Groupe de travail Réseau G. Choudhury, éd., AT&T
Request for Comments : 4222
BCP 112
Traduction Claude Brière de L'Isle octobre 2005

Traitement par priorité des paquets spécifiques de OSPF version 2 et évitement d'encombrement

Résumé
Le présent document recommande des méthodes qui sont destinées à améliorer l'adaptabilité et la stabilité des grands réseaux qui utilisent le protocole de plus court chemin ouvert en premier (OSPF, *Open Shortest Path First*) version 2. Les méthodes incluent de traiter les Hellos OSPF et les accusés de réception d'annonce d'état de liaison (LSA, *Link State Advertisement*) à une priorité supérieure par rapport aux autres paquets OSPF, et d'autres procédures d'évitement d'encombrement.

Table des matières

1. Introduction

Dans le présent document, OSPF se réfère à OSPFv2 [Ref1]. Les techniques d'amélioration de l'adaptabilité et de la stabilité décrites ici peuvent aussi s'appliquer à OSPFv3 [Ref2], mais elles exigeront des études et des expériences de fonctionnement complémentaires.

Un grand réseau qui fonctionne avec le protocole OSPF peut rencontrer à l'occasion la mise à jour simultanée ou presque simultanée d'un grand nombre d'annonces d'état de liaison, ou LSA. C'est particulièrement vrai si l'extension d'ingénierie du trafic OSPF [Ref3] est utilisée car elle peut augmenter significativement le nombre de LSA dans le réseau. On appelle cet événement une tempête de LSA et elle peut être initiée par une défaillance non programmée ou par un événement de maintenance programmé. La défaillance peut être du matériel, du logiciel, ou de nature procédurale.

La tempête de LSA cause une forte utilisation de CPU et de mémoire dans le routeur causant des retards ou des abandons de paquets entrants. Des accusés de réception retardés (au delà de la valeur du temporisateur de retransmission) résultent en des retransmissions, et des paquets Hello retardés (au delà de l'intervalle de mort du routeur) résultent en la déclaration de mort d'adjacences de voisinage. Les retransmissions et les générations de LSA supplémentaires résultent en plus d'usage de CPU et de mémoire, causant essentiellement une boucle de rétroaction positive, qui, dans les cas extrêmes, peut conduire le réseau à un état instable.

La valeur par défaut du temporisateur de retransmission est de 5 secondes et celle de l'intervalle de routeur mort est de 40 secondes. Cependant, on s'est récemment intéressé à une réduction significative du temps de convergence d'OSPF. Au titre de ce plan, des intervalles beaucoup plus courts (en dessous de la seconde) pour le Hello et le routeur mort ont été proposés [Ref4]. Dans ce scénario, il serait plus probable que les paquets Hello soient retardés au delà de l'intervalle de routeur mort durant l'encombrement de réseau causé par une tempête de LSA.

Afin d'améliorer l'adaptabilité et la stabilité des réseaux, on recommande des mesures pour attribuer des priorités aux paquets OSPF critiques et éviter l'encombrement. Les détails des recommandations figurent à la Section 2. Une étude de simulation est rapportée dans [Ref13] qui quantifie le phénomène d'encombrement et son impact. Elle étudie aussi plusieurs des recommandations et montre qu'elles améliorent bien l'adaptabilité et la stabilité des réseaux qui utilisent le protocole OSPF. La [Ref13] est disponible sur demande en contactant l'éditeur ou un des auteurs.

L'Appendice A explique en plus grand détail les scénarios de tempête de LSA, leur impact, et insiste que quelques exemples réels de tempêtes de messages de contrôle. L'Appendice B donne la liste des variables utilisées dans les recommandations et les valeurs à titre d'exemple. L'Appendice C donne quelques autres recommandations et suggestions avec des objectifs similaires.

2. Recommandations

Les recommandations qui suivent sont destinées à améliorer l'adaptabilité et la stabilité de grands réseaux qui utilisent le protocole OSPF. Durant des périodes d'encombrement du réseau, elles vont réduire les retransmissions, éviter qu'une adjacence soit déclarée morte parce que les paquets Hello ont été retardés au delà de l'intervalle de routeur mort *(RouterDeadInterval)* et prendre d'autres mesures d'évitement d'encombrement. Les recommandations n'ont pas d'ordre sauf que la Recommandation 2 n'est à mettre en œuvre que si la Recommandation 1 n'est pas mise en œuvre.

(1) Classer tous les paquets OSPF en deux groupes : une classe "haute priorité" comprenant les paquets OSPF Hello et les paquets d'accusé de réception d'état de liaison, et une classe de "faible priorité" comprenant tous les autres paquets. La classification est réalisée en examinant l'en-tête de paquet OSPF. Quand on reçoit un paquet d'un voisin ou quand on transmet un paquet à un voisin, on essaye de traiter un paquet "haute priorité" avant un paquet "faible priorité".

Le traitement priorisé lors de la transmission peut être cause que des paquets OSPF provenant d'un voisin vont être reçus en désordre. Si l'authentification cryptographique (AuType = 2) est utilisée (comme spécifié dans [Ref1]), les paquets OSPF valides reçus successivement d'un voisin doivent avoir un "numéro de séquence cryptographique" non décroissant. Pour se conformer à cette exigence, on recommande qu'au cas où l'authentification cryptographique (AuType = 2) est utilisée [Ref1], le traitement priorisé ne soit pas fait à l'émetteur. Cela évitera que les paquets arrivent hors séquence au receveur. Cependant, après le traitement de sécurité au receveur (incluant la vérification de numéro de séquence) est achevé, les paquets OSPF peuvent être gardés dans une file d'attente de "haute priorité" ou de "faible priorité" sur la base de leur classe et être traités en conséquence. Le bénéfice du traitement priorisé est clairement supérieur en l'absence de l'authentification cryptographique car dans ce cas, la priorisation peut être mise en œuvre chez l'émetteur et chez le receveur. Cependant, même avec l'authentification cryptographique, il sera bénéfique de n'avoir la priorisation que chez le receveur (à la suite du traitement de sécurité).

(2) Si la recommandation 1 ne peut pas être mise en œuvre, on remet à zéro le temporisateur d'inactivité pour une adjacence chaque fois qu'un paquet OSPF en envoi individuel ou un paquet OSPF envoyé à tous les routeurs OSPF *(AllSPFRouters)* sur une liaison point à point est reçu sur cette adjacence au lieu de ne remettre à zéro le temporisateur d'inactivité qu'à réception du paquet Hello. Ainsi OSPF ne déclarerait l'adjacence morte que si aucun paquet OSPF en envoi individuel ou aucun paquet OSPF envoyé à AllSPFRouters sur une liaison point à point n'est reçue sur cette adjacence pour une période égalant ou excédant l'intervalle de routeur mort *(RouterDeadInterval)*. La raison pour ne pas recommander cette proposition en conjonction avec la recommandation 1 est d'éviter de potentiels effets collatéraux indésirables. Un de ces effets est le délai pour découvrir l'état de mort d'une adjacence dans le cas où aucun paquet Hello de haute priorité n'a été reçu mais où le temporisateur d'inactivité est remis à zéro par d'autres paquets périmés dans la file d'attente de faible priorité.

(3) Utiliser un algorithme de retard exponentiel pour déterminer la valeur de l'intervalle de retransmission de LSA (RxmtInterval). Soit R(i) qui représente la valeur de RxmtInterval utilisée durant la ième retransmission d'un LSA. Utiliser l'algorithme suivant pour calculer R(i).

R(1) = Rmin
R(i+1) = Min(KR(i),Rmax) pour i ≤ 1

où K, Rmin, et Rmax sont des constantes et la fonction Min(.,.) représente la valeur minimum de ses deux arguments. Des exemples de valeurs pour K, Rmin, et Rmax peuvent être respectivement 2, 5, et 40 secondes. Noter que l'exemple de valeur pour Rmin, l'intervalle initial de retransmission, est le même que la valeur d'échantillon de RxmtInterval dans [Ref1].

Cette recommandation est motivée par l'observation que durant un événement d'encombrement de réseau causé par des messages de contrôle, une source majeure d'entretien de l'encombrement est la retransmission répétée des LSA. L'utilisation d'un algorithme de retard exponentiel pour les intervalles de retransmission de LSA réduit le taux de retransmission des LSA alors que le réseau subit de l'encombrement (durant lequel il est très probable que vont se produire de multiples retransmissions du même LSA). Ceci aide aussi le réseau à sortir de l'état d'encombrement.

(4) Détection implicite d'encombrement et action fondée sur elle : si il y a un encombrement de messages de contrôle à un routeur, ses voisins ne le savent pas explicitement. Cependant, ils peuvent implicitement le détecter sur la base du nombre de LSA non acquittés sur ce routeur. Si ce nombre excède un certain seuil haut, le taux auquel les LSA sont envoyés à ce routeur devrait être réduit progressivement en utilisant un mécanisme de retard exponentiel mais pas en dessous d'un certain taux minimum. Ultérieurement, si le nombre de LSA non acquittés à ce routeur tombe en dessous d'un certain seuil bas, le taux d'envoi des LSA à ce routeur devrait alors être augmenté progressivement, là encore en utilisant un mécanisme de retard exponentiel, mais pas au dessus d'un certain taux maximum. L'algorithme complet est donné ci-dessous. Noter que cet algorithme est à appliquer indépendamment à chaque voisin et seulement pour les LSA en envoi individuel à un voisin ou les LSA envoyés à AllSPFRouters sur une liaison point à point.

Soit,
U(t) = nombre de LSA non acquittés au voisin à l'instant t.
H = un seuil haut (en nombre de LSA non acquittés).
L = seuil bas (en nombre de LSA non acquittés).
G(t) = durée entre les envois successifs de LSA au voisin à l'instant t.
F = facteur d'augmentation de la durée ci-dessus en cas d'encombrement et de diminution après la fin de l'encombrement.
T = durée minimale qui doit s'écouler avant que le changement de la durée existante soit considéré.
Gmin = valeur minimum admise de la durée.
Gmax = valeur maximum admise de la durée.

L'équation ci-dessous montre comment la durée va être changée après l'écoulement d'un temps T depuis le dernier changement :

$$G(t+T) = \begin{cases} \text{Min}(FG(t), Gmax) & \text{si } U(t+T) > H \\ G(t) & \text{if } H \leq U(t+T) \geq L \\ \text{Max}(G(t)/F, Gmin) & \text{si } U(t+T) < L \end{cases}$$

Min(.,.) et Max(.,.) représentent les valeurs, respectivement minimum et maximum, des deux arguments

Exemple de valeurs pour les divers paramètres de l'algorithme : H = 20, L = 10, F = 2, T = 1 s, Gmin = 20 ms, Gmax = 1 s.

Les recommandations 3 et 4 ralentissent toutes deux les LSA pour les voisins encombrés sur la base d'une détection implicite de l'encombrement, mais elles ont des différences importantes. La recommandation 3 ralentit progressivement les retransmissions successives du même LSA, tandis que la recommandation 4 ralentit progressivement tous les LSA (nouveaux ou retransmissions) à un voisin encombré.

(5) Étranglement des adjacences à activer simultanément : si un routeur essaye d'activer simultanément un grand nombre d'adjacences avec ses voisins, cela peut causer un encombrement sévère dû à la synchronisation de la base de données et aux activités d'arrosage de LSA. Il est recommandé que durant une telle situation pas plus de "n" adjacences ne devraient être activées simultanément. Une fois qu'un sous ensemble des adjacences a été activé, de nouvelles adjacences peuvent l'être ensuite pour autant que le nombre d'adjacences activées simultanément n'excède pas "n". La valeur appropriée de "n" va dépendre de la puissance de traitement du routeur, de la bande passante totale disponible pour le trafic de plan de contrôle, et du délai de propagation. La valeur de "n" devrait être configurable.

En présence d'étranglement, une question importante est l'ordre dans lequel les adjacences vont être formées. On recommande une politique de premier arrivé, premier servi (FCFS, *First Come First Served*) sur la base de l'ordre dans lequel les demandes de formation d'adjacence arrivent. Les demandes peuvent être soit de voisins, soit auto générées. Parmi les demandes auto générées, une liste de priorité peut être utilisée pour décider de l'ordre dans lequel les demandes seront traitées. Cependant, une fois que commence un processus de formation d'adjacence, il n'est pas modifiable sauf pour des circonstances exceptionnelles comme des erreurs ou des fins de temporisation.

Dans certaines des recommandations ci-dessus, on se réfère à des liaisons point à point. Ces références devraient aussi inclure les cas où un réseau de diffusion est à traiter comme une connexion point à point du point de vue de l'acheminement IP [Ref5]

3. Considérations sur la sécurité

Le présent mémoire ne crée aucune nouveau problème de sécurité pour le protocole OSPF.

4. Remerciements

Nous tenons à remercier de leur soutien et de leurs utiles commentaires les présidents du groupe de travail OSPF Rohit Dube, Acee Lindem, et John Moy, les directeurs de la zone Acheminement Alex Zinin et Bill Fenner, et les relecteurs de l'IESG. Nous remercions Vivek Dube, Mitchell Erblich, Mike Fox, Tony Przygienda, et Krishna Rao de leurs commentaires sur les précédentes versions de ce document. Nous remercions aussi Margaret Chiosi, Elie Francis, Jeff Han, Beth Munson, Roshan Rao, Moshe Segal, Mike Wardlow, et Pat Wirth de leurr collaboration et de leurs encouragements à nos efforts d'améliorations de l'adaptabilité des réseaux fondés sur des protocoles d'état de liaison.

5. Références normatives

[Ref1] [RFC2328] J. Moy, "OSPF version 2", STD 54, avril 1998. *(MàJ par la* RFC6549, RFC8042)

[Ref2] [RFC2740] R. Coltun, D. Ferguson, J. Moy, "OSPF pour IPv6", décembre 1999. (*Obsolète, voir* RFC5340) *(P.S.)*

6. Références pour information

[Ref3] [RFC3630] D. Katz, K. Kompella et D. Yeung, "Extensions d'ingénierie de trafic à OSPF version 2", septembre 2003.

[Ref4] C. Alaettinoglu, V. Jacobson and H. Yu, "Towards Millisecond IGP Convergence", Travail en cours.

[Ref5] N. Shen, A. Lindem, J. Yuan, A. Zinin, R. White and S. Previdi, "Point-to-point operation over LAN in link-state routing protocols", Travail en cours.

[Ref6] Pappalardo, D., "AT&T, customers grapple with ATM net outage", Network World, 26 février 2001.

[Ref7] "AT&T announces cause of frame-relay network outage," AT&T Press Release, 22 avril 1998.

[Ref8] Cholewka, K., "MCI Outage Has Domino Effect", Inter@ctive Week, 20 août 1999.

[Ref9] Jander, M., "In Qwest Outage, ATM Takes Some Heat", Light Reading, 6 avril 2001.

[Ref10] A. Zinin and M. Shand, "Flooding Optimizations in Link-State Routing Protocols", Travail en cours.

[Ref11] [RFC4136] P. Pillay-Esnault, "Rafraîchissement et réduction d'arrosage OSPF dans les topologies stables", juillet 2005. *(Info.)*

[Ref12] G. Ash, G. Choudhury, V. Sapozhnikova, M. Sherif, A. Maunder, V. Manral, "Congestion Avoidance & Control for OSPF Networks", Travail en cours.

[Ref13] G. Choudhury, G. Ash, V. Manral, A. Maunder and V. Sapozhnikova, "Prioritized Treatment of Specific OSPF Packets and Congestion Avoidance: Algorithms and Simulations", AT&T Technical Report, août 2003.

[Ref14] [RFC2474] K. Nichols, S. Blake, F. Baker et D. Black, "Définition du champ Services différenciés (DS) dans les en-têtes IPv4 et IPv6", décembre 1998. *(P.S. ; MàJ par* RFC3168, RFC3260, RFC8436)

Appendice A. Tempête de LSA : causes et impact

Une tempête de LSA peut être initiée pour de nombreuses raisons. Voici quelques exemples :
(a) une ou plusieurs défaillances de liaison dues à des coupures de fibre,
(b) une ou plusieurs défaillances de routeur pour des raisons diverses, par exemple, défaillance du logiciel ou divers types de desastres (incluant une panne de courant) dans un complexe de bureaux hébergeant de nombreux routeurs,
(c) fluctuations de liaison/routeur,
(d) exigence d'arrêter et ensuite remettre en service de nombreux routeurs durant une mise à niveau de logiciel/matériel,
(e) presque synchronisation du rafraîchissement périodique de 1800 secondes de LSA d'un sous ensemble de LSA,
(f) rafraîchissement de tous les LSA du système durant un changement de version de logiciel,
(g) injection d'un grand nombre de chemins externes à OSPF due à une erreur de procédure,
(h) changement de l'identifiant de routeur causant un grand nombre de nouvelles générations de LSA (éventuellement aussi des purges de LSA selon la mise en œuvre).

En plus des LSA générés par suite directe de défaillances de chemin/routeur, il peut y avoir aussi d'autres LSA indirects. Un exemple dans les réseaux MPLS est celui des LSA d'ingénierie du trafic [Ref3] générés sur d'autres liaisons par suite de changements significatifs de la bande passante réservée. Ils résultent du réacheminement des chemins de commutation d'étiquette (LSP, *Label Switched Path*) qui sont tombés durant la défaillance de liaison/routeur. La tempête de LSA cause une forte utilisation de CPU et de mémoire au processeur du routeur causant la suppression ou l'abandon de paquets entrants. Les accusés de réceptions retardés (au delà de la valeur du temporisateur de retransmission) résultent en retransmissions, et les paquets Hello retardés (au delà de l'intervalle de routeur mort) résultent en liaisons déclarées mortes. Un événement de circuit mort cause la génération de LSA de routeur par ses routeurs de point d'extrémité. Si les LSA d'ingénierie du trafic sont utilisés pour chaque liaison, ce type de LSA va alors aussi être généré par les routeurs de point d'extrémité et potentiellement aussi ailleurs à cause de changements significatifs de la bande passante réservée sur d'autres liaisons causés par la défaillance et le réacheminement des LSP qui à l'origine utilisaient le circuit défaillant. Finalement, quand la liaison récupère, cela va aussi déclencher des LSA de routeur supplémentaires et des LSA d'ingénierie du trafic.

Les retransmissions et les générations de LSA supplémentaires résultent en plus d'usage de CPU et de mémoire, causant essentiellement une boucle de rétroactions positive. On définit la taille de la tempête de LSA comme le nombre de LSA dans la tempête d'origine, en ne comptant pas les LSA supplémentaires résultant de la boucle de rétroactions décrite ci-dessus. Si la tempête de LSA est trop grosse, la boucle de rétroactions mentionnée plus haut peut être assez grosse pour entretenir indéfiniment une grosse utilisation de CPU et de mémoire sur de nombreux routeurs du réseau, conduisant par là le réseau à un état instable. Dans le passé, des événements de panne de réseau ont été rapportés dans des réseaux IP et ATM utilisant des protocoles d'état de liaison tels que OSPF, système intermédiaire à système intermédiaire (IS-IS), interface de réseau privé à réseau (PNNI, *Private Network-Network Interface*) ou autres variantes propriétaires. Voir par exemple [Ref6-Ref9]. Dans beaucoup de ces exemples, des inondations à grande échelle de LSA ou autres messages de contrôle similaires (soit naturelles, soit déclenchées par une erreur ou procédure inappropriée) ont été partiellement ou entièrement responsables de l'instabilité et de la panne du réseau.

Dans [Ref13], un modèle de simulation est utilisé pour montrer qu'il y a un certain seuil de taille de tempête de LSA au dessus duquel le réseau peut présenter un comportement instable causé par un grand nombre de retransmissions, de défaillances de liaisons dues à des paquets Hello manqués, et des récupérations suivantes de liaison. Il montre aussi que la taille de tempête de LSA qui cause l'instabilité peut être substantiellement augmentée en fournissant un traitement priorisé aux paquets Hello et aux LSA d'accusé de réception et en utilisant un algorithme de retard exponentiel pour déterminer l'intervalle de retransmission de LSA. Si il n'est pas possible de donner une priorité aux paquets Hello, remettre alors à zéro le temporisateur d'inactivité à réception de tout paquet OSPF valide peut aussi assurer le même avantage. De plus, si on donne une priorité aux paquets Hello, même quand le réseau fonctionne un peu au dessus du seuil de stabilité, les liaisons ne sont alors pas déclarées mortes à cause des Hellos manqués. Cela implique que même si il y a encombrement du plan de contrôle à cause de trop de retransmissions, le plan des données reste debout et aucun nouveau LSA n'est généré (en plus de ceux de la tempête originale et des rafraîchissements). Ces observations viennent à l'appui des trois premières recommandations de la Section 2. Les auteurs de ce document ont aussi fait des simulations pour vérifier que les autres recommandations de la Section 2 aident à éviter l'encombrement et permettent une sortie en douceur d'un état d'encombrement.

On peut argumenter que la question de l'adaptabilité des grands réseaux ne devrait être résolue que par une division hiérarchique du réseau en plusieurs zones afin que l'inondation de LSA reste localisée au sein de ces zones. Cependant, cette approche augmente la complexité de la gestion et de la conception du réseau et peut résulter en un acheminement moins optimal entre les zones. Aussi, les LSA de système autonome externe (ASE, *Autonomous System External*) sont arrosés à travers les AS, et cela peut être un problème si il y en a un grand nombre. De plus, un grand nombre de LSA de résumé peuvent devoir être arrosés à travers les zones, et leur nombre va augmenter significativement si plusieurs routeurs de zone bordure sont employés pour assurer la fiabilité. Donc, il est important de permettre que le réseau croisse vers la plus grande taille possible sous une seule zone.

Les recommandations du document entrent en synergie avec les plus larges propositions d'amélioration de l'adaptabilité et de la stabilité. [Ref10] propose une réduction de la redondance d'inondation dans les cas où plus d'une interface va au même voisin. [Ref11] propose un mécanisme pour réduire considérablement les rafraîchissements de LSA dans des topologies stables.

La [Ref12] propose une large gamme de mécanismes de contrôle d'encombrement et de récupération de défaillance (certaines de ces idées sont reprises dans le présent document, mais [Ref12] a d'autres idées qui ne sont pas couvertes ici).

Appendice B. Liste des variables et des valeurs

F = facteur d'accroissement de la durée entre l'envoi des LSA successifs à un voisin durant l'encombrement et de diminution après la fin de l'encombrement (utilisé dans la recommandation 4). L'exemple de valeur est 2.

G(t) = durée entre l'envoi de LSA successifs à un voisin à l'instant t (utilisé dans la recommandation 4).

Gmax = valeur maximum admise de la durée entre l'envoi de LSA successifs à un voisin (utilisé dans la recommandation 4). L'exemple de valeur est 1 seconde.

Gmin = valeur minimum admise de la durée entre l'envoi de LSA successifs à un voisin (utilisé dans la recommandation 4). L'exemple de valeur est 20 ms.

H = seuil haut (l'unité est le nombre de LSA non acquittés). Excéder ce seuil déclenche une augmentation potentielle de la durée entre l'envoi de LSA successifs à un voisin (utilisé dans la recommandation 4). L'exemple de valeur est 20.

K = constante multiplicative utilisée pour augmenter la valeur RxmtInterval utilisée durant les retransmissions successives du même LSA (utilisé dans la recommandation 3). L'exemple de valeur est 2.

L = seuil bas (l'unité est le nombre de LSA non acquittés) Tomber en dessous de ce seuil déclenche une diminution potentielle de la durée entre l'envoi de LSA successifs à un voisin (utilisé dans la recommandation 4). L'exemple de valeur est 10.

n = limite supérieure du nombre d'adjacences à activer simultanément (utilisé dans la recommandation 5).

R(i) = valeur de RxmtInterval utilisée durant la ième retransmission d'un LSA (utilisé dans la recommandation 3).

Rmax = valeur maximum admise de RxmtInterval (utilisé dans la recommandation 3). L'exemple de valeur est 40 secondes.

Rmin = valeur minimum admise de RxmtInterval (utilisée dans la recommandation 3). L'exemple de valeur est 5 secondes.

T = délai minimum qui doit s'écouler avant que la durée existante entre l'envoi de LSA successifs à un voisin soit changée (utilisée dans la recommandation 4). L'exemple de valeur est 1 seconde.

U(t) = nombre de LSA non acquittés pour un voisin à l'instant t (utilisé dans la recommandation 4).

Appendice C. Autres recommandations et suggestions

(1) Marquage explicite : dans la Section 2, on recommande que les paquets OSPF soient classés en priorité "haute" et "faible" sur la base d'un examen de l'en-tête de paquet OSPF. Dans certains cas (en particulier chez le receveur) cet examen peut être coûteux en terme de calcul. Une solution de remplacement serait d'utiliser des réglages différents du champ de TOS/Préséance pour les deux classes de priorité. [Ref1] recommande le réglage du champ TOS à 0 et du champ Préséance à 6 pour tous les paquets OSPF. On recommande ce même réglage pour les paquets de "faible" priorité OSPF et un réglage différent pour les paquets OSPF de "haute" priorité afin d'être capable de les classer séparément sans avoir à examiner l'en-tête de paquet OSPF. On donne deux exemples ci-dessous :

Exemple 1 : pour les paquets de priorité "faible", on règle le champ TOS à 0 et le champ Préséance à 6, et pour les paquets de priorité "haute" on règle le champ TOS à 4 et le champ Préséance à 6.

Exemple 2 : pour les paquets de priorité "faible", on règle le champ TOS à 0 et le champ Préséance à 6, et pour les paquets de priorité "haute" on règle le champ TOS à 4 et le champ Préséance à 7.

Noter que les bits TOS/Préséance ont été redéfinis par Diffserv (RFC 2474, [Ref14]). Noter aussi que les différents réglages du champ TOS/Préséance suggérés ci-dessus ont seulement à faire l'objet d'un accord par les systèmes sur la liaison. Il n'est pas nécessaire de suivre cette recommandation si il est aisé d'examiner l'en-tête de paquet OSPF et donc de classer ainsi séparément les paquets de priorité "haute" et "faible".

(2) Plus de priorités pour les paquets OSPF : à côté des paquets désignés comme de "haute" priorité dans la recommandation 1 de la Section 2, il peut y avoir un besoin de plus séparer les priorités parmi les paquets OSPF de priorité "faible". On recommande l'utilisation de trois classes de priorité : "haute", "moyenne" et "faible". Lors de la réception d'un paquet d'un voisin et lors de la transmission d'un paquet à un voisin, on essaye de traiter un paquet de priorité "haute avant des paquets de priorité "moyenne" et "faible" et un paquet de priorité"moyenne" avant les paquets de priorité "faible". Les paquets de priorité "haute" sont comme les désigne la recommandation 1 de la Section 2. On fournit ci-dessous deux exemples de candidats pour les paquets de priorité "moyenne". Tous les paquets OSPF qui ne sont pas désignés comme de priorité "haute" ou "moyenne" sont de priorité "faible". Si on utilise l'authentification cryptographique (AuType = 2) (comme spécifié dans [Ref1]) le traitement priorisé n'est à fournir que chez le receveur et après le traitement de sécurité, mais pas chez l'émetteur car cela peut causer un déclassement des paquets à l'arrivée et violer les exigences de "Autype = 2".

Un exemple de paquet de priorité "moyenne" est le paquet de description de base de données (DBD, *Database Description*) provenant d'un esclave (durant le processus de synchronisation de la base de données) qui est utilisé comme accusé de réception.

Un second exemple est un LSA portant des informations de changement de topologie intra zone (cela peut déclencher un calcul de SPF et un réacheminement des chemins à commutation d'étiquettes, de sorte qu'un traitement rapide de ce paquet peut améliorer les temps de convergence de OSPF et des protocoles de distribution d'étiquettes (LDP, *Label Distribution Protocol*)). Cependant, si le coût de traitement d'identifier et mettre séparément en file d'attente les LSA de cet exemple est réputé élevé, la mise en œuvre peut décider de ne pas le faire.

(3) Traitement d'un grand nombre de purges de LSA : occasionnellement, certains événements du réseau, comme un changement d'identifiant de routeur, peuvent résulter en un grand nombre de nouvelles générations de LSA et de purges de LSA. Dans ce scénario, on peut considérer de traiter les LSA dans un ordre différent, par exemple, de traiter les purges de LSA avant les générations de LSA. On ne recommande cependant pas le traitement hors séquence des LSA pour plusieurs raisons. D'abord, supprimer le type de LSA avant la mise en file d'attente peut être coûteux du point de vu calcul. Le traitement hors séquence peut aussi causer des erreurs subtiles. On ne veut pas recommander un changement majeur du paradigme de traitement de LSA pour un événement relativement rare comme le changement d'identifiant d'un routeur. Cependant, un routeur avec un identifiant qui change peut purger graduellement les vieux LSA sans causer de tempête.

Adresse des auteurs contributeurs

En plus de l'éditeur, plusieurs personnes ont contribué à ce document. Les noms et informations de contact de tous les auteurs figurent ci-dessous :

Anurag S. Maunder
Erlang TechnologyAT&T
2880 Scott Boulevard
Santa Clara, CA 95052
US
téléphone : (408) 420-7617
mél : anuragm@erlangtech.com

Gerald R. Ash
AT&T
Room D5-2A01
200 Laurel Avenue
Middletown, NJ, 07748
téléphone : (732) 420-4578
mél : gash@att.com

Vishwas Manral
Sinett Corp,
2/1 Embassy Icon Annex,
Infantry Road,
Bangalore 560 001
India
mél : vishwas@sinett.com

Vera D. Sapozhnikova
AT&T
Room C5-2C29
200 Laurel Avenue
Middletown, NJ, 07748
USA
téléphone : (732) 420-2653
mél : sapozhnikova@att.com

Adresse de l'éditeur

Gagan L. Choudhury
AT&T
Room D5-3C21
200 Laurel Avenue
Middletown, NJ, 07748
USA
téléphone : (732) 420-3721
mél : gchoudhury@att.com

Déclaration complète de droits de reproduction

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