stéphane Bortzmeyer
<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 22 mai 2012

https://www.bortzmeyer.org/slaving-the-root.html

Joli titre, non ? Il fait référence à un débat récurrent chez les opérateurs DNS : faut-il qu'un **résolveur** DNS récupère la zone racine et la serve, afin d'être indépendant des serveurs racine et, ainsi, pouvoir survivre à une panne de ceux-ci ?

Le débat a été récemment relancé par un article intéressant de Pingdom <http://royal.pingdom.com/2012/05/07/the-very-uneven-distribution-of-dns-root-servers-on-the-internet/> à propos de la répartition des serveurs racine. Mais il réapparait régulièrement (voir par exemple l'ancien article « "*The Root of the Matter : Hints or Slaves*" <http://www.maths.tcd.ie/~dwmalone/p/rootslave03.pdf> » de David Malone). L'idée est la suivante : la zone racine est très petite (au 2012-05-22, seulement 3 505 enregistrements, sans compter les signatures DNSSEC, 1 498 noms dont 271 TLD), publiquement disponible en FTP <ftp://rs.internic.net/domain/root.zone.gz> ou via un transfert de zone <https://www.bortzmeyer.org/recuperer-zone-dns.html>. Pourquoi ne pas la récupérer sur ses machines, et modifier la configuration de son serveur résolveur pour faire autorité sur cette zone ? Ainsi, la racine pourrait tomber en panne ou bien être attaquée par des méchants (comme c'est arrivé, sérieusement, en 2007 <https://www.bortzmeyer.org/attaque-serveurs-racine.html> ou, virtuellement, en 2012 <https://www.bortzmeyer.org/racine-dns-opblackout.html>) sans que cela m'affecte.

Mais il n'y a pas de consensus chez les experts sur cette question. Est-ce une bonne idée ou pas ? Le principal problème avec ce système est que la plupart des gens qui le déploient n'assurent pas le service après-vente. Cela semble cool, on configure son résolveur, on se vante après sur Facebook ou sur un forum en http://je-suis-un-neuneu/blaireaux.php mais on ne met pas en place de surveillance automatique. Quelques mois après, le système de mise à jour ne marche plus, la résolution DNS échoue, les utilisateurs pleurent et certains crient « la racine du DNS est en panne, les Mayas l'avaient prédit, nous allons tous mourir ». Configurer un résolveur en esclave de la racine n'est pas un problème lorsque c'est le résolveur individuel d'un gourou qui sait ce qu'il fait. Mais c'est plus problématique lorsque Jean-Kevin Boulet le fait en cinq minutes et s'éclipse après, laissant son successeur régler les problèmes. Rappelez-vous, en administration système, installer et configurer un truc, c'est facile. Le maintenir sur le long terme, c'est le vrai défi.

Quelques autres questions à se poser :

— Est-ce vraiment nécessaire ? La racine n'est jamais tombé une seule fois en panne, malgré plusieurs attaques. Elle est gérée par des gens autrement plus compétents que Jean-Kevin Boulet, il existe déjà beaucoup d'instances physiques de la racine <https://www.bortzmeyer.org/combien-serveurs-racines.html> et les probabilités d'une panne sont très faibles.
— Avoir une copie locale de la racine peut aussi améliorer les performances, si tous les serveurs racine existants sont très loin. (Cela peut se tester <https://www.bortzmeyer.org/le-plus-rapide-dns.html>.)
— De toute façon, sur un résolveur typique, la plupart des données de la racine sont rapidement mises en cache et les serveurs racine ne sont plus interrogés que pour les noms inexistants (suite à une faute de frappe, par exemple). On peut dire que les serveurs de la racine répondent surtout <http://dns.icann.org/cgi-bin/dsc-grapher.pl?binsize=60&window=86400&plot=rcode&server=L-root> NXDOMAIN... (Sur les noms demandés, voir l'étude de Verisign <http://www.verisignlabs.com/notes/DITL-2011-TLDS/>.) Ceci dit, ce n'est pas une raison suffisante pour rejeter l'idée de servir la racine localement : il y a aussi les colles (les adresses IP des serveurs), que BIND retourne chercher à la racine lorsque leur durée de vie expire.
— C'est bien joli d'avoir une copie locale de la racine, et d'être ainsi à l'abri des éventuels problèmes de celle-ci. Mais la plupart des utilisateurs voudraient aussi pouvoir continuer la résolution et accéder à des noms sous `.com` ou `.fr`. Ces TLD ne sont pas publiquement accessibles et, de toute façon, sont trop gros pour tenir sur un résolveur typique.

Si vous voulez le faire quand même (je n'ai pas dit que c'était une bonne idée), voici quelques points à garder en tête :

— Les risques dépendent de la méthode de mise à jour utilisée. Si votre résolveur valide avec DNSSEC (une bonne idée), le principal risque est celui que, en cas de non mise à jour, les signatures expirent et soient donc refusées. Actuellement, la durée de vie des signatures à la racine est de seulement huit jours... Si votre résolveur ne valide pas, et qu'il récupère la zone racine par un transfert de zones depuis les serveurs racine, le principal risque est celui d'expiration de la zone : la durée indiquée dans l'enregistrement SOA de la racine est actuellement de sept jours. Enfin, si vous récupérez la racine en dehors du DNS, par un wget nocturne lancé par cron, il n'y a plus de risque d'expiration. Par contre, si la mise à jour ne se fait plus, vous risquez de servir à vos utilisateurs des données dépassées, un problème qui sera très difficile à déboguer (au moins, avec DNSSEC, le plantage sera franc).
— Si vous choisissez (cela semble la méthode la plus courante) de configurer votre résolveur en esclave de la racine, rappelez-vous que la racine ne sait pas que vous existez et vous ne recevrez donc pas les `NOTIFY` (RFC 1996 [1]). Votre copie sera donc peut-être un peu en retard (le délai d'interrogation dans le SOA de la racine est d'actuellement trente minutes).
— L'ICANN essaie actuellement d'augmenter considérablement le nombre de TLD. Toutefois, même une zone de plusieurs centaines de milliers de noms est largement dans les capacités d'un serveur typique. Pas besoin de grosse machine pour servir la racine localement.
— Il n'existe pas actuellement de serveur maître officiel pour la racine. Certains serveurs permettent le transfert de zones (RFC 5936) comme `F.root-servers.net` mais rien ne garantit que cela durera. L'ICANN a, pour l'instant, des serveurs spéciaux pour ces transferts, `xfr.{cjr|lax}.dns.icann.or`

Une configuration possible pour le résolveur BIND serait :

```
zone "." {
        type slave;
        masters { 192.5.5.241; 2001:500:2f::f; // F
                  193.0.14.129; 2001:7fd::1; // K
                  192.0.47.140; 2620:0:2830:202::140; 192.0.32.140; 2620:0:2d0:202::140; // ICANN
        }; // RFC 4085 says it's bad to hardwire IP addresses.
        file "slave/root.zone";
        notify no;  // Don't bother the offical roots with NOTIFY.
};
```

1. Pour voir le RFC de numéro NNN, `https://www.ietf.org/rfc/rfcNNN.txt`, par exemple `https://www.ietf.org/rfc/rfc1996.txt`

Mais rappelez-vous : il est indispensable de surveiller que tout se passe bien et de réagir vite si la mise à jour ne se fait plus (la racine expirera sur le résolveur en une semaine...) Pour aider à la surveillance, et détecter une expiration proche, vous pouvez essayer le programme dns-slave-expire-checker <http://code.google.com/p/dns-slave-expire-checker/> ou bien attendre BIND 9.10 qui aura un `rndc zonestatus`.

Si vous préférez la solution d'un wget lancé de temps en temps par cron, pensez dans votre script à bien tout tester (ici, on suppose qu'il existe une commande `report_error` qui va envoyer le message d'erreur à un endroit où il sera lu - pas dans un fichier journal que tout le monde ignore) :

```
(wget --quiet ftp://rs.internic.net/domain/root.zone.gz.sig &&
 wget --quiet ftp://rs.internic.net/domain/root.zone.gz) > $LOG 2>&1
if [ $? != 0 ]; then
        report_error "Cannot retrieve root zone file" < $LOG
        exit 1
fi
# Now that the root is DNSSEC-signed, checking with PGP is less important
gpg --quiet --verify root.zone.gz.sig > $LOG 2>&1
if [ $? != 0 ]; then
        report_error "[SECURITY] Bad signature of the root zone file" < $LOG
        exit 1
fi
gunzip --to-stdout root.zone.gz > root.zone
named-checkzone . root.zone > $LOG 2>&1
if [ $? != 0 ]; then
        report_error "Format invalid for root zone file" < $LOG
        exit 1
fi
```

Pour ceux qui aiment se plonger dans l'histoire des débats, le plus gros débat sur la question avait été provoqué en 2007 par la décision de Doug Barton de mettre un système de zone racine local dans FreeBSD. Notez que l'utilisateur voyait donc le comportement de FreeBSD changer soudainement sans qu'il l'ait explicitement accepté, ce qui a beaucoup fait râler <http://lists.freebsd.org/pipermail/freebsd-stable/2007-August/036470.html>. Le débat public avait été lancé sur la liste dns-operations <https://lists.dns-oarc.net/pipermail/dns-operations/2007-July/001803.html> et s'était conclu par un retour en arrière pour FreeBSD <https://lists.dns-oarc.net/pipermail/dns-operations/2007-August/001891.html>. La configuration qui permet de mettre BIND en esclave de la racine est toujours là <http://www.freebsd.org/cgi/cvsweb.cgi/src/etc/namedb/named.conf> mais en commentaire seulement. Notez que les choses ont évolué depuis (la racine est signée, par exemple...) et donc des gens ont changé d'avis sur ce débat.