

Groupe de travail Réseau
Request for Comments : 947

Ken Lebowitz
David Mankins
BBN Laboratories
juin 1985

Traduction Claude Brière de L'Isle

Diffusion multi-réseau au sein de l'Internet

1. Statut de ce mémoire

La présente RFC décrit l'extension du domaine de diffusion d'un réseau pour inclure plus d'un réseau physique au moyen d'un répéteur de diffusion de paquet.

L'article qui suit présente le problème de la diffusion multi-réseau et les motivations de la résolution de ce problème qui est dans le contexte du développement d'un système d'exploitation réparti. On expose les différentes solutions pour étendre un domaine de diffusion et les raisons du choix de celui qui a été mis en œuvre. De plus, il y a des informations sur la mise en œuvre elle-même et des notes sur ses performances.

On espère que les idées présentées ici aideront les gens de l'Internet qui ont des applications qui utilisent la diffusion et se heurtent à la limitation de n'être capable de diffuser que dans un seul réseau.

Les informations présentées ici sont précises à la date de la présente publication mais des détails spécifiques, en particulier ceux qui concernent la mise en œuvre, pourraient changer à l'avenir. La distribution du présent mémoire n'est soumise à aucune restriction.

2. Le problème

La communication entre des hôtes qui sont sur des réseaux séparés a été largement traitée par l'utilisation des protocoles et des routeurs de l'Internet. Un aspect de la communication inter-réseaux qui n'a pas été résolu dans l'Internet est celui de l'extension de la diffusion pour englober deux réseaux ou plus. La diffusion est un moyen efficace pour envoyer des informations à de nombreux hôtes tout en ayant à transmettre un seul paquet. Beaucoup des architectures actuelles de réseau de zone locale (LAN, *local area network*) prennent directement en charge un mécanisme de diffusion. Malheureusement, ce mécanisme de diffusion présente des défauts lorsque il est utilisé dans des environnements de réseautage qui comportent plusieurs LAN connectés par des routeurs tels que dans le DARPA Internet. Ce défaut est que les paquets diffusés ne sont reçus que par les hôtes qui sont sur le réseau physique sur lequel le paquet est diffusé. Il en résulte que toute application qui tire parti de la diffusion de LAN ne peut diffuser qu'aux hôtes qui sont sur son réseau physique.

On s'appuie sur le développement du système d'exploitation réparti Cronus [1] pour la diffusion. Cronus fournit des services et communications à des processus répartis entre différents types de systèmes informatiques. Cronus est construit autour de grappes logiques d'hôtes connectés à un ou plusieurs LAN à haut débit. La communication dans Cronus est construite sur les protocoles TCP et UDP. Cronus utilise la diffusion pour la localisation dynamique des ressources sur d'autres hôtes et pour collecter les informations d'état à partir d'une collection de serveurs. Comme les capacités de diffusion de Cronus ne sont pas destinées à se limiter aux frontières d'un seul LAN, on avait besoin de trouver un moyen d'étendre notre domaine de diffusion de façon à inclure les hôtes de LAN distants afin d'expérimenter avec des grappes qui s'étendent sur plusieurs réseaux physiques. Cronus utilise de façon prédominante la diffusion pour communiquer avec un sous-ensemble des hôtes qui reçoivent effectivement le message diffusé. Un mécanisme de diffusion groupée serait plus approprié, mais était indisponible dans certaines de nos mises en œuvre de réseau, de sorte qu'on a choisi la diffusion pour la mise en œuvre initiale des utilitaires Cronus.

3. Notre solution

La technique qu'on a choisi d'expérimenter avec le problème de la diffusion multi-réseau peut être décrite comme un "répéteur de diffusion". Un répéteur de diffusion est un mécanisme qui relaie de façon transparente les paquets de diffusion d'un LAN à un autre, et peut aussi transmettre les paquets en diffusion aux hôtes qui sont sur un réseau qui ne prend pas en charge la diffusion au niveau liaison. Ce mécanisme donne de la souplesse tout en tirant quand même parti de la diffusion de LAN.

Notre répéteur de diffusion est un processus sur un hôte réseau qui écoute les paquets en diffusion. Ces paquets sont pris et retransmis, en utilisant un simple protocole de répéteur à répéteur, à un ou plusieurs répéteurs qui sont connectés aux LAN distants. Le répéteur sur l'extrémité de réception va rediffuser le paquet sur son LAN, conservant l'adresse de source du paquet original. Le répéteur de diffusion peut être rendu très intelligent dans son choix de messages à transmettre. Le répéteur ne transmet actuellement que les messages en diffusion envoyés grâce aux accès UDP utilisés par Cronus, mais les messages peuvent être choisis en utilisant tout champ des en-têtes UDP ou IP, ou bien tous les messages de diffusion de niveau IP peuvent être transmis.

4. Solutions de remplacement au répéteur de diffusion

On a exploré plusieurs solutions avant de décider de la technique de transmission des messages en diffusion. Une de ces méthodes était de mettre des fonctions supplémentaires dans les routeurs Internet. Les routeurs pourraient écouter au niveau liaison les paquets en diffusion et les relayer à un ou plusieurs routeurs sur les LAN distants. Ces routeurs pourraient alors transmettre le même paquet sur leurs réseaux en utilisant la capacité de diffusion de niveau liaison du réseau local, si il en est une disponible. Tous les routeurs qui participent à ce schéma vont devoir tenir des tableaux de tous les autres routeurs qui doivent recevoir des diffusions. Si le routeur receveur dessert un réseau sans capacité de diffusion, il pourrait transmettre les messages directement à un ou plusieurs hôtes désignés sur son réseau, mais là encore, cela exigerait que des tableaux soient tenus au routeur. Mettre cette sorte de fonction dans les routeurs a été rejeté pour un certain nombre de raisons : (a) cela exigerait des extensions au protocole de contrôle de routeur pour permettre de mettre à jour les listes que les routeurs devraient tenir, (b) comme tous les messages (par exemple, les messages de résolution d'adresse de LAN) n'ont pas besoin d'être transmis, le besoin de contrôler la transmission devrait être sous la responsabilité de protocoles de niveau supérieur qui pourraient être à la disposition des routeurs, (c) Cronus pourrait être mis dans des environnements où les routeurs peuvent être fournis par d'autres fabricants qui ne mettent pas en œuvre la propagation de la diffusion, (d) au titre du réseau sous-jacent, les routeurs vont vraisemblablement être contrôlés par une agence différente de celle qui contrôle la configuration d'un système Cronus, ajoutant une complexité bureaucratique à la reconfiguration.

Une autre idée qui a été rejetée était de mettre la fonctionnalité de diffusion dans le noyau Cronus. Le noyau Cronus est un processus qui fonctionne sur chaque hôte participant à Cronus, et sa tâche est d'acheminer tous les messages passés entre les processus Cronus. Le noyau Cronus est le seul programme dans le système Cronus qui utilise directement la capacité de diffusion (les autres parties de Cronus communiquent en utilisant des mécanismes fournis par le noyau). On pourrait aussi bien retirer entièrement la dépendance de la diffusion du noyau Cronus, ou ajouter un mécanisme qui émule la diffusion en utilisant des messages transmis en série lorsque le réseau sous-jacent ne fournit pas lui-même une facilité de diffusion. L'une ou l'autre solution exige que tous les processus de noyau Cronus connaissent les adresses de tous les autres participants d'un système Cronus, ce qu'on considère être une limite indésirable à la souplesse de configuration. De plus, cette solution serait spécifique de Cronus, tandis que la solution du répéteur de diffusion est applicable aux autres protocoles fondés sur la diffusion.

5. Mise en œuvre

Le répéteur de diffusion est mis en œuvre comme deux processus séparés – le transmetteur et le répéteur. Le processus transmetteur attend que des paquets UDP en diffusion viennent à travers son réseau local et correspondent à un ou plusieurs numéros d'accès spécifiques (ou adresses de destination). Lorsque arrive un tel paquet, il est encapsulé dans un message de transmetteur-répéteur envoyé à un processus répéteur sur un réseau étranger. Le répéteur relaie alors le paquet transmis sur son LAN en utilisant l'adresse de niveau liaison de ce réseau dans le champ Destination du paquet, mais en préservant l'adresse de source du paquet d'origine.

Lorsque le processus transmetteur démarre pour la première fois, il lit un fichier de configuration. Ce fichier spécifie les adresses des processus répéteurs, et choisit les paquets qui devraient être transmis à chaque processus répéteur (différents répéteurs peuvent choisir des ensembles différents de paquets UDP). Le transmetteur tente d'établir une connexion TCP avec chaque répéteur énuméré dans le fichier de configuration. Si une liaison TCP avec un répéteur a une défaillance, le transmetteur va ressayer périodiquement de s'y connecter. Les hôtes non répéteurs peuvent aussi être énumérés dans le fichier de configuration. Pour ces hôtes, le transmetteur va simplement remplacer l'adresse de destination de diffusion dans le paquet UDP par l'adresse de l'hôte et envoyer ce nouveau datagramme directement à l'hôte non répéteur.

Si un répéteur et un transmetteur coexistent sur le même LAN, il peut survenir un problème si le transmetteur prend des paquets qui ont été rediffusés par le répéteur. Une précaution contre la rediffusion de paquets transmis ("en retour" ou "en boucle") serait que le transmetteur ne se connecte à aucun répéteur énuméré dans son fichier de configuration qui est sur le même réseau que le transmetteur lui-même. Aussi, pour éviter une boucle de diffusion impliquant deux LAN, ayant chacun un transmetteur parlant à un répéteur sur l'autre LAN, les transmetteurs ne transmettent pas les paquets dont l'adresse de source n'est pas sur le LAN du transmetteur.

6. Expérience

Aujourd'hui, le répéteur de diffusion a été mis en œuvre sur le VAX qui fonctionne avec le système d'exploitation UNIX 4.2 BSD avec le logiciel de réseautage de BBN et qui s'est révélé fonctionner assez bien pour notre objet. Notre configuration actuelle comporte deux Ethernets qui sont physiquement séparés par deux autres LAN. Dans les quelques derniers mois, le répéteur de diffusion a réussi à étendre notre domaine de diffusion pour inclure les deux Ethernets même si les messages entre les deux réseaux doivent passer à travers au moins deux routeurs. Nous avons été forcés d'ajouter une capacité spéciale à la mise en œuvre TCP/IP de BBN qui permet à des processus privilégiés d'envoyer des paquets IP avec l'adresse de source d'un autre hôte.

Le répéteur impose une certaine quantité de redondance aux hôtes partagés qui le prennent actuellement en charge à cause de la nécessité de faire fonctionner le processus transmetteur sur tous les paquets UDP qui arrivent à l'hôte, car la décision de rejeter un paquet est prise par un logiciel de niveau utilisateur, plutôt que dans les pilotes de protocole réseau. Une solution à ce problème serait de mettre en œuvre le filtrage de paquets dans le noyau du système (laissant la gestion de configuration et le mécanisme de rediffusion dans le code d'utilisateur) comme il a été fait par Stanford/CMU dans un filtre de paquets UNIX qu'ils ont développé. Comme autre solution, nous prévoyons d'héberger la mise en œuvre de la fonction de répéteur sur un autre hôte comme un service réseau spécialisé fourni par un micro-ordinateur fondé sur un système en temps réel qui fait déjà partie de notre configuration Cronus.

Une telle machine convient mieux pour cette tâche car la redondance de programmation est bien moindre pour eux qu'elle n'est pour un système multi-utilisateur en temps partagé.

7. Référence

- [1] R. Schantz, R. Thomas, R. Gurwitz, G. Bono, M. Dean, K. Lebowitz, K. Schroder, M. Barrow and R. Sands, "Cronus, A Distributed Operating System: Phase 1 Final Report", Rapport technique n° 5885, Bolt Beranek and Newman, Inc., janvier 1985. Le projet Cronus est soutenu par le Rome Air Development Center.

8. Note des éditeurs

Voir aussi sur ce sujet les RFC 919 et 940.