Sécurité dans le système de résolution de noms -DNS

Constantin Yamkoudougou

28 décembre 2005

Table des matières

Définitions et Abréviations :
DNS : Domain Name System
TSIG : Transaction Signature
IP : Internet Protocol
CERT : Computer Emergency Response Team
IETF : Internet Engineering Task Force, organisme de standardisation en charge des spécifications pour le fonctionnement d'Internet.
RFC : Request For Comment, document de publication finale émis par l'IETF qui fait office standard pour toute nouvelle spécification.
DNSSEC : DNS Security Extension, extension de sécurité spécifiée par l'IETF pour sécuriser les programmes de gestion de noms.

SRI-NIC Stanford Research Instute Network Information Center

1 Généralités dans les systèmes de résolution de noms

Sur internet, une machine est identifiée par son adresse appelée adresse IP (Internet Protocol). Chaque adresse IP est un nombre de 32 bits selon la version 4 du protocole IP[référence fiche adressageIP-initiation]. A titre d'exemple 11000000101000000000101001111111 est une représentation possible d'une adresse IP. Si cette représentation ne pose aucun problème au niveau machine, on ne peut en dire autant à l'échelle humaine pour sa mémorisation. L'IETF a proposé un moyen mnémotechnique pour la représentation d'une telle adresse en la divisant en quatre parties. Chaque partie est alors séparées des autres par un point. L'adresse précédente peut donc être à nouveau représentée par 11000000.10100000.00001010.01111111. En effectuant une conversion des valeurs binaires en décimales on obtient la valeur 192.160.10.127 beaucoup plus facile à retenir pour une mémoire humaine. Une adresse IP selon ce format comporte au maximum 12 chiffres.

1.1 Petite anecdote de la résolution des noms par carnet d'adresse

Retenir deux ou trois adresses comportant 12 chiffres n'est pas très difficile pour une mémoire humaine. En revanche, se remémorer un plus grand nombre nécessiterait la tenu d'un carnet d'adresses. Une organisation de ce carnet d'adresses peut se faire par exemple en deux colonnes. A ce niveau, une analogie peut être faite avec l'organisation d'un répertoire classique de téléphone mobile. Dans la première colonne on y mettrait un nom plus facile à retenir. La seconde colonne contiendrait les adresses IP associées

aux différents noms . Pour une petite et moyenne entreprise comportant un réseau, la structure du carnet d'adresses des machines pour l'administrateur système et réseau pourrait se représenter comme suit :

appelation machine	adresse
directeur	192.168.10.12
comptable	192.168.10.11
stagiaire	192.168.10.13
administrateur	192.168.10.1

Dans ce carnet d'adresses, chaque ligne représente un enregistrement. Pour connaître par exemple l'adresse du poste du directeur, il suffit de consulter la bonne ligne. Dans les années 1970, lorsque internet se résumait à quelques grands réseaux tels que ARPANET et TYMNET, sa gestion reposait sur un grand carnet d'adresses géré au sein d'un fichier texte nommé HOSTS.TXT. Ce fichier porte toujours le même nom sur les machines Unix. La gestion de ce fichier était dévolue au Stanford Research Instute Network Information Center (SRI-NIC) en Californie (Etats Unis). Le système de gestion de noms se résumait donc à la gestion centralisée d'un unique fichier qui était ensuite mis à disposition des administrateurs par courrier électronique ou par téléchargement.

1.2 Bref rappel de la genèse du DNS

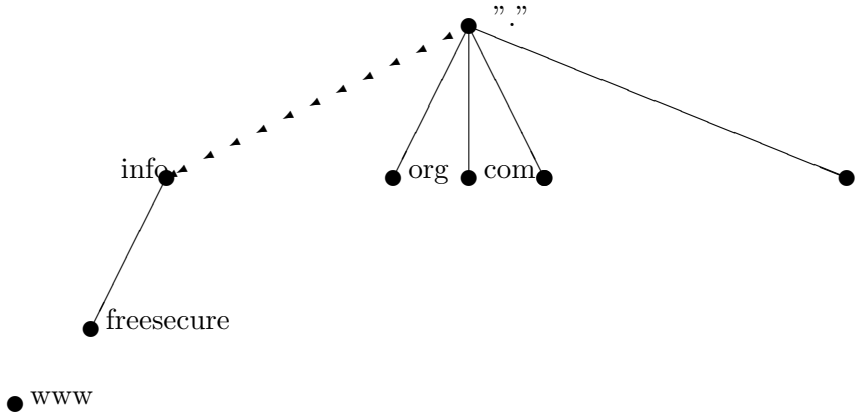
Le maintient de la liste du fichier HOSTS.TXT contenant l'ensemble des ordinateurs devenait très fastidieux compte tenu de l'extraordinaire croissance du réseau Internet. Il était donc impératif de mettre en place un nouveau système de gestion de noms de machines dont les caractéristiques principales étaient les suivantes : éviter les collisions de noms en assurant l'unicité du nommage des machines, assurer un temps de réponse aussi bref que possible à toute sollicitation (temps de réponse à une requête), assurer une gestion décentralisée. Les premières spécifications de ce nouveau système de gestion de noms appelé DNS (Domain Name System) furent écrites à partir de 1983 par Paul MOCKAPETRIS. Ses documents Moc83a et Moc83b fournirent à L'IETF (RFC 882, 883) les premiers canevas de travail pour la spécification du système de nom actuel (RFC 1034, 1035).

Le nouveau système propose une méthode distribuée de gestion des données par opposition à la méthode centralisée du SRI-NIC. Les données peuvent désormais être gérées par chaque administrateur de réseau en utilisant un programme spécialisé appelé serveur de noms. Les données sont organisées sous forme d'enregistrements. Chaque enregistrement peut avoir une typologie particulière qui permet de décrire des alias, des adresses email,

des références à d'autres serveurs de noms, etc. Ce nouveau standard définit également le protocole de communication permettant à un serveur de dialoguer avec d'autres serveurs de noms lorsqu'il ne connait pas l'adresse IP correspondant à un nom donné.

1.3 Organisation du système DNS

Pour permettre la gestion distribuée et l'indépendance de la gestion des serveurs de noms, une stratégie d'organisation sous forme d'arbre inversé a été proposée par Paul MOCKAPETRIS. Un domaine dans cet arbre est l'ensemble des sous arbres et des noeuds rattachés à un noeud donné. Nommé un noeud de cet arbre consiste à suivre le chemin qui mène de ce noeud jusqu'à la racine de l'arbre. A titre d'exemple, pour donner le nom **www** à une machine appartenant au domaine **freesecure** qui est lui même appartient au domaine **info**, il suffira de lui donner le nom du chemin partant du noeud picsi1 à la racine, en mentionnant un point chaque fois qu'un nouveau noeud est franchi dans le parcours. Ainsi, picsi1 sera reconnue sur Internet par le nom pics1.unilim.fr. De cette manière, pour chaque nouveau noeud (machine) connecté à l'arbre, on peut en garantir ainsi l'unicité.



Le système DNS a également introduit la notion de **zone**. Pour des raisons de commodité de gestion, les noeuds ont été organisés en zones. Une zone est l'ensemble des machines appartenant à un même domaine et gérées par un même serveur de noms. On dit que le serveur de noms a autorité sur cette zone. La zone peut coincider avec le domaine s'il n'y a qu'un seul serveur de noms qui gère tous les noeuds d'un domaine donné. D'autres concepts introduisant la notion de hiérarchie entre les serveurs de noms ainsi qu'un arbre permettant de retrouver un nom sur la base à partir d'une adresse IP ont également été introduits. Ces derniers concepts ne sont pas détaillés dans ce document.

1.4 Concept de résolution itérative, résolution récursive

On appelle résolution de nom, le processus qui consiste à trouver une padresse IP pour un nom de noeud donné. Le processus inverse est appelé résolution inverse. Lorsque Dupont veut consulter le site internet de la Redoute, son ordinateur doit d'abord trouver l'adresse IP de ce site avant de lui adresser sa requête. Le programme local sur sa machine qui prend en charge cette opération s'appelle le **résolveur**. Le résolveur s'adresse au serveur de noms qui a autorité sur la zone de Dupont. Deux types de requêtes peuvent alors être exprimées : une requête récursive ou une requête dite itérative.

1.4.1 résolution récursive

Lorsque la requête est récursive, le serveur de noms prend en charge le mécanisme de résolution intégrale en émettant éventuellement lui même des requêtes vers d'autres serveurs de noms. Une réponse du serveur de noms au résolveur dans ce cas a un caractère définitif. Le résolveur de Dupont n'est pas invité à contacter d'autres serveurs de noms.

1.4.2 résolution récursive

Dans le cas d'une requête itérative,, lorsque le serveur de noms ne gère pas lui même l'adresse IP demandée, il retourne au résolveur la meilleure réponse qu'il a du serveur de nom le plus proche qui pourrait résoudre le nom. Le résolveur doit alors à nouveau émettre une requête vers le nouveau serveur de noms qui ainsi été référencé. Le résolveur procède ainsi de proche en proche de manière itérative afin d'arriver à résoudre le nom du site de la Redoute.

1.5 Messages de résolution des noms

Connaître la structure des messages est important pour la compréhension des attaques sur la résolution des noms. Le format des messages échangés entre un résolveur et un serveur de noms peut être résumé brièvement comme suit :

identifiant
Question
Réponse
Autorité
Données additives

identifiant : correspond à un numéro identifiant du message. Ce numéro varie de 0 à 65534.
Question : cette partie du message contient le nom du site de la Redoute que Dupont souhaite résoudre.
Reponse : cette partie est réservée à la réponse du serveur de noms.
Autorité : dans le cas d'une requête itérative, cette partie permet au serveur de noms de transmettre au résolveur, la référence du serveur de noms le plus proche du nom que l'on veut résoudre. D'autres données non indispensables pour la compréhension de la suite de ce document sont ensuite ajoutées à chaque message.

1.6 Comment oscar peut-il usurper l'identité d'un serveur en utilisant le protocole de résolution standard des noms ?

Lorsque le résolveur de Dupont demande la résolution d'un nom au son serveur de nom, il accepte toute réponse en retour à condition qu'il possède le même identifiant que celui de la requête. En resumé, une réponse est fiable si elle possède le même identifiant que la requête. Oscar qui a la possibilité d'écouter le résolveur de Dupont demander la résolution du nom du site de la Redoute peut forger un message et répondre plus rapidement que le serveur de nom de Dupont en lui transmettant une fausse adresse de site qui sera utilisée par son navigateur. Monsieur Dupont pourrait par exemple transmettre son numéro de carte bleue pour une transaction sur un site qu'il croit être celui de la Redoute. On imagine aisément la suite de la mésaventure.

L'usurpation de noms qui vient d'être décrite peut aller plus loin, si le serveur de nom de Dupont ne possède pas le nom et doit lui aussi le résoudre récursivement ou par itération auprès d'autres serveurs. En effet après une résolution de noms, les serveurs de noms stockent en mémoire les adresses des noms qu'ils ont résolus pendant un certain délai. Si Oscar

réussit à transmettre au serveur de noms une mauvaise adresse, il le stockera et le transmettra à tous les internautes qui le lui demanderont pendant un certain laps de temps avant de rafraichir ses données. La duperie peut alors être collective et ne pas concerner uniquement Monsieur Dupont. Une étude du CERT/CC en avril 2002 [ref1] a montré que 80 serveurs sur 100 des TLD (Top Level Domain gérant les domaines ".com", ".edu", ".fr") sont vulnérables à des problèmes de compromission
ref1 : http ://www.cert-ist.com/

1.7 Peut-on sécuriser la résolution des noms ?

Les spécifications de Paul MOCKAPETRIS pour la conception d'un système de gestion de noms n'ont pas tenu compte des problèmes de sécurité qui pouvaient faire défaut. Deux solutions de sécurité ont été proposées pour pallier la vulnérabilité des serveurs de nom précédemment illustrée. La première repose sur l'authentification des messages utilisant des techniques de cryptologie à clé secrète. Il s'agit du protocole TSIG (Transaction Signature). La seconde et plus plausible utilise la cryptologie à clé publique [ref 1].

2 sécurisation des serveurs de noms

2.1 Protocole TSIG

Dans cette première approche de sécurité des serveurs, on utilise un secret partagé entre deux serveurs de noms en communication. Le secret partagé permet de chiffrer une emprunte de chaque message échangé. Les serveurs de noms qui connaissent le secret peuvent vérifier l'authenticité des messages en déchiffrant les empruntes transmises avec les messages. La génération de chaque emprunte se fait à l'aide d'une fonction de hashage md5. Deux réserves peuvent être émises pour le protocole TSIG. Au sujet de md5, des collisions ont été mise en évidence en février 2005 par des chercheurs chinois [ref 1] qui avaient également trouvé des collision sur les algorithmes MD4, MD5, HAVAL-128, et RIPEMD. Utiliser des clé secrêtes nécessiterait une gestion quasiment impossible des clés distribuées. Pour n serveurs de noms sur Internet, il faudra être capable de distribuer en ensemble n^2 clés secrètes pour qu'il puisse dialoguer deux à deux. Si l'on avait un million de serveurs de nom, il est évident que la notion de secret perdrait tout son sens. Le manque de fiabilité de cette première tentative de sécurisation des serveurs de noms par TSIG a conduit à l'élaboration du standard DNSSEC.

ref http ://eprint.iacr.org/2004/199.pdf

2.2 Le protocole DNSSEC

Le projet DNSSEC [ref1] repose sur l'utilisation de la cryptologie à clé publique [ref 1] pour sécuriser les enregistrements et les transactions entre les serveurs de noms. Chaque serveur de noms possède une clé publique et une clé privée. Pour un serveur de nom appartenant à un domaine donné, l'autorité de certification est celui du niveau supérieur. Par exemple, la clé publique du serveur de noms ayant autorité sur le domaine unlim, aura une clé publique signée par l'autorité du domaine fr. Chaque serveurs de nom dans ce nouveau standard signe tous les enregistrements de sa zone à l'aide de sa clé privée. Un enregistrement spécifique est créé pour stocker la clé publique qui est transmise aux autres serveurs de noms pour la vérification des signatures. Les enregistrements peuvent être signés individuellement ou par groupe. On peut par exemple signer l'ensembles des enregistrements d'une zone faisant référence à d'autres serveurs de nom, l'ensemble des enregistrements faisant référence à des serveurs de messagerie électronique, etc.
ref2 http ://www.dnssec.net/drafts

2.3 Les performances du DNSSEC

Le nouveau standard DNSSEC introduit au sein du protocole de communication des serveurs de noms un ralentissement supplémentaire à cause de la temporisation induite par la génération et la vérification des signatures de zone. Il augmente également la taille des enregistrements du fait du stockage des signatures. On estime à environ [ref1] une heure, le temps de signature d'une zone de 300 MB et à 8 ou 9 fois le foisonnement des données à gérer pour un serveur de noms conforme à ce standard par rapport à un serveur de noms de première génération. L'utilisation des technologies à base de courbes elliptiques [ref 2] pourraient améliorer ces performances
ref 1 projet pilote dnssec au pays bas http ://www.nlnetlabs.nl/dnssec/
ref 2 http ://tools.ietf.org/wg/dnsext/draft-ietf-dnsext-ecc-key/draft-ietf-dnsext-ecc-key-07-from-06.diff.html

Références documentaires

DNS and BIND Paul Albitz & Cricket Liu
Adressing weaknesses in the domain name system protocol - thesis submitted for a master of science Purdue University - Chistoph Suba
Paul MOCKAPETRIS, RFC 882 Domain Names - Concepts and Facilities. Network Working group, November 1983
Paul MOCKAPETRIS, RFC 883 Domain Names - Concepts and Facilities. Implementation and Specifications. ,Network Working group, November 1983
Paul MOCKAPETRIS, RFC 1035 Domain Names - Concepts. Implementation and Specifications. ,Network Working group, November 1987
Paul MOCKAPETRIS, RFC 1035 Domain Names - Concepts and Facilities

,Network Working group, November 1987
Eastlake D., RFC 2535, Domain Name System Security Extensions, 1999