Groupe de travail Réseau
Request for Comments : 3258
Catégorie : Information

T. Hardie, Nominum, Inc.
avril 2002
Traduction Claude Brière de L'Isle

Distribution des serveurs de noms d'autorité via des adresses partagées en envoi individuel

Résumé
Le présent mémoire décrit un ensemble de pratiques destinées à permettre à un opérateur de serveur de noms d'autorité de fournir l'accès à un seule serveur désigné dans plusieurs localisations. Le principal motif du développement et déploiement de ces pratiques est d'augmenter la distribution des serveurs du système des noms de domaine (DNS, *Domain Name System*) sur les zones précédemment mal desservies de la topologie de réseau et de réduire la latence des interrogations et réponses du DNS dans ces zones.

Table des matières

1. Introduction

Le présent mémoire décrit un ensemble de pratiques destinées à permettre à un opérateur de serveur de noms d'autorité de fournir l'accès à un seul serveur désigné dans plusieurs localisations. Le principal motif du développement et du déploiement de ces pratiques est d'augmenter la distribution des serveurs DNS sur les zones précédemment mal desservies de la topologie de réseau et de réduire la latence des interrogations/réponses du DNS dans ces zones. Le présent document suppose une transposition biunivoque entre les serveurs d'autorité désignés et les entités administratives (opérateurs). Le présent document ne contient pas de lignes directrices ni recommandations pour les serveurs de noms de mise en antémémoire. Le système d'envoi individuel partagé décrit ici est spécifique de IPv4 ; l'applicabilité à IPv6 est un domaine qui fera l'objet d'une étude ultérieure. On devrait aussi noter que le système décrit ici est relatif à celui décrit dans la [RFC1546], mais il n'exige pas d'espace d'adresses dédié, de changements d'acheminement, ou d'autre éléments d'une pleine infrastructure d'envoi à la cantonade que décrit ce document.

2. Architecture

2.1 Exigences pour le serveur

Les opérateurs de serveurs de noms d'autorité peuvent souhaiter de référer aux [RFC2182] et [RFC2870] pour les directives générales sur les pratiques appropriées pour les serveurs de noms d'autorité. En plus d'une configuration appropriée comme serveur de noms d'autorité standard, chaque hôte participant à un système d'envoi individuel partagé devrait être configuré avec deux interfaces réseau. Ces interfaces peuvent être deux interfaces physiques ou une interface physique transposée sur deux interfaces logiques. Une des interfaces réseau devrait utiliser l'adresse d'envoi individuel IPv4 partagée associée au serveur de noms d'autorité. L'autre interface, appelée ci-dessous une interface administrative, devrait utiliser une adresse IPv4 distincte spécifique de cet hôte. L'hôte devrait répondre aux interrogations DNS seulement sur l'interface d'envoi individuel partagé. Afin de fournir l'ensemble le plus cohérent de réponses à partir du maillage d'hôte d'envoi à la cantonade, il est de bonne pratique de limiter les réponses sur cette interface aux zones pour lesquelles l'hôte est d'autorité.

2.2 Livraison de fichier de zone

Afin de minimiser les risques d'attaques par interposition, les fichiers de zone devraient être livrés à l'interface administrative des serveurs participant au maillage. Des méthodes sûres de transfert de fichier et d'authentification forte devraient être utilisées pour tous les transferts. Si les hôtes dans le maillage rendent leurs zones disponibles pour le transfert de zone, les interfaces administratives devraient aussi être utilisées pour ces transferts, afin d'éviter les problèmes de potentiels changements d'acheminement pour le trafic TCP notés au paragraphe 2.5.

2.3 Synchronisation

Les serveurs de noms d'autorité peuvent être synchronisés de façon lâche ou stricte, selon les pratiques de l'organisation qui opère. Comme noté au paragraphe 4.1.2, le manque de synchronisation entre serveurs utilisant la même adresse d'envoi individuel partagée pourrait créer des problèmes à certains des utilisateurs de ce service. Afin de minimiser ce risque, les basculements d'un ensemble de données à un autre ensemble de données devraient être coordonnés autant que possible. L'utilisation d'horloges synchronisées chez les hôtes participants et des heures fixes pour les basculements fournissent un niveau de base de coordination. Un processus plus complet de coordination impliquerait :
a) un récépissé des zones chez l'hôte de distribution,
b) la confirmation de l'intégrité des zones reçues,
c) la distribution des zones à tous les serveurs dans le maillage,
d) la confirmation de l'intégrité des zones à chaque serveur,
e) la coordination de l'heure de basculement pour les serveurs dans le maillage,
f) l'institution d'un traitement d'échec pour s'assurer que les serveurs qui n'ont pas reçu des données correctes ou pourraient ne pas basculer sur les nouvelles données cessent de répondre aux interrogations entrantes jusqu'à ce que le problème puisse être résolu.

Selon la taille du maillage, l'hôte de distribution peut aussi être un participant ; pour les serveurs d'autorité, ce peut aussi être l'hôte sur lequel les zones sont générées.

Le présent document présume que les méthodes usuelles de récupération sur échec du DNS sont les seules utilisées pour assurer l'accessibilité des données pour les clients. Il ne conseille pas que les chemins soient supprimés dans le cas d'un échec; il conseille plutôt que le processus DNS ferme afin que les serveurs sur d'autres adresses soient interrogés. Cette recommandation reflète un choix entre performances et complexité de fonctionnement. Bien qu'il serait possible d'avoir un processus de suppression de chemin pour une instance spécifique de serveur quand il n'est pas disponible, il y a une complexité opérationnelle considérable qui est impliquée par le fait de s'assurer que cela se fait de façon fiable. Étant données les méthodes existantes de reprise sur échec du DNS, l'amélioration marginale des performances ne serait pas suffisante pour justifier la complexité accrue pour la plupart des utilisations.

2.4 Placement du serveur

Bien que la diversité géographique du placement des serveurs aide à réduire les effets des interruptions de service dues à des problèmes locaux, c'est la diversité du placement dans la topologie de réseau qui est la force conductrice derrière ces pratiques de distribution. Le placement des serveurs devrait souligner cette diversité. Idéalement, les serveurs devraient être placés topologiquement près des points auxquels l'opérateur échange des routes et du trafic avec les autres réseaux.

2.5 Acheminement

L'organisation qui administre le maillage des serveurs partageant une adresse d'envoi individuel doit avoir un numéro de système autonome et parler BGP à ses homologues. À ces homologues, l'organisation annonce un chemin pour le réseau qui contient l'adresse d'envoi individuel partagée du serveur de noms. Les routeurs frontière de l'organisation doivent alors livrer le trafic destiné au serveur de noms à la plus proche instanciation. L'acheminement aux interfaces administratives pour les serveurs peut utiliser les méthodes normales d'acheminement pour l'organisation qui administre.

Un problème potentiel de l'utilisation des adresses d'envoi individuel partagées est que les routeurs qui leur transmettent le trafic peuvent avoir plus d'un chemin disponible, et ces chemins peuvent, en fait, atteindre différentes instances de l'adresse d'envoi individuel partagée. Des applications comme le DNS, dont les communications consistent normalement en messages indépendants de demandes-réponses dont chacun tient dans un seul paquet UDP ne présente pas de problème. D'autres applications, dans lesquelles plusieurs paquets doivent atteindre le même point d'extrémité (par exemple, TCP) peuvent échouer ou présenter des caractéristiques de performances inacceptables dans certaines circonstances. Des échecs de destination partagée peuvent se produire quand un routeur fait du partage de charge par paquet (ou du round-robin) qu'un changement de topologie change les métriques relatives des deux chemins pour la même destination d'envoi à la cantonade, etc.

Quatre choses atténuent la sévérité de ce problème. La première est que UDP représente une très grande proportion du trafic d'interrogations aux serveurs de noms. La seconde est que le but de cette proposition est de diversifier le placement topologique ; pour la plupart des utilisateurs, cela signifie que la coordination du placement va assurer que de nouvelles instances de serveur de noms vont être à une métrique de coût significativement différente des instances existantes. Un certain ensemble d'utilisateurs peut finir au milieu, mais cela devrait être relativement rare. La troisième est que le partage de charge par paquet est seulement un des mécanismes possibles de partage de charge, et d'autres mécanismes sont d'une popularité croissante.

Enfin, dans le cas où le trafic est TCP, le partage de charge par paquet est utilisé, et des chemins de coût égal pour les différentes instances de serveur de noms sont disponibles, toute mise en œuvre du DNS qui mesure les performances des serveurs pour choisir un serveur préféré va rapidement préférer un serveur pour lequel ce problème ne se pose pas. Cependant, pour que les mécanismes de reprise sur défaillance du DNS évitent de façon fiable ce problème, ceux qui utilisent le mécanisme de distribution en envoi individuel partagé doivent veiller à ce que tous les serveurs pour une zone spécifique ne soient pas des participants au même maillage d'envoi individuel partagé. Pour se garder même contre le cas où plusieurs maillages ont un ensemble d'usagers affectés par le partage de charge par paquet le long de chemin de coût égal, les organisations qui mettent en œuvre ces pratiques devraient toujours fournir au moins un serveur d'autorité qui n'est pas un participant à un maillage d'envoi individuel partagé. Ceux qui déploient des maillages d'envoi individuel partagé devraient noter que tout hôte spécifique peut devenir inaccessible à un client si un serveur a une défaillance, si un chemin a une défaillance, ou si le chemin pour un hôte est supprimé. Ces conditions d'erreur ne sont cependant pas spécifiques des distributions en envoi individuel partagé, mais vont se produire pour les hôtes d'envoi individuel standard.

Comme les paquets de réponse ICMP peuvent aller à un membre du maillage différent de celui qui envoie un paquet, les paquets envoyés avec une adresse de source d'envoi individuel partagé devraient aussi éviter d'utiliser la découverte de la MTU de chemin.

L'Appendice A. contient un diagramme ASCII d'un exemple de mise en œuvre simple de ce système. Dans celui ci, les routeurs de numéro impair livrent le trafic à l'interface réseau d'envoi individuel partagé et filtrent le trafic provenant du réseau administratif ; les routeurs de numéro pair livrent le trafic au réseau administratif et filtrent le trafic provenant du réseau en envoi individuel partagé. Ils sont décrits comme des routeurs séparés pour la facilité de l'explication, mais ils pourraient aisément être des interfaces séparées sur le même routeur. De même, une source locale NTP est décrite pour la synchronisation, mais le niveau de synchronisation nécessaire n'exige pas que cette source soit locale ou un serveur NTP de strate un.

3. Administration

3.1 Points de contact

Un seul point de contact pour rapporter les problèmes est crucial pour l'administration correcte de ce système. Si un utilisateur externe du système a besoin de rapporter un problème relatif au service, il doit n'y avoir aucune ambiguïté sur qui contacter. Si la surveillance interne n'indique pas un problème, le contact peut, bien sûr, devoir travailler avec l'utilisateur externe pour identifier quel serveur a généré l'erreur.

4. Considérations de sécurité

Comme pièce centrale de l'infrastructure Internet, les serveurs de noms d'autorité sont des cibles courantes d'attaques. Les pratiques mentionnées ici augmentent le risque de certaines sortes d'attaques et réduisent le risque d'autres.

4.1 Risques accrus

4.1.1 Augmentation des serveurs physiques

L'architecture mentionnée dans le présent document augmente le nombre de serveurs physiques, ce qui pourrait augmenter la possibilité que se produise une mauvaise configuration d'un serveur et permettre une rupture de la sécurité. En général, l'entité qui administre un maillage devrait s'assurer que les correctifs et les mécanismes de sécurité appliqués à un seul membre du maillage sont appropriés pour et appliqués à tous les membres du maillage. La "diversité générique" (code provenant de bases de code différentes) peut être une mesure de sécurité utile pour éviter des attaques fondées sur des vulnérabilité d'une base de code spécifique ; afin d'assurer la cohérence des réponses provenant d'un seul serveur désigné, cependant, cette diversité devrait être appliquée aux différents maillages d'envoi individuel partagé ou entre un maillage et un serveur d'autorité relatif à l'envoi individuel.

4.1.2 Problèmes de synchronisation des données

Le niveau de synchronisation systémique décrit ci-dessus devrait être augmenté par la synchronisation des données présentes à chaque serveur. Bien que le DNS lui-même soit un système à couplage lâche, les problèmes de débogage des données dans des zones spécifiques seraient bien plus difficiles si deux serveurs différents partageant une seule adresse d'envoi individuel pouvaient retourner des réponses différentes à la même interrogation. Par exemple, si les données associées à www.exemple.com ont changé et si les administrateurs du domaine sont en train de vérifier les changements aux serveurs de noms d'autorité exemple.com, ils ne devraient pas avoir besoin de vérifier chaque instance d'un serveur d'autorité désigné. L'utilisation de NTP pour fournir un temps synchronisé pour les basculements élimine certains aspects de ce problème, mais des mécanismes pour traiter les défaillances durant le basculement sont nécessaires. En particulier, un serveur qui ne peut pas faire le basculement ne doit pas revenir à une version antérieure ; il doit cesser de répondre aux interrogations afin que les autres serveurs soient interrogés.

4.1.3 Risques de distribution

Si le mécanisme utilisé pour distribuer les fichiers de zone parmi les serveurs n'est pas bien sécurisé, une attaque par interposition pourrait résulter en l'injection de fausses informations. Les signatures numériques vont alléger ce risque, mais le chiffrement du transport et des listes d'accès strictes y sont un ajout nécessaire. Comme les fichiers de zone vont être distribués aux interfaces administratives des serveurs maillés, la liste de contrôle d'accès pour la distribution des fichiers de zone devrait inclure l'interface administrative du ou des serveurs, plutôt que leurs adresses d'envoi individuel partagées.

4.2 Risques diminués

L'augmentation du nombre de serveurs physiques réduit la probabilité qu'une attaque de déni de service puisse neutraliser une portion significative de l'infrastructure du DNS. L'augmentation des serveurs réduit aussi les effets des défaillances de machines, des coupures de fibres, et de désastres localisés en réduisant le nombre d'utilisateurs dépendants d'une machine spécifique.

5. Remerciements

Masataka Ohta, Bill Manning, Randy Bush, Chris Yarnell, Ray Plzak, Mark Andrews, Robert Elz, Geoff Huston, Bill Norton, Akira Kato, Suzanne Woolf, Bernard Aboba, Casey Ajalat, et Gunnar Lindberg ont tous fourni des apports et commentaires sur le présent travail. l'éditeur souhaite rappeler en particulier la contribution du regretté Scott Tucker, dont la grande expérience des systèmes et le parfait bon sens ont tous deux largement contribué à la propre expérience de déploiement de l'éditeur et il manque à tous ceux qui l'ont connu.

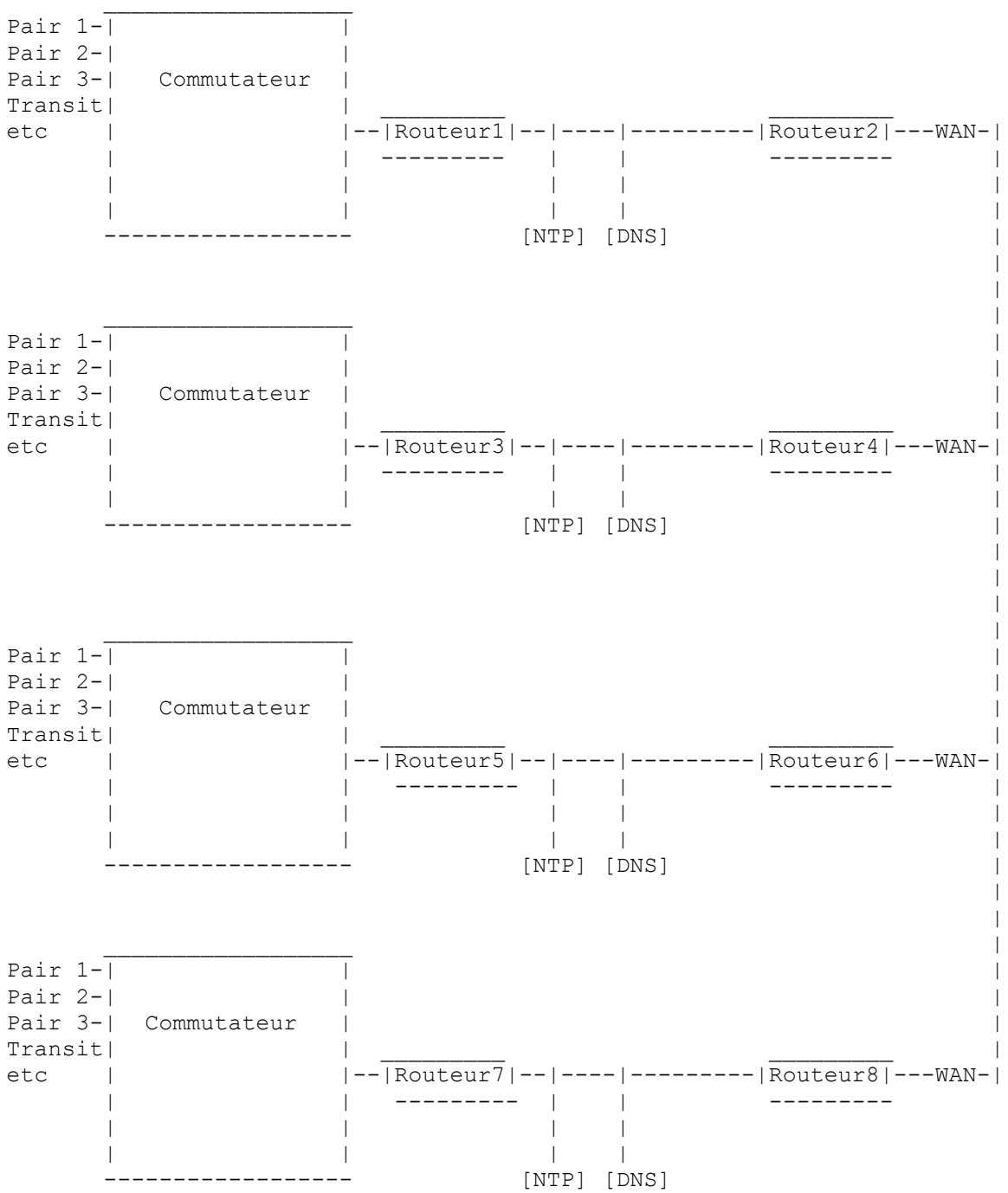
6. Références

[RFC1546] C. Partridge, T. Mendez, W. Milliken, "Service d'envoi à la cantonade pour les hôtes", novembre 1993. *(Information)*

[RFC2182] R. Elz et autres, "Sélection et fonctionnement des serveurs secondaires du DNS", juillet 1997. (BCP0016)

[RFC2870] R. Bush et autres, "Exigences pour le fonctionnement de serveur de nom racine", juin 2000. *(Remplacé par RFC7720)*

Appendice A.



8. Adresse de l'éditeur

Ted Hardie
Nominum, Inc.
2385 Bay Road.
Redwood City, CA 94063
USA
téléphone : 1.650.381.6226
mél : Ted.Hardie@nominum.com

8. Déclaration complète de droits de reproduction

Remerciement
Le financement de la fonction d'édition des RFC est actuellement fourni par l'Internet Society.