Groupe de travail Réseau
Request for Comments : 2814
Catégorie : En cours de normalisation

R. Yavatkar, Intel
D. Hoffman, Teledesic
Y. Bernet, Microsoft
F. Baker, Cisco
M. Speer, Sun Microsystems

Traduction Claude Brière de L'Isle
mai 2000

SBM : Protocole de contrôle d'admission fondé sur RSVP pour réseaux de style IEEE 802

Statut du présent mémoire
Le présent document spécifie un protocole de l'Internet en cours de normalisation pour la communauté de l'Internet, et appelle à des discussions et suggestions pour son amélioration. Prière de se référer à l'édition en cours des "Normes officielles des protocoles de l'Internet" (STD 1) pour connaître l'état de la normalisation et le statut de ce protocole. La distribution du présent mémoire n'est soumise à aucune restriction.

Résumé
Le présent document décrit une méthode et un protocole de signalisation pour le contrôle d'admission fondé sur RSVP sur des LAN de style IEEE 802. Le protocole est conçu pour fonctionner aussi bien avec la génération actuelle de LAN IEEE 802 qu'avec ceux des travaux récemment achevés par le comité IEEE 802.1.

Table des matières

1. Introduction

De nouvelles extensions à l'architecture de l'Internet et aux modèles de service ont été définies pour un Internet à intégration de services [RFC1633], [RFC2205], [RFC2210] afin que les applications puissent demander des qualités ou niveaux de service spécifiques à l'interréseau en plus du service IP actuel au mieux. Ces extensions incluent RSVP, un protocole d'établissement de réservation de ressource, et la définition de nouvelles classes de service à prendre en charge par les routeurs de services intégrés. RSVP et les définitions de classes de service sont largement indépendantes des technologies de réseautage sous-jacentes et il est nécessaire de définir la transposition de RSVP et des spécifications de services intégrés dans les technologies spécifiques des sous-réseaux. Par exemple, une définition des transpositions de service et des protocoles d'établissement de réservations est nécessaire pour des technologies spécifiques de couche réseau comme les technologies de LAN de style IEEE-802 partagés et commutés.

Le présent document définit le gestionnaire de bande passante de sous-réseau (SBM, *Subnet Bandwidth Manager*) un protocole de signalisation pour le contrôle d'admission fondé sur RSVP sur réseaux de style IEEE 802. SBM fournit une méthode pour la transposition d'un protocole d'établissement au niveau internet, comme RSVP, en réseau de style IEEE 802. En particulier, il décrit le fonctionnement d'hôtes/routeurs à capacité RSVP et d'appareils de couche liaison (commutateurs, ponts) pour prendre en charge la réservation de ressources de LAN pour les flux de données à capacité RSVP. Un cadre pour la fourniture de services intégrés sur des technologies de LAN de style IEEE-802 partagés et commutés et une définition des transpositions de service ont été décrits dans des documents distincts [RFC2816], [RFC2815].

2. Objectifs et hypothèses

Le protocole de gestionnaire de bande passante de sous-réseau (SBM, *Subnet Bandwidth Manager*) et son utilisation pour le contrôle d'admission et la gestion de bande passante dans les réseaux IEEE 802 de niveau 2 se fonde sur les hypothèses et objectifs archicturaux suivants :

I. Bien que la tendance actuelle soit à une utilisation accrue des topologies de LAN commutés consistant en nouveaux commutateurs qui prennent en charge les mécanismes de mise en file d'attente avec priorité spécifiés par IEEE 802.1p, on suppose que les technologies de LAN vont continuer d'être un mélange de segments de LAN traditionnels partagés/commutés et de nouveaux segments commutés fondés sur la spécification IEEE 802.1p. Nous spécifions donc un protocole de signalisation pour la gestion de la bande passante sur les deux topologies de LAN traditionnels et plus récents et nous tirons parti de ces fonctionnalités supplémentaires (comme la prise en charge explicite de différentes classes de trafic ou de classes de services intégrés) lorsque elles deviennent disponibles dans la nouvelle génération de commutateurs, concentrateurs, ou ponts. Il en résulte que le protocole SBM devrait permettre toute une gamme de solutions de gestion de bande passante de LAN qui vont de celles qui n'exercent qu'un contrôle purement administratif (sur la quantité de bande passante consommée par les flux de trafic à capacité RSVP) à celles qui exigent la coopération (et l'application) par tous les systèmes d'extrémité ou commutateurs d'un LAN IEEE 802.

II. Le présent document ne spécifie qu'une méthode et un protocole de signalisation pour le contrôle d'admission fondé sur le LAN sur les flux RSVP. Aucun mécanisme de contrôle du trafic n'est défini ici pour la couche liaison ; le protocole est conçu pour utiliser les mécanismes définis par IEEE 802. De plus, on suppose que les systèmes d'extrémité de couche 3 (par exemple, un hôte ou un routeur) vont exercer le contrôle de trafic en régulant les flux de trafic de services intégrés pour assurer que chaque flux reste dans les spécifications de trafic stipulées dans la demande de réservation précédemment soumise pour le contrôle d'admission. Cela permet alors à un système qui utilise le contrôle d'admission SBM combiné avec le formatage par flux chez les systèmes d'extrémité et au contrôle de trafic défini par l'IEEE à la couche liaison de réaliser des approximations des services à charge contrôlée (et même garantie) sur les LAN de style IEEE 802.

III. En l'absence de tout contrôle de trafic de couche liaison ou de mécanisme de mise en file d'attente avec priorité dans le LAN sous-jacent (comme un segment de LAN partagé) le mécanisme de contrôle d'admission fondé sur SBM limite seulement la quantité totale de charge de trafic imposée par les flux à capacité RSVP sur un LAN partagé. Dans un tel environnement n'existe aucun mécanisme de séparation des flux de trafic pour protéger les flux à capacité RSVP contre le trafic au mieux sur le même support partagé et cela soulève la question de l'utilité d'un tel mécanisme en dehors d'une topologie ne comportant que des commutateurs conformes à 802.1p. Cependant, on suppose que le mécanisme de contrôle d'admission fondé sur SBM servira quand même à quelque chose d'utile dans une topologie de LAN partagé traditionnelle pour deux raisons. D'abord, en supposant que tous les nœuds qui génèrent des flux de trafic à services intégrés utilisent la procédure de contrôle d'admission fondée sur SBM pour demander la réservation de ressources avant d'envoyer du trafic, le mécanisme va restreindre la quantité totale de trafic généré par les flux de services intégrés au sein des limites désirées par un administrateur de LAN (voir la discussion du paramètre NonResvSendLimit à l'Appendice C). Ensuite, le trafic au mieux généré par les sources de trafic fondé sur TCP/IP est généralement adaptable au débit (en utilisant un mécanisme d'évitement d'encombrement de "démarrage lent" de style TCP ou un mécanisme d'adaptation du débit fondé sur les rétroactions comme celui des flux audio/vidéo fondés sur les protocoles RTP/RTCP) et s'adapte pour rester dans la bande passante disponible sur le réseau. Donc, la combinaison du contrôle d'admission et de l'adaptation du débit devrait éviter l'encombrement persistent du trafic. Cela ne garantit cependant pas que le trafic qui n'est pas à intégration de services ne va pas interférer avec le trafic à intégration de services en l'absence de prise en charge du contrôle du trafic dans l'infrastructure de LAN sous-jacente.

3. Organisation de la suite du document

Le reste de ce document donne une description détaillée des procédures de contrôle d'admission fondé sur SBM pour les technologies de LAN IEEE 802. Le document est organisé comme suit :

* La Section 4 définit d'abord les divers termes utilisés dans le document puis donne une vue d'ensemble de la procédure de contrôle d'admission avec un exemple de son application à un échantillon de réseau.

* La Section 5 décrit les règles du traitement et de transmission des messages PATH (et PATH_TEAR) aux gestionnaires de bande passante de sous-réseau désignés (DSBM, *Designated Subnet Bandwidth Managers*), SBM, et clients DSBM.

* La Section 6 traite des questions d'inter fonctionnement lorsque un DSBM peut fonctionner en l'absence de signalisation RSVP à la couche 3 ou lorsque un autre protocole de signalisation (comme SNMP) est utilisé pour réserver des ressources sur un segment de LAN.

* L'Appendice A décrit les détails de l'algorithme de choix de DSBM qui est utilisé pour choisir un SBM désigné sur un segment de LAN lorsque plus d'un SBM est présent. Il décrit aussi comment les clients DSBM découvrent la présence d'un DSBM sur un segment géré.

* L'Appendice B spécifie les formats des messages spécifiques de SBM utilisés et les formats des nouveaux objets RSVP nécessaires pour le fonctionnement de SBM.

* L'Appendice C décrit l'usage du DSBM pour distribuer les informations de configuration aux envoyeurs sur un segment géré.

4. Généralités

4.1 Définitions

Couche liaison, ou couche 2 ou L2 : se réfère aux technologies de couche liaison des données telles que IEEE 802.3/Ethernet comme L2 ou couche 2.

Domaine de couche liaison ou domaine de couche 2 ou domaine L2 : ensemble de nœuds et de liaisons interconnectées sans passer à travers une fonction de transmission de L3. Un ou plusieurs sous-réseaux IP peuvent être superposés sur un domaine L2.

Appareil de couche 2 ou L2 : se réfère aux appareils qui ne mettent en œuvre que la fonction de couche 2 comme appareils de couche 2 ou appareils L2. Cela inclut les ponts ou commutateurs 802.1D.

Couche interréseau ou couche 3 ou L3 : couche 3 du modèle en sept couches de l'ISO. Ce document est principalement concerné par les réseaux qui utilisent le protocole Internet (IP) à cette couche.

Appareil de couche 3 ou appareil L3 ou station d'extrémité : cela inclut les hôtes et routeurs qui utilisent des protocoles de couche 3 et supérieures ou des programmes d'application qui ont besoin de faire des réservations de ressource.

Segment : un segment physique de couche qui est partagé par un ou plusieurs envoyeurs. Des exemples de segments incluent (a) un Ethernet partagé ou anneau à jetons inter réseau qui résout la concurrence pour l'accès aux supports en utilisant le CSMA *(accès multiple avec surveillance de signal et détection de collision)* ou le passage de jetons ("segment L2 partagé"), (b) une liaison unidirectionnelle entre deux stations ou commutateurs, (c) une direction d'une liaison bidirectionnelle commutée.

Segment géré : c'est un segment avec un DSBM présent et chargé d'exercer le contrôle d'admission sur les demandes de réservations de ressource. Un segment géré comporte les parties interconnectées d'un LAN partagé qui ne sont pas séparées par des DSBM.

Classe de trafic : agrégation de flux de données à qui sont donnés un service similaire au sein d'un réseau commuté.

Priorité d'utilisateur : c'est une valeur associée à la transmission et la réception de toutes les trames dans le modèle de service IEEE 802 : elle est fournie par l'envoyeur qui utilise le service MAC. Elle est fournie avec les données à un receveur qui utilise le service MAC. Elle peut être portée ou non sur le réseau : l'anneau à jetons /802.5 porte cette valeur (codée dans son octet FC), L'Ethernet/802.3 de base ne le fait pas, 802.12 peut ou non le faire selon le format de trame utilisé. 802.1p définit une façon cohérente de porter cette valeur sur le réseau ponté sur Ethernet, les anneaux à jetons, la priorité de demande, FDDI ou autres supports de couche MAC qui utilisent un format de trame étendu. L'usage de priorité d'utilisateur est complètement décrit au paragraphe 2.5 de 802.1D [IEEE8021D] et de 802.1p [IEEE8021P] "Prise en charge du service de couche interne par des procédures spécifiques de MAC".

Sous-réseau : utilisé dans le présent mémoire pour indiquer un groupe d'appareils de couche 3 qui partagent un préfixe commun d'adresse réseau de couche 3 ainsi que l'ensemble des segments qui constituent le domaine de couche 2 dans lequel ils sont situés.

Pont/Commutateur : un appareil de transmission de couche 2 comme défini par IEEE 802.1D. Les termes pont et commutateur sont utilisés l'un pour l'autre dans ce document.

DSBM : SBM désigné (DSBM) est une entité de protocole qui réside dans un appareil de couche 2 ou 3 et qui gère les ressources sur un segment de couche 2. Il en existe au plus un pour chaque segment L2.

SBM : c'est une entité de protocole qui réside dans un appareil de L2 ou L3 et qui est capable de gérer les ressources sur un segment. Cependant, seul un DSBM gère les ressources pour un segment géré. Lorsqu'il existe plus d'un SBM sur un segment, un des SBM est choisi comme DSBM.

Segment étendu : un segment étendu inclut les parties d'un réseau qui sont membres du même sous-réseau IP et ne sont donc séparées par aucun appareil de couche 3. Plusieurs segments gérés, interconnectés par des appareils de couche 2, constituent un segment étendu.

Domaine L2 géré : un domaine de couche 2 consistant en segments gérés est désigné comme un domaine géré de couche 2 pour le distinguer d'un domaine L2 sans DSBM présent pour exercer le contrôle d'admission sur les ressources des segments dans le domaine L2.

Clients DSBM : ce sont des entités qui transmettent le trafic sur un segment géré et utilisent les services d'un DSBM sur le segment géré pour le contrôle d'admission sur un segment de LAN. Seules les entités de couche 3 ou supérieure sur des appareils de couche 3 tels que les hôtes et les routeurs sont supposés envoyer du trafic qui exige des réservations de ressources, et donc, les clients DSBM sont des entités de couche 3.

Appareils transparents à SBM : un appareil "transparent à SBM" ne connaît pas les SBM ou DSBM (bien qu'il puisse ou non avoir connaissance de RSVP) et donc, ne participe pas à la procédure de contrôle d'admission fondée sur le SBM sur un segment géré. Un tel appareil utilise les règles standard de transmission appropriées pour cet appareil et il est transparent par rapport au SBM. Un exemple d'un tel appareil L2 est un commutateur traditionnel qui ne participe pas à la réservation de ressource.

Adresses de couche 2 et de couche 3: on se réfère aux adresses de couche 3 des appareils L3/L2 comme à des "adresses L3" et aux adresses de couche 2 comme à des "adresses L2". Cette convention sera utilisée dans le reste du document pour distinguer entre les adresses de couche 3 et de couche 2 utilisées pour se référer aux appareils de prochain bond RSVP (NHOP) et du bond précédent (PHOP). Par exemple, dans le traitement conventionnel de message RSVP, l'objet RSVP_HOP dans un message PATH porte l'adresse L3 de l'appareil du bond précédent. On se réfèrera à l'adresse contenue dans l'objet RSVP_HOP sous le nom d'adresse RSVP_HOP_L3 et l'adresse MAC correspondante de l'appareil de bond précédent sera désignée comme adresse RSVP_HOP_L2.

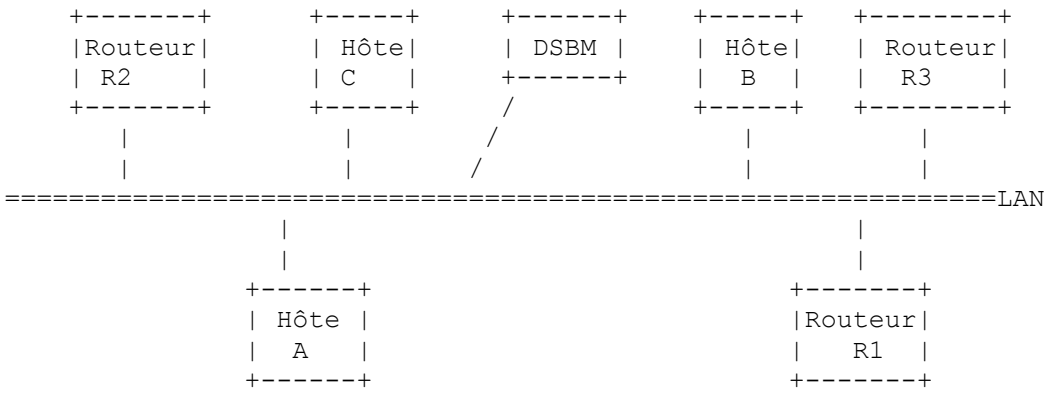
4.2 Généralités sur la procédure de contrôle d'admission fondée sur SBM

Une entité de protocole appelée "SBM désigné" (DSBM) existe pour chaque segment géré et est chargée du contrôle d'admission sur les demandes de réservation de ressource générées par les clients du DSBM dans ce segment. Sur un segment, un ou plusieurs SBM peuvent exister. Par exemple, de nombreux appareils à capacité de SBM peuvent être rattachés à un segment de couche 2 partagé tandis que deux commutateurs à capacité SBM peuvent partager un segment commuté unidirectionnel. Dans ce cas, un seul DSBM est choisi pour le segment. La procédure de choix dynamique du DSBM est décrite à l'Appendice A. La seule autre méthode approuvée pour spécifier un DSBM pour un segment géré est la configuration statique des appareils à capacité SBM.

La présence d'un DSBM fait du segment un "segment géré". Parfois, deux segments L2 ou plus peuvent être interconnectés par des appareils transparents à SBM. Dans ce cas, un seul DSBM va gérer les ressources pour les segments qui traitent la collection de ces segments comme un seul segment géré pour les besoins du contrôle d'admission.

4.2.1 Algorithme de base

Figure 1 - Exemple de segment géré



La Figure 1 montre l'exemple d'un segment géré dans un domaine L2 qui interconnecte un ensemble d'hôtes et de routeurs. Pour les besoins de l'exposé, on ignore la topologie physique réelle du domaine L2 (on suppose que c'est un segment L2 partagé et qu'un seul segment géré représente le domaine L2 entier). Un seul appareil SBM est désigné comme DSBM

pour le segment géré. On donnera des exemples de fonctionnement de DSBM sur des segments commutés et partagés plus loin dans ce document.

La procédure de base de contrôle d'admission fondée sur le DSBM fonctionne de la façon suivante :

1. Initialisation du DSBM : Au titre de sa configuration initiale, le DSBM obtient des informations comme les limites des fractions de ressources disponibles qui peuvent être réservées sur chaque segment géré sous son contrôle. Par exemple, la bande passante est une de ces ressources. Bien que des méthodes telles que l'auto-négociation du débit de liaison et la connaissance de la topologie des liaisons permette la découverte des capacités des liaisons, la configuration peut être nécessaire pour limiter la fraction de la capacité de liaison qui peut être réservée sur une liaison. La configuration va vraisemblablement être statique avec les appareils actuels de L2/L3. Des travaux futurs pourront permettre la découverte dynamique de ces informations. Le présent document ne spécifie pas le mécanisme de configuration.

2. Initialisation du client DSBM : Pour chaque interface rattachée, un client DSBM détermine si un DSBM existe sur l'interface. La procédure de découverte et de vérification de l'existence du DSBM pour un segment rattaché est décrite à l'Appendice A. Si le client lui-même est capable de servir de DSBM sur le segment, il peut choisir de participer au choix du DSBM. Au démarrage, un client DSBM vérifie d'abord qu'il existe un DSBM dans son domaine L2 afin qu'il puisse communiquer avec lui pour les besoins du contrôle d'admission.

Dans le cas d'un segment bidirectionnel, un choix peut n'être pas nécessaire car le SBM de chaque extrémité va normalement agir comme DSBM pour le trafic sortant dans chaque direction.

3. Contrôle d'admission fondé sur le DSBM : pour demander la réservation de ressources (par exemple, la bande passante du LAN dans un domaine L2) les clients DSBM (des appareils L3 à capacité RSVP tels que des hôtes et des routeurs) suivent les étapes suivantes :

   a) Lorsque un client DSBM envoie ou transmet un message PATH RSVP sur une interface rattachée à un segment géré, il envoie le message PATH au DSBM du segment au lieu de l'envoyer à l'adresse de destination de la session RSVP (comme cela se fait dans un traitement RSVP conventionnel). Après le traitement (et l'éventuelle mise à jour d'une ADSPEC) le DSBM va transmettre le message PATH vers son adresse de destination. Au titre de ce traitement, le DSBM construit et entretient un état PATH pour la session et note le bond L2/L3 précédent qui lui a envoyé le message PATH.

      Considérons le segment géré de la Figure 1. Supposons qu'un envoyeur à une session RSVP (l'adresse de session spécifie l'adresse IP de l'hôte A sur le segment géré de la Figure 1) réside en-dehors du domaine L2 du segment géré et envoie un message PATH qui arrive au routeur R1 qui est sur le chemin vers l'hôte A.

      Le client DSBM sur le routeur R1 transmet le message PATH de l'envoyeur au DSBM. Le DSBM traite le message PATH et le transmet vers le receveur RSVP (les règles de traitement et de transmission détaillées du message sont décrites à la Section 5). Dans le processus, le DSBM construit l'état PATH, mémorise le routeur R1 (ses adresses L2 et l3) comme bond précédant pour la session, met sa propre adresse de L2 et de L3 dans les objets PHOP (voir les explications plus loin) et s'insère effectivement comme nœud intermédiaire entre l'envoyeur (ou R1 dans la Figure 1) et le receveur (hôte A) sur le segment géré.

   b) Lorsque une application sur l'hôte A souhaite faire une réservation pour la session RSVP, l'hôte A suit les règles standard de traitement de message RSVP et envoie un message RSVP RESV à l'adresse L2/L3 du bond précédent (l'adresse des DSBM) obtenue du ou des objets PHOP dans le message PATH précédemment reçu.

   c) Le DSBM traite le message RSVP RESV sur la base de la bande passante disponible et retourne un message RESV_ERR au demandeur (l'hôte A) si la demande ne peut pas être accordée. Si des ressources suffisantes sont disponibles et si la demande de réservation est accordée, le DSBM transmet le message RESV vers le ou les PHOP sur la base de son état PATH local pour la session. Le DSBM fusionne les demandes de réservation pour la même session pour autant et lorsque c'est possible en utilisant des règles similaires à celles utilisées dans le traitement RSVP conventionnel (excepté pour un critère supplémentaire qui est décrit au paragraphe 5.8).

   d) Si le domaine L2 contient plus d'un segment géré, le demandeur (l'hôte A) et l'émetteur (le routeur R1) peuvent être séparés par plus d'un segment géré. Dans ce cas, le message PATH d'origine va être propagé à travers de nombreux DSBM (un pour chaque segment géré sur le chemin de R1 à A) établissant l'état PATH sur chaque DSBM. Donc, le message RESV va se propager de bond en bond en sens inverse à travers les DSBM intermédiaires et atteindre finalement l'émetteur d'origine (le routeur R1) sur le domaine L2 si le contrôle d'admission réussit sur tous les DSBM.

4.2.2 Améliorations au fonctionnement RSVP conventionnel

Les (D)SBM et les clients DSBM mettent en œuvre des ajouts mineurs au protocole RSVP standard. Ils sont résumés dans ce paragraphe. Une description détaillées des règles de traitement et de transmission du message figurent à la section 5.

4.2.2.1 Envoi de messages PATH au DSBM sur un segment géré

Les règles normales de transmission RSVP s'appliquent à un client DSBM lorsque il ne transmet pas un message PATH sortant sur un segment géré. Cependant, les messages PATH sortants sur un segment géré sont envoyés au DSBM pour le segment géré correspondant (le paragraphe 5.2 décrit comment les messages PATH sont envoyés au DSBM sur un segment géré).

4.2.2.2 Objets LAN_NHOP

Dans le traitement RSVP conventionnel sur des liaisons point à point, les nœuds RSVP (hôtes/routeurs) utilisent les objets RSVP_HOP (les informations de NHOP et de PHOP) pour garder trace des nœuds de prochain bond (nœuds vers l'aval sur le chemin des paquets de données dans un flux de trafic) et des nœuds du bond précédent (les nœuds vers l'amont par rapport au flux des données) sur le chemin entre un envoyeur et un receveur. Les routeurs le long du chemin d'un message PATH transmettent le message vers l'adresse de destination sur la base des tableaux d'acheminement de couche 3 (transmission de paquet).

Par exemple, considérons le domaine L2 de la Figure 1. Supposons que l'envoyeur (un hôte X) et le receveur (un hôte Y) dans une session RSVP résident tous deux en-dehors du domaine L2 montré sur la figure, mais que le messages PATH de l'envoyeur passe vers son receveur à travers les routeurs dans le domaine L2 en l'utilisant comme sous-réseau de transit. Supposons que le message PATH provenant de l'envoyeur X arrive au routeur R1. R1 utilise ses informations d'acheminement locales pour décider quel routeur de prochain bond (le routeur R2 ou le routeur R3) utiliser pour transmettre le message PATH vers l'hôte Y. Cependant, lorsque le chemin traverse un domaine L2 géré, on exige que les messages PATH et RESV passent à travers un DSBM pour chaque segment géré. Un tel domaine L2 peut s'étendre sur de nombreux segments gérés (et DSBM) et, normalement, les entités de protocole SBM sur des appareils L2 (comme un commutateur) vont servir de DSBM pour les segments gérés dans une topologie commutée. Lorsque R1 transmet le message PATH au DSBM (un appareil L2) le DSBM peut ne pas avoir les informations d'acheminement de couche 3 nécessaires pour choisir le routeur de sortie (entre R2 et R3) avant de transmettre le message PATH. Pour assurer un fonctionnement et un acheminement corrects des messages RSVP, on doit fournir des informations de transmission supplémentaires aux DSBM.

À cette fin, on introduit de nouveaux objets RSVP appelés objets d'adresse LAN_NHOP qui gardent trace du prochain bond de couche 3 lorsque le message PATH traverse un domaine L2 entre deux entités L3 (les nœuds RSVP PHOP et NHOP).

4.2.2.3 Inclusion d'adresses de couche 2 et de couche 3 dans le LAN_NHOP

Lorsque un client DSBM (hôte ou routeur agissant comme origine d'un message PATH) envoie un message PATH au DSBM, il doit inclure des informations de LAN_NHOP dans le message. Dans le cas d'une destination en envoi individuel, l'adresse de LAN_NHOP spécifie l'adresse de destination (si la destination est locale pour son domaine L2) ou l'adresse du routeur de prochain bond vers la destination. Dans notre exemple d'une session RSVP impliquant l'envoyeur X et le receveur Y avec le domaine L2 de la Figure 1 agissant comme sous-réseau de transit, R1 est le nœud d'entrée qui reçoit le message PATH. R1 détermine d'abord que R2 est le routeur de prochain bond (ou le nœud de sortie dans le domaine L2 pour l'adresse de session) puis insère un objet LAN_NHOP qui spécifie l'adresse IP de R2. Lorsque un DSBM reçoit un message PATH, il peut maintenant regarder l'adresse dans l'objet LAN_NHOP et transmettre le message PATH vers le nœud de sortie après avoir traité le message PATH. Cependant, on s'attend à ce que les appareils L2 (comme les commutateurs) agissent comme des DSBM sur le chemin au sein du domaine L2 et il peut n'être pas raisonnable d'attendre de ces appareils qu'ils aient une capacité ARP pour déterminer l'adresse MAC (on l'appelle L2ADDR pour l'adresse de couche 2) correspondant à l'adresse IP dans l'objet LAN_NHOP.

Donc, on exige que les informations de LAN_NHOP (générées par l'appareil L3) comportent à la fois l'adresse IP (adresse LAN_NHOP_L3) et l'adresse MAC correspondante (adresse LAN_NHOP_L2) pour le prochain bond L3 sur le domaine L2. L'adresse LAN_NHOP_L3 est utilisée par les entités de protocole SBM sur les appareils L3 pour transmettre le message PATH vers sa destination alors que l'adresse L2 est utilisée par les entités de protocole SBM sur les appareils L2 pour déterminer comment transmettre le message PATH vers le NHOP L3 (point de sortie du domaine L2). Le format exact des informations de LAN_NHOP et les objets pertinents sont décrits plus loin à l'Appendice B.

4.2.2.4 Similarités avec le traitement de message RSVP standard

- Lorsque un DSBM reçoit un message PATH RSVP, il traite le message PATH conformément aux règles de traitement de PATH décrites dans la spécification RSVP. En particulier, le DSBM restitue l'adresse IP du bond précédent à partir de l'objet RSVP_HOP dans le message PATH et mémorise l'adresse PHOP dans son état PATH. Il transmet alors le message PATH avec l'objet PHOP (RSVP_HOP) modifié pour refléter sa propre adresse IP (adresse RSVP_HOP_L3). Donc, le DSBM s'insère comme bond intermédiaire dans la chaîne des nœuds sur le chemin entre deux nœuds L3 à travers le domaine L2.

- L'état PATH dans un DSBM est utilisé pour transmettre les messages RESV suivants selon les règles standard de traitement de message RSVP. Lorsque le DSBM reçoit un message RESV, il traite le message et le transmet aux PHOP appropriés sur la base de son état PATH.

- Parce qu'un DSBM s'insère comme bond entre deux nœuds RSVP dans le chemin d'un flux RSVP, tous les messages en rapport avec RSVP (comme les PATH, PATH_TEAR, RESV, RESV_CONF, RESV_TEAR, et RESV_ERR) s'écoulent maintenant à travers le DSBM. En particulier, un message PATH_TEAR est acheminé à travers le ou les DSBM intermédiaires exactement comme le message PATH correspondant et l'état PATH local est le premier purgé à chaque bond intermédiaire avant que le message PATH_TEAR soit transmis.

- Jusque là, nous avons décrit comment le message PATH se propage à travers le domaine L2 en établissant l'état PATH à chaque DSBM le long des segments gérés sur le chemin. L'adresse de couche 2 (adresse LAN_NHOP_L2) dans l'objet LAN_NHOP devrait être utilisée par les appareils L2 le long du chemin pour décider comment transmettre le message PATH vers le prochain bond L3. De tels appareils appliqueront à l'adresse LAN_NHOP_L2 les règles standard de transmission IEEE 802.1D (par exemple, l'envoyer sur un seul accès déterminé d'après sa base de données de filtrage, ou en arroser tous les accès actifs sur l'arborescence de diffusion si l'adresse L2 n'apparaît pas dans la base de données de filtrage) comme elles sont appliquées normalement aux paquets de données destinés à l'adresse.

4.2.2.5 Inclusion d'adresses de couche 2 et de couche 3 dans les objets RSVP_HOP

Dans le traitement conventionnel de message RSVP, l'état PATH établi le long des nœuds sur un chemin est utilisé pour acheminer le message RESV d'un receveur à un envoyeur dans une session RSVP. Lorsque chaque nœud intermédiaire construit l'état Path, il se souvient du bond précédent (il mémorise l'adresse IP PHOP disponible dans l'objet RSVP_HOP d'un message entrant) qui lui a envoyé le message PATH et, lorsque le message RESV arrive, le nœud intermédiaire utilise simplement l'adresse PHOP mémorisée pour transmettre la RESV après la réussite de son traitement.

Dans notre cas, on s'attend à ce que les entités SBM qui résident sur les appareils L2 agissent comme des DSBM (et donc, soient des bonds RSVP intermédiaires dans un domaine L2) le long du chemin entre un envoyeur (PHOP) et un receveur (NHOP). Donc, lorsque un message RESV arrive à un DSBM, il doit utiliser l'adresse IP de PHOP mémorisée pour transmettre le message RESV à son bond précédent. Cependant, il peut n'être pas raisonnable de s'attendre à ce que les appareils L2 aient une antémémoire ARP ou la capacité ARP pour transposer l'adresse IP de PHOP en son adresse L2 correspondante avant de transmettre le message RESV.

Pour pallier le besoin d'une telle transposition d'adresse aux appareils L2, on utilise un objet RSVP_HOP_L2 dans le message PATH. L'objet RSVP_HOP_L2 inclut l'adresse de couche 2 (L2ADDR) du bond précédent et complète les informations d'adresse de couche 3 incluses dans l'objet RSVP_HOP (l'adresse RSVP_HOP_L3).

Lorsque un appareil L3 construit et transmet un message PATH sur un segment géré, il inclut son adresse IP (adresse IP de l'interface sur laquelle PATH est envoyé) dans l'objet RSVP_HOP et ajoute un objet RSVP_HOP_L2 qui comporte l'adresse L2 correspondante pour l'interface. Lorsque un appareil dans le domaine L2 reçoit un tel message PATH, il se souvient des adresses dans les objets RSVP_HOP et RSVP_HOP_L2 dans son état PATH et il réécrit alors les objets RSVP_HOP et RSVP_HOP_L2 avec ses propres adresses avant de transmettre le message PATH sur un segment géré.

Le format exact de l'objet RSVP_HOP_L2 est spécifié à l'Appendice B.

4.2.2.6 Détection de boucle

Lorsque une adresse de session RSVP est une adresse de diffusion groupée et qu'un SBM, un DSBM, et des clients DSBM partagent le même segment L2 (un segment partagé) il est possible qu'un SBM ou client DSBM reçoive une ou plusieurs copies d'un message PATH qu'il a transmis précédemment lorsque un DSBM sur le même circuit le transmet (voir au paragraphe 5.7 un exemple de ce cas). Pour faciliter la détection de telles boucles, on utilise un nouvel objet RSVP appelé LAN_LOOPBACK. Les clients DSBM ou les SBM (mais pas les DSBM qui reflètent un message PATH sur l'interface sur laquelle il est arrivé antérieurement) doivent modifier (ou ajouter si le message PATH n'en comporte pas déjà un)

l'objet LAN_LOOPBACK dans le message PATH avec leur propre adresse IP d'envoi individuel.

Maintenant, un SBM ou un client DSBM peut facilement détecter et éliminer les duplicatas en vérifiant le contenu de l'objet LAN_LOOPBACK (un message PATH dupliqué va faire la liste des propres adresses d'interface d'un appareil dans l'objet LAN_LOOPBACK). L'Appendice B spécifie le format exact de l'objet LAN_LOOPBACK.

4.2.2.7 802.1p, priorité d'utilisateur et TCLASS

Le modèle proposé par le groupe de travail Services intégrés exige l'isolement de chaque flux de trafic durant son transit à travers un réseau. Le motif de la séparation des flux de trafic est de fournir aux flux de services intégrés la protection contre les flux qui se conduisent mal et autre trafic au mieux qui partagent le même chemin. Les réseaux IEEE 802.3/Ethernet de base ne connaissent aucune notion de classes de trafic à discriminer parmi différents flux qui demandent des services différents. Cependant, IEEE 802.1p définit un moyen pour que les commutateurs différencient entre plusieurs valeurs de "priorité d'utilisateur" codées dans les paquets, qui représentent des classes de trafic différentes (voir les détails complémentaires dans [IEEE802Q], [IEEE8021p]). Les valeurs de priorité d'utilisateur peuvent être codées dans les paquets qui sortent d'un LAN (par exemple, dans l'octet FC exact de IEEE 802.5) ou en utilisant une encapsulation par dessus la couche MAC (par exemple, dans le cas de Ethernet, la valeur de priorité d'utilisateur allouée à chaque paquet sera portée dans l'en-tête de trame en utilisant le nouveau format de trame étendu défini par [IEEE8021Q]. L'IEEE ne fait, cependant, aucune recommandation sur la façon dont un envoyeur ou un réseau devrait utiliser les valeurs de priorité d'utilisateur. Un document d'accompagnement fait des recommandations sur l'usage des valeurs de priorité d'utilisateur (voir les détails dans la [RFC2815]).

Avec le modèle de services intégrés, les entités L3 (ou supérieures) qui transmettent des flux de trafic sur un segment L2 devraient effectuer une régulation par flux pour s'assurer que les flux n'excèdent pas leur spécification de trafic comme spécifié durant le contrôle d'admission. De plus, les appareils L3 peuvent étiqueter les trames de ces flux avec une valeur de priorité d'utilisateur pour identifier leur classe de service.

Pour les besoins de cet exposé, on se réfèrera à la valeur de priorité d'utilisateur portée dans l'en-tête de trame étendu comme à la "classe de trafic" d'un paquet. Avec le modèle ISSLL, les entités L3, qui envoient le trafic et qui utilisent le protocole SBM, peuvent choisir la classe de trafic appropriée pour les paquets sortants [RFC2815]. Ce choix peut être outrepassé par les appareils DSBM de la façon suivante. Une fois qu'un envoyeur envoie un message PATH, les DSBM vers l'aval vont insérer un nouvel objet de classe de trafic (l'objet TCLASS) dans le message PATH qui voyage vers le prochain appareil L3 (NHOP L3 pour le message PATH). Dans une certaine mesure, le contenu de l'objet TCLASS est traité comme l'objet ADSPEC dans les messages PATH RSVP. L'appareil L3 qui reçoit le message PATH doit retirer et mémoriser les objets TCLASS au titre de son état PATH pour la session. Plus tard, lorsque le même appareil L3 a besoin de transmettre un message RESV RSVP vers l'envoyeur d'origine, il doit inclure l'objet TCLASS dans le message RESV. Lorsque le message RESV arrive à l'envoyeur d'origine, celui-ci doit utiliser la valeur de priorité d'utilisateur provenant de l'objet TCLASS pour outrepasser le choix de classe de trafic marquée dans les paquets sortants.

Le format de l'objet TCLASS est spécifié à l'Appendice B. Noter que les objets TCLASS et autres objets spécifiques de SBM sont portés dans un message RSVP en plus de tous les autres objets RSVP normaux selon la [RFC2205].

4.2.2.8 Traitement de l'objet TCLASS

En résumé, l'utilisation des objets TCLASS exige les ajouts suivants au traitement conventionnel du message RSVP aux DSBM, SBM, et clients DSBM :

* Lorsque un DSBM reçoit un message PATH sur un segment géré et que le message PATH ne comporte pas un objet TCLASS, le DSBM PEUT ajouter un objet TCLASS au message PATH avant de le transmettre. Le DSBM détermine la valeur appropriée de priorité d'utilisateur pour l'objet TCLASS. Un mécanisme de sélection de la valeur de priorité d'utilisateur appropriée est décrit dans un document d'accompagnement [RFC2815].

* Lorsque un SBM ou DSBM reçoit un message PATH avec un objet TCLASS sur un segment géré dans un domaine L2 et qu'il a besoin de le transmettre sur un segment géré dans le même domaine L2, il va le mémoriser dans son état Path et normalement transmettre le message sans changer le contenu de l'objet TCLASS. Cependant, si le DSBM/SBM ne peut pas prendre en charge la classe de service représentée par la valeur de priorité d'utilisateur spécifiée par l'objet TCLASS dans le message PATH, il peut changer la valeur de priorité dans le TCLASS en une valeur de service sémantiquement "inférieure" pour refléter sa capacité et mémoriser la valeur de TCLASS modifiée dans son état Path.

Note : Un document d'accompagnement qui définit les transpositions d'Intserv sur les réseaux IEEE 802 [RFC2815] donne une définition précise des valeurs de priorité d'utilisateur et décrit comment les valeurs de priorité d'utilisateur sont

comparées pour déterminer "la plus basse" des deux valeurs ou "la plus basse" parmi toutes les valeurs de priorité d'utilisateur.

* Lorsque un SBM reçoit un message RESV avec un objet TCLASS, il peut utiliser les informations de classe de trafic (en plus des informations de flowspec usuelles dans le message RSVP) pour son propre contrôle d'admission pour le segment géré.

Noter que le présent document ne spécifie pas l'algorithme ou la politique réels utilisés pour le contrôle d'admission. D'un côté, un DSBM peut utiliser la demande de réservation par flux comme spécifié par la flowspec pour un contrôle d'admission pointu. À l'opposé, un DSBM peut ne considérer les informations de classe de trafic que pour un contrôle d'admission très grossier fondé sur une allocation statique des capacité de liaison pour chaque classe de trafic. Toute combinaison des options représentées par ces deux extrêmes peut aussi être utilisée.

* Lorsque un appareil de DSBM (à la L2 ou L3) reçoit un message RESV sans objet TCLASS et qu'il a besoin de transmettre le message RESV sur un segment géré au sein du même domaine L2, il devrait d'abord vérifier son état de chemin et vérifier si il a mémorisé une valeur de TCLASS. Si oui, il devrait inclure l'objet TCLASS dans le message RESV sortant après avoir effectué son propre contrôle d'admission. Si aucune valeur de TCLASS n'est mémorisée, il doit transmette le message RESV sans insérer d'objet TCLASS.

* Lorsque un client DSBM (résidant sur un appareil de L3 tel qu'un hôte ou un routeur bordure) reçoit l'objet TCLASS dans un message PATH qu'il accepte sur une interface, il devrait mémoriser l'objet TCLASS au titre de son état PATH pour l'interface. Plus tard, lorsque le client transmet un message RESV pour la même session sur l'interface, le client doit inclure l'objet TCLASS (inchangé par rapport à ce qu'il a reçu dans le message PATH précédent) dans le message RESV qu'il transmet sur l'interface.

* Lorsqu'un client DSBM reçoit un objet TCLASS dans un message RESV entrant sur un segment géré et que le contrôle d'admission local réussit pour la session sur l'interface sortante sur le segment géré, le client doit passer la valeur de priorité d'utilisateur contenue dans l'objet TCLASS à son classeur de paquet local. Cela va garantir que les paquets de données dans le flux RSVP admis qui sont ultérieurement transmis sur l'interface sortante vont contenir la valeur appropriée codée dans leur en-tête de trame.

* Lorsque un appareil de L3 reçoit un message PATH ou RESV sur un segment géré dans un domaine L2 et qu'il a besoin de transmettre le message PATH/RESV sur une interface en dehors de ce domaine, l'appareil L3 doit retirer l'objet TCLASS (ainsi que les objets LAN_NHOP, RSVP_HOP_L2, et LAN_LOOPBACK dans le cas du message PATH) avant de transmettre le message PATH/RESV. Si l'interface sortante est sur un domaine L2 séparé, ces objets peuvent être régénérés conformément aux règles de traitement applicables à cette interface.

5. Règles détaillées du traitement de message

5.1 Notes supplémentaires sur la terminologie

* Un appareil L2 peut avoir plusieurs interfaces avec des segments rattachés qui font partie du même domaine L2. Un commutateur dans un domaine L2 est un exemple d'un tel appareil. Un appareil qui a plusieurs interfaces peut contenir une entité de protocole SBM qui agit avec des capacités différentes sur chaque interface. Par exemple, une entité de protocole SBM pourrait agir comme un SBM sur l'interface A, et agir comme un DSBM sur l'interface B.

* Une entité de protocole SBM sur un appareil de couche 3 peut être un client DSBM, et un SBM, un DSBM, ou aucun d'eux (transparent à SBM). Les appareils L3 non transparents peuvent mettre en œuvre toute combinaison de ces rôles simultanément. Les clients DSBM résident toujours sur des appareils L3.

* Une entité de protocole SBM qui réside sur un appareil de couche 2 peut être un SBM, un DSBM ou aucun d'eux (transparent à SBM). Un appareil de couche 2 ne va jamais héberger un client DSBM.

5.2 Utilisation des adresses de diffusion groupée IP réservées

Comme il est dit plus haut, on exige que les clients DSBM transmettent les messages RSVP PATH à leur DSBM dans un domaine L2 avant qu'ils atteignent le prochain bond L3 sur le chemin. Les messages RSVP PATH sont adressés, conformément à la [RFC2205], à leur adresse de destination (qui peut être une adresse IP en envoi individuel ou en diffusion groupée). Lorsque un appareil L2 héberge un DSBM, un mécanisme simple à mettre en œuvre doit être fourni pour que l'appareil capture un message PATH entrant et le passe à l'agent DSBM local sans exiger que l'appareil L2 surveille les messages RSVP de couche 3.

De plus, les clients DSBM ont besoin de savoir comment adresser les messages SBM au DSBM. Pour faciliter le fonctionnement et permettre une liaison dynamique entre DSBM et client, il devrait être possible de détecter facilement et de s'adresser au DSBM existant sur un segment géré.

Pour faciliter la liaison dynamique entre DSBM et client ainsi que pour permettre aux appareils L2 une détection et une capture facile des messages PATH, on exige qu'on s'adresse à un DSBM en utilisant une adresse logique plutôt qu'une adresse physique. On utilise la ou les adresses réservées de diffusion groupée IP pour les besoins de communication avec un DSBM. En particulier, on exige que lorsque un client DSBM ou un SBM transmet un message PATH sur un segment géré, il soit adressé à une adresse réservée de diffusion groupée IP. Donc, un DSBM sur un appareil L2 a besoin d'être configuré d'une façon qui rende facile d'intercepter le message PATH et de le transmettre à l'entité de protocole SBM locale. Par exemple, cela peut impliquer simplement d'ajouter une entrée statique dans la base de données de filtrage (FDB, *filtering database*) de l'appareil pour l'adresse MAC de diffusion groupée correspondante pour assurer que les messages PATH sont interceptés et ne sont pas transmis plus loin sans l'intervention du DSBM.

De même, un DSBM envoie toujours les messages PATH sur un segment géré en utilisant une adresse réservée de diffusion groupée IP, et donc, les SBM ou les clients DSBM sur les segments gérés doivent simplement être configurés pour intercepter les messages destinés à l'adresse réservée de diffusion groupée sur les interfaces appropriées pour recevoir facilement les messages PATH.

Les messages RSVP RESV continuent d'être en envoi individuel pour l'adresse de bond précédent mémorisée au titre de l'état PATH à chaque bond intermédiaire.

On définit l'utilisation de deux adresses réservées de diffusion groupée. On appelle ces adresses "AllSBMAddress" et "DSBMLogicalAddress". Elles sont choisies dans la gamme des adresses de diffusion groupée locale, telles que :

* Elles ne soient pas passées à travers les appareils de couche 3.

* Elles soient passées de façon transparente à travers les appareils de couche 2 qui sont transparents à SBM.

* Elles soient configurées dans la base de données permanente des appareils de couche 2 qui héberge les SBM ou DSBM, de façon telle qu'elles soient dirigées sur l'entité de gestion SBM dans ces appareils. Cela pallie le besoin que ces appareils espionnent explicitement les paquets de contrôle en relation avec SBM.

* Les deux adresses réservées sont 224.0.0.16 (DSBMLogicalAddress) et 224.0.0.17 (AllSBMAddress).

Ces adresses sont utilisées comme décrit dans le tableau suivant :

Type	DSBMLogicaladdress	AllSBMAddress
Client DSBM	* Envoie les messages PATH à cette adresse	* Surveille cette adresse pour détecter la présence d'un DSBM
		* Surveille cette adresse pour recevoir les messages PATH transmis par le DSBM
SBM	* Envoie les messages PATH à cette adresse	* Surveille et envoie sur cette adresse pour participer au choix du DSBM
		* Surveille cette adresse pour recevoir les messages PATH transmis par le DSBM
DSBM	* Surveille cette adresse pour les messages PATH qui lui sont adressés	* Surveille et envoie sur cette adresse pour participer au choix du DSBM
		* Envoie les messages PATH à cette adresse

Les adresses L2 ou MAC qui correspondent aux adresses de diffusion groupée IP sont calculées par un algorithme qui utilise un bloc de l'adresse L2 réservée (les 24 bits de poids fort sont 00:00:5e). La RFC Numéros alloués [RFC1700] donne des détails supplémentaires.

5.3 Transposition d'adresse de couche 3 en adresse de couche 2

Comme on l'a dit précédemment, les DSBM ou clients DSBM qui résident sur des appareils L3 doivent inclure une adresse LAN_NHOP_L2 dans les informations de LAN_NHOP afin que les appareils L2 le long du chemin d'un message PATH n'aient pas besoin de déterminer séparément la transposition entre l'adresse LAN_NHOP_L3 dans l'objet LAN_NHOP et l'adresse L2 correspondante (par exemple, en utilisant ARP).

Pour les besoins d'une telle transposition aux appareils L3, on suppose une fonction de transposition appelée "map_address" qui effectue la transposition nécessaire :

objet L2ADDR = map_addr(L3Addr)

On ne spécifie pas comment la fonction est mise en œuvre ; la mise en œuvre peut simplement impliquer l'accès à l'entrée d'antémémoire ARP locale ou peut exiger d'effectuer une fonction ARP. La fonction retourne un objet L2ADDR qu'il n'est pas besoin qu'un appareil L3 interprète et qui peut être traité comme un objet opaque. Le format de l'objet L2ADDR est spécifié à l'Appendice B.

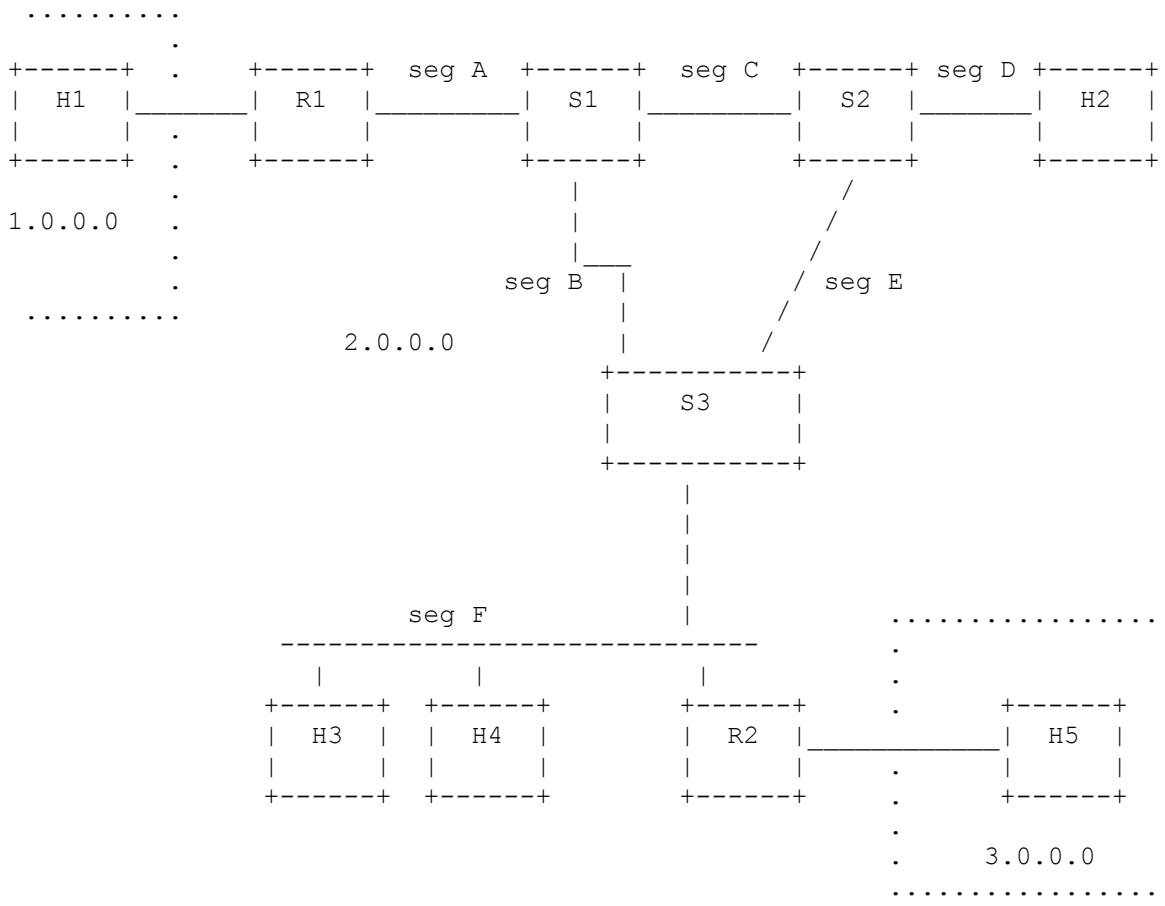
5.4 IP brut contre encapsulation UDP

On suppose que les DSBM, clients DSBM, et SBM utilisent seulement IP brut pour encapsuler les messages RSVP qui sont transmis sur un domaine L2. Donc, lorsque une entité de protocole SBM sur un appareil L3 transmet un message RSVP sur un segment L2, elle va utiliser seulement l'encapsulation IP RAW.

5.5 Règles de transmission

Les règles de traitement et transmission de message seront décrites dans le contexte de l'échantillon de réseau illustré dans la Figure 2.

Figure 2 – Échantillon de réseau ou domaine L2 consistant en segments L2 commutés et partagés



La Figure 2 illustre un échantillon de topologie de réseau consistant en trois sous-réseaux IP (1.0.0.0, 2.0.0.0, et 3.0.0.0) interconnectés en utilisant deux routeurs. Le sous-réseau 2.0.0.0 est un exemple de domaine L2 consistant en commutateurs, hôtes, et routeurs interconnectés à l'aide de segments commutés et d'un segment L2 partagé. L'échantillon de réseau contient les appareils suivants :

Appareil	Type	Type SBM
H1, H5	Hôte (couche 3)	Transparent à SBM
H2-H4	Hôte (couche 3)	Client DSBM
R1	Routeur (couche 3)	SBM
R2	Routeur (couche 3)	DSBM pour le segment F
S1	Commutateur (couche 2)	DSBM pour les segments A, B
S2	Commutateur (couche 2)	DSBM pour les segments C, D, E
S3	Commutateur (couche 2)	SBM

Les paragraphes qui suivent décrivent les règles que chacun de ces appareils devraient utiliser pour transmettre les messages PATH (les règles s'appliquent aussi bien aux messages PATH_TEAR). Elles sont décrites dans le contexte du réseau général illustré ci-dessus. Bien que les exemples ne traitent pas tous les scénarios, ils visent la plupart de ceux qui nous intéressent. Les exceptions seront exposées à part.

Les règles de transmission sont appliquées aux messages PATH reçus (routeurs et commutateurs) ou aux messages PATH d'origine (hôtes), comme suit :

1. Déterminer la ou les interfaces sur lesquelles transmettre le message PATH en utilisant les règles de transmission standard :
   * Si il y a un objet LAN_LOOPBACK dans le message PATH, et s'il porte l'adresse de cet appareil, éliminer le message en silence. (Voir au paragraphe ci-dessous les "notes supplémentaires" sur la transmission du message PATH sur un segment géré").
   * Les appareils de couche 3 utilisent l'adresse de session RSVP et effectuent un examen de l'acheminement pour déterminer la ou les interfaces de transmission.
   * Les appareils de couche 2 utilisent l'adresse de LAN_NHOP_L2 dans les informations de LAN_NHOP et les tableaux MAC de transmission pour déterminer la ou les interfaces de transmission. (Voir au paragraphe ci-dessous les "notes supplémentaires" sur la transmission du message PATH sur un segment géré").
2. Pour chaque interface de transmission :
   * Si l'appareil est de couche 3, déterminer si l'interface est sur un segment géré piloté par un DSBM, sur la base de la présence ou l'absence de messages I_AM_DSBM. Si l'interface n'est pas sur un segment géré, supprimer les objets RSVP_HOP_L2, LAN_NHOP, LAN_LOOPBACK, et TCLASS (si ils sont présents) et transmettre à la destination d'envoi individuel ou de diffusion groupée.

(Noter que les numéros de classe RSVP pour ces nouveaux objets sont choisis de telle sorte que si un message RSVP comporte ces objets, les nœuds qui sont à capacité RSVP mais ne participent pas au protocole SBM vont ignorer et supprimer en silence de tels objets.)

   * Si l'appareil est de couche 2 ou de couche 3 *et* si l'interface est sur un segment géré, passer à la règle n° 3.

3. Transmettre le message PATH sur le segment géré :
   * Si l'appareil est de couche 3, insérer les objets d'adresse LAN_NHOP, un objet LAN_LOOPBACK, et un objet RSVP_HOP_L2 dans le message PATH. Les objets LAN_NHOP portent les adresses LAN_NHOP_L3 et LAN_NHOP_L2 du prochain bon de couche 3. L'objet RSVP_HOP_L2 porte l'adresse propre de couche 2 de l'appareil, et l'objet LAN_LOOPBACK contient l'adresse IP de l'interface sortante.

     Un appareil L3 devrait utiliser la fonction map_addr() décrite plus haut pour obtenir une adresse de couche 2 correspondant à une adresse IP.

   * Si l'appareil héberge le DSBM pour le segment auquel l'interface de transmission est rattachée, faire ce qui suit :
   - Restituer les informations de PHOP à partir de l'objet HOP standard RSVP dans le message PATH, et les mémoriser. Elles seront utilisées pour acheminer les messages RESV en retour à travers le réseau L2. Si le message PATH est arrivé sur un segment géré, il va aussi contenir l'objet RSVP_HOP_L2 ; restituer alors et mémoriser aussi l'adresse de couche 2 du bond précédent dans l'état PATH.
   - Copier l'adresse IP de l'interface de transmission (les appareils de couche 2 doivent aussi avoir des adresses IP) dans l'objet HOP standard RSVP et l'adresse de couche 2 de l'interface de transmission dans l'objet RSVP_HOP_L2.
   - Si le message PATH reçu ne contient pas l'objet TCLASS, insérer un objet TCLASS. La valeur priorité_d'utilisateur insérée dans l'objet TCLASS se fonde sur des transpositions de service internes à l'appareil qui sont configurées conformément aux lignes directrices dont la liste figure dans la [RFC2815]. Si le message contient déjà l'objet TCLASS, la valeur de priorité_d'utilisateur peut être changée là encore sur la base des transpositions de service internes à l'appareil.

   * Si l'appareil est de couche 3 et héberge un SBM pour le segment auquel l'interface de transmission est

rattachée, il *est exigé* que les informations de PHOP soient restituées et mémorisées.

Si l'appareil est de couche 2 et héberge un SBM pour le segment auquel l'interface de transmission est rattachée, il *n'est pas exigé* de restituer et mémoriser les informations de PHOP. Si il ne le fait pas, le SBM doit laisser l'objet HOP standard RSVP et les objets RSVP_HOP_L2 dans le message PATH intacts et il ne va pas recevoir de messages RESV.

Si le SBM sur un appareil L2 choisit de réécrire les objets HOP RSVP et RSVP_HOP_L2 avec les adresses IP et de L2 de son interface de transmission, il va recevoir des messages RESV. Dans ce cas, il doit mémoriser les informations d'adresse de PHOP reçues dans le champ RSVP_HOP standard et dans les objets RSVP_HOP_L2 du message PATH incident.

Dans les deux cas mentionnés ci-dessus (appareils L2 ou L3) le SBM doit transmettre l'objet TCLASS inchangé dans le message PATH reçu.

* Copier l'adresse IP de l'interface de transmission dans l'objet LAN_LOOPBACK, sauf si l'entité de protocole SBM est un DSBM qui reflète un message PATH sur l'interface incidente. (Voir au paragraphe suivant "Notes supplémentaires sur la transmission d'un message PATH sur un segment géré").

* Si l'entité de protocole SBM est le DSBM pour le segment auquel l'interface de transmission est rattachée, elle doit envoyer le message PATH à AllSBMAddress.

* Si l'entité de protocole SBM est un SBM ou un client DSBM sur le segment auquel l'interface de transmission est rattachée, elle doit envoyer le message PATH à DSBMLogicalAddress.

5.5.1 Notes supplémentaires sur la transmission d'un message PATH sur un segment géré

La règle n° 1 déclare que les règles normales de transmission IEEE 802.1D devraient être utilisées pour déterminer les interfaces sur lesquelles le message PATH devrait être transmis. Dans le cas de paquets de données, les règles de transmission standard à un appareil de couche 2 imposent que le paquet ne devrait pas être transmis sur l'interface d'où il a été reçu. Cependant, dans le cas d'un DSBM qui reçoit un message PATH sur un segment géré, les exceptions suivantes s'appliquent :

E1. Si l'adresse dans l'objet LAN_NHOP est en envoi individuel, consulter la base de données de filtrage (FDB, *filtering database*) pour déterminer si l'adresse de destination figure sur la liste sur la même interface sur laquelle le message a été reçue. Si oui, suivre la règle ci-dessous sur "refléter un message PATH sur une interface" décrite ci-dessous ; autrement, poursuivre comme d'habitude le reste du traitement du message.

E2. Si il y a des membres de l'adresse du groupe de diffusion groupée (spécifiés par les adresses dans l'objet LAN_NHOP) sur le segment d'où le message a été reçu, le message devrait être retransmis sur l'interface d'où il a été reçu et on devrait suivre la règle sur "refléter un message PATH sur une interface" décrite ci-dessous.

*** Refléter un message PATH sur une interface ***

Dans les circonstances décrites ci-dessus, lorsque un DSBM reflète le message PATH sur une interface sur laquelle il a été reçu, il doit l'adresser en utilisant AllSBMAddress.

Comme il est possible à un DSBM de refléter un message PATH sur l'interface d'où il a été reçu, des précautions doivent être prises pour éviter de faire une boucle infinie de ces messages. L'objet LAN_LOOPBACK traite cette question. Toutes les entités de protocole SBM (sauf les DSBM qui reflètent un message PATH) réécrivent l'objet LAN_LOOPBACK dans le message PATH avec l'adresse IP de l'interface sortante. Les DSBM qui reflètent un message PATH, laissent l'objet LAN_LOOPBACK inchangé. Donc, les entités de protocole SBM vont toujours être capables de reconnaître un message en envoi individuel reflété par la présence de leur propre adresse dans l'objet LAN_LOOPBACK. Ces messages devraient être éliminés en silence.

5.6 Application des règles – Session en envoi individuel

Voyons comment les règles sont appliquées dans le réseau général illustré précédemment à la Figure 2.

Supposons que H1 envoie un PATH pour une session en envoi individuel pour laquelle H5 est le receveur. Le message PATH suivant est composé par H1 :

	Contenu RSVP
adresse IP de session RSVP	adresse IP de H5 (3.0.0.35)
Gabarit d'envoyeur	adresse IP de H1 (1.0.0.11)
PHOP	adresse IP de H1 (1.0.0.11)
RSVP_HOP_L2	n/a (H1 n'envoie pas sur un segment géré)
LAN_NHOP	n/a (H1 n'envoie pas sur un segment géré)
LAN_LOOPBACK	n/a (H1 n'envoie pas sur un segment géré)
	En-tête IP
Adresse de source	adresse IP de H1 (1.0.0.11)
Adresse de destination	adresse IP de H5 (3.0.0.35, en supposant le mode brut et l'alerte de routeur)
	En-tête MAC
Adresse de destination	adresse L2 correspondant à R1 (déterminée par map_addr() et les tableaux d'acheminement chez H1)

Comme H1 n'envoie pas sur un segment géré, le message PATH est composé et transmis conformément aux règles de traitement RSVP standard.

À réception du message PATH, R1 compose et transmet un message PATH comme suit :

	Contenu RSVP
Adresse IP de session RSVP	adresse IP de H5
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de R1 (2.0.0.1) (sert de chemin de retour pour les messages RESV)
RSVP_HOP_L2	adresse L2 de R1
LAN_NHOP	LAN_NHOP_L3 (2.0.0.2) et adresse LAN_NHOP_L2 de R2 (L2ADDR) (c'est le prochain bond de couche 3)
LAN_LOOPBACK	adresse IP de R1 (2.0.0.1)
	En-tête IP
Adresse de source	adresse IP de H1
Adresse de destination	adresse IP DSBMLogical (224.0.0.16)
	En-tête MAC
Adresse de destination	adresse MAC DSBMLogical

* R1 fait une recherche d'acheminement sur l'adresse de session RSVP, pour déterminer l'adresse IP du prochain bond de couche3, R2.

* Il détermine que R2 est accessible via le segment A et que le segment A est géré par un DSBM, S1.

* Donc, il conclut qu'il envoie sur un segment géré, et compose les objets LAN_NHOP pour porter les adresses de couche 3 et de couche 2 de prochain bond. Pour composer l'objet LAN_NHOP L2ADDR, il invoque la fonction de transposition d'adresse de L3 en L2 ("map_address") pour découvrir l'adresse de MAC pour l'appareil L3 de prochain bond, et insère alors un objet LAN_NHOP_L2ADDR (qui porte l'adresse MAC) dans le message.

* Comme R1 n'est pas le DSBM pour le segment A, il envoie le message PATH à DSBMLogicalAddress.

À réception du message PATH, S1 compose comme suit un message PATH et le transmet :

	Contenu RSVP
Adresse IP de session RSVP	adresse IP de H5
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de S1 (pose le chemin de retour pour les messages RESV)
RSVP_HOP_L2	adresse de L2 de S1
LAN_NHOP	adresse LAN_NHOP_L3 (IP) et LAN_NHOP_L2 de R2 (les appareils de couche 2 ne modifient pas le LAN_NHOP)
LAN_LOOPBACK	adresse IP de S1
	En-tête IP
Adresse de source	adresse IP de H1
Adresse de destination	AllSBMIPaddr (224.0.0.17, car S1 est le DSBM pour le segment B).

En-tête MAC

Adresse de destination	adresse MAC Tous_les_SBM (car S1 est le DSBM pour le segment B).

* S1 cherche les informations d'adresse de LAN_NHOP pour déterminer l'adresse L2 vers laquelle il devrait transmettre le message PATH.

* À partir des tableaux de transmission de pontage, il détermine que l'adresse L2 est accessible via le segment B.

* S1 insère l'objet RSVP_HOP_L2 et réécrit l'objet RSVP HOP (PHOP) avec ses propres adresses.

* Comme S1 est le DSBM pour le segment B, il adresse le message PATH à AllSBMAddress.

À réception du message PATH, S3 compose comme suit un message PATH et le transmet.

Contenu RSVP

Adresse IP de session RSVP	adresse IP de H5
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de S3 (porte le chemin de retour pour les messages RESV)
RSVP_HOP_L2	adresse L2 de S3
LAN_NHOP	adresse LAN_NHOP_L3 (IP) et LAN_NHOP_L2 (MAC) de R2. (Les appareils L2 ne modifient pas LAN_NHOP.)
LAN_LOOPBACK	adresse IP de S3

En-tête IP

Adresse de source	adresse IP de H1
Adresse de destination	adresse IP DSBMLogical (car S3 n'est pas le DSBM pour le segment F)

En-tête MAC

Adresse de destination	adresse MAC DSBMLogical

* S3 recherche les informations de l'adresse de LAN_NHOP pour déterminer l'adresse L2 vers laquelle il devrait transmettre le message PATH.

* À partir des tableaux de transmission de pontage, il détermine que l'adresse L2 est accessible via le segment F.

* Il a découvert que R2 est le DSBM pour le segment F. Il envoie donc le message PATH à DSBMLogicalAddress.

* Noter que S3 peut choisir ou non de réécrire les objets PHOP avec ses propres adresses IP et L2. Si il le fait, il va recevoir les messages RESV. Dans ce cas, il doit aussi mémoriser les informations de PHOP reçues dans le message PATH incident afin qu'il soit capable de transmettre les messages RESV sur le chemin correct.

À réception du message PATH, R2 compose comme suit un message et le transmet :

Contenu RSVP

Adresse IP de session RSVP	adresse IP de H5
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de R2 (porte le chemin de retour pour les messages RESV)
RSVP_HOP_L2	Retiré par R2 (R2 n'envoie pas sur un segment géré)
LAN_NHOP	Retiré par R2 (R2 n'envoie pas sur un segment géré)

En-tête IP

Adresse de source	adresse IP de H1
Adresse de destination	adresse IP de H5, l'adresse de la session RSVP

En-tête MAC

Adresse de destination	adresse L2 correspondant à H5, le prochain bond de couche 3

* R2 fait une recherche d'acheminement sur l'adresse de la session RSVP, pour déterminer l'adresse IP du prochain bond de couche 3, H5.

* Il détermine que H5 est accessible via un segment pour lequel il n'y a pas de DSBM (ce n'est pas un segment géré).

* Donc, il retire les objets LAN_NHOP et RSVP_HOP_L2 et place l'adresse de la session RSVP dans l'adresse de destination de l'en-tête IP. Il place l'adresse L2 du prochain bond de couche 3 dans l'adresse de destination de l'en-tête MAC et transmet le message PATH à H5.

5.7 Application des règles – Session en diffusion groupée

Les règles décrites ci-dessus s'appliquent aussi aux sessions en diffusion groupée. Pour les besoins de l'exposé, on suppose que les appareils de couche 2 suivent les membres du groupe de diffusion groupée individuellement sur chaque accès. Les appareils de couche 2 qui ne le font pas vont simplement générer du trafic en dehors de la diffusion groupée. C'est le cas pour les appareils L2 qui ne mettent pas en œuvre le filtrage de diffusion groupée ou la capacité GARP/GMRP.

Supposons que H1 envoie un PATH pour une session en diffusion groupée pour laquelle H3 et H5 sont les receveurs. Les règles sont appliquées comme dans le cas de l'envoi individuel décrit précédemment, jusqu'à ce que le message PATH atteigne R2, avec l'exception suivante : l'adresse de session RSVP et le LAN_NHOP portent les adresses de diffusion groupée de destination plutôt que les adresses d'envoi individuel portées dans l'exemple de l'envoi individuel.

Regardons maintenant les processus qui s'appliquent à R2 à réception du message PATH. On rappelle que R2 est le DSBM pour le segment F. Donc, S3 aura transmis son message PATH à DSBMLogicalAddress, pour qu'il soit collecté par R2. Le message PATH n'aura pas été vu par H3 (un des receveurs de la diffusion groupée) car il surveille seulement AllSBMAddress, et pas DSBMLogicalAddress pour les messages PATH entrants. On s'appuie sur R2 pour réfléchir le message PATH sur le segment f, et pour le transmettre à H5. R2 transmet le message PATH suivant au segment F :

	Contenu RSVP
Adresse de session RSVP	adresse de session en diffusion groupée
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de R2 (porte le chemin de retour pour les messages RESV)
RSVP_HOP_L2	adresse L2 de R2
LAN_NHOP	adresse de session en diffusion groupée et adresse L2 correspondante
LAN_LOOPBACK	adresse IP de S3 (les DSBM reflétant un message PATH ne modifient pas cet objet)
	En-tête IP
Adresse de source	adresse IP de H1
Adresse de destination	adresse IP Tous_les_SBM (car R2 est le DSBM pour le segment F)
	En-tête MAC
Adresse de destination	Adresse MAC Tous_les_SBM (car R2 est le DSBM pour le segment F)

Comme H3 surveille l'adresse Tous_les_SBM, il va recevoir le message PATH reflété par R2. Noter que R2 a violé ici les règles standard de transmission en renvoyant un message entrant à l'interface d'où il avait été reçu. Il s'est protégé contre les boucles en laissant l'adresse de S3 inchangée dans l'objet LAN_LOOPBACK.

R2 transmet le message PATH suivant à H5 :

	Contenu RSVP
Adresse de session RSVP	adresse de session en diffusion groupée
Gabarit d'envoyeur	adresse IP de H1
PHOP	adresse IP de R2 (porte le chemin de retour pour les messages RESV)
RSVP_HOP_L2	Retiré par R2 (R2 n'envoie pas sur un segment géré)
LAN_NHOP	Retiré par R2 (R2 n'envoie pas sur un segment géré)
LAN_LOOPBACK	Retiré par R2 (R2 n'envoie pas sur un segment géré)
	En-tête IP
Adresse de source	adresse IP de H1
Adresse de destination	adresse de session en diffusion groupée
	En-tête MAC
Adresse de destination	adresse MAC correspondant à l'adresse de session en diffusion groupée

* R2 détermine qu'il y a un receveur de diffusion groupée accessible via un segment pour lequel il n'y a pas de DSBM. Donc, il retire les objets LAN_NHOP et RSVP_HOP_L2 et place l'adresse de session RSVP dans l'adresse de destination de l'en-tête IP. Il place l'adresse L2 correspondante dans l'adresse de destination de l'en-tête MAC et envoie

en diffusion groupée le message vers H5.

5.8 Fusion des objets de classe de trafic

Lorsque un client DSBM reçoit des objets TCLASS de différents envoyeurs (différents messages PATH) dans la même session RSVP et qu'il a besoin de les combiner pour renvoyer un seul message RESV (comme dans une réservation de style générique) le client DSBM doit choisir une valeur appropriée qui corresponde à la classe de trafic du délai désiré. Un document d'accompagnement discute les lignes directrices pour le choix de la classe de trafic sur la base du service désiré et des informations de la TSpec [RFC2815].

De plus, lorsque un SBM ou DSBM a besoin de fusionner les RESV provenant de différents prochains bonds à un point de fusion, il doit décider comment traiter les valeurs de TCLASS dans les RESV entrantes si elles ne correspondent pas. Considérons le cas où une réservation est en place pour un flux à un DSBM (ou SBM) avec un contrôle d'admission réussi fait pour la TCLASS demandée dans la première RESV pour le flux. Si une autre RESV pour le même flux (pas le rafraîchissement de la RESV précédemment admise) arrive au DSBM, le DSBM doit d'abord vérifier la valeur de TCLASS dans la nouvelle RESV par rapport à la valeur de TCLASS dans la RESV déjà installée. Si les deux valeurs sont les mêmes, les demandes RESV sont fusionnées et la nouvelle RESV, fusionnée, est installée et transmise en utilisant les règles normales du traitement de message. Cependant, si les deux valeurs ne sont pas identiques, le DSBM doit générer et envoyer un message RESV_ERR vers l'envoyeur (NHOP) du message RESV le plus récent. Le RESV_ERR doit spécifier le code d'erreur correspondant au "erreur de contrôle de trafic" de RSVP (RESV_ERR code 21) qui indique l'échec d'une fusion de deux demandes de service incompatibles (sous-code 01 pour l'erreur de contrôle de trafic RSVP) [RFC2205]. Le message RESV_ERR peut inclure des objets supplémentaires pour aider les nœuds vers l'aval à récupérer de cette condition. La définition et l'usage de tels objets sort du domaine d'application du présent mémoire.

5.9 Fonctionnement des appareils transparents à SBM

Les appareils transparents à SBM ne connaissent pas la totalité du protocole SBM/DSBM. Ils n'interceptent pas les messages adressés à l'une des adresses du groupe local en relation avec SBM (les adresses DSBMLogicalAddress et ALLSBMAddress) mais les font plutôt passer plus loin. Il en résulte qu'ils ne divisent pas la portée du domaine d'élection du DSBM, ils ne participent pas explicitement à l'acheminement des messages PATH ou RESV, et ils ne participent pas au contrôle d'admission. Ils sont entièrement transparents par rapport au fonctionnement de SBM.

Conformément aux définitions fournies, les segments physiques interconnectés par des appareils transparents à SBM sont considérés comme un seul segment géré. Donc, les DSBM doivent effectuer le contrôle d'admission sur de tels segments gérés, avec une connaissance limitée de la topologie du segment. Dans ce cas, l'administrateur de réseau devrait configurer le DSBM pour chaque segment géré, avec une approximation raisonnable de la capacité du segment. Une politique prudente va configurer le DSBM pour le chemin de plus faible capacité à travers le segment géré. Une politique libérale va configurer le DSBM pour le chemin de plus forte capacité à travers le segment géré. Un administrateur de réseau va vraisemblablement choisir une valeur entre les deux, sur la base du niveau de garanties requis et d'une certaine connaissance des schémas de trafic vraisemblables.

Le présent document ne spécifie pas le mécanisme de configuration ou le choix d'une politique.

5.10 Fonctionnement des appareils SBM qui ne sont pas DSBM

Dans l'exemple illustré, S3 héberge un SBM, mais le SBM sur S3 n'a pas gagné l'élection pour agir comme DSBM sur un segment. On peut se demander quelle est la fonction d'une entité de protocole comme un SBM. Ces SBM assurent en fait deux fonctions utiles. Tout d'abord, les SBM supplémentaires restent passifs à l'arrière plan pour la tolérance aux fautes. Ils écoutent les annonces périodiques provenant du DSBM en cours pour le segment géré (l'Appendice A décrit cela plus en détail) et ils se manifestent pour élire un nouveau DSBM lorsque le DSBM actuel a une défaillance ou cesse d'être opérationnel pour une raison quelconque. Ensuite, ces SBM fournissent aussi l'important service de se partager le domaine d'élection et de réduire la taille et la complexité des segments gérés. Par exemple, considérons à nouveau l'exemple de topologie de la Figure 3 : l'appareil S3 contient un SBM qui n'est un DSBM pour aucun des segments, B, E, ou F, qui lui sont rattachés. Cependant, si l'entité de protocole SBM n'est pas présente sur S3, les segments B et F ne seraient pas des segments séparés du point de vue du protocole SBM. Ils constitueraient plutôt un seul segment géré, géré par un seul DSBM. Comme l'entité SBM sur S3 divise le domaine d'élection, le segment B et le segment F sont chacun gérés par des DSBM distincts. Chacun de ces segments a une topologie triviale et une capacité bien définie. Il en résulte que les DSBM pour ces segments n'ont pas besoin d'effectuer le contrôle d'admission sur la base d'approximations (comme se serait le cas si S3 était un SBM transparent).

Noter que les entités de protocole SBM qui ne sont pas des DSBM ne sont pas obligées de réécrire le PHOP dans les messages PATH incidents avec leur propre adresse. Cela parce qu'il n'est pas nécessaire que les messages RESV soient

acheminés à travers ces appareils. Les message RESV sont seulement obligés d'être acheminés à travers la séquence de DSBM correcte. Les SBM ne peuvent pas traiter les messages RESV qui passent bien à travers eux, autrement qu'en les transmettant vers leur adresse de destination, en utilisant les règles standard de transmission.

Les entités de protocole SBM qui ne sont pas des DSBM sont obligées de réécrire l'adresse dans l'objet LAN_LOOPBACK avec leur propre adresse, afin d'éviter de mettre en boucle les messages de diffusion groupée. Cependant, aucun état n'est à mémoriser.

6. Considérations d'inter fonctionnement

Il n'y a que peu de questions d'inter opérabilité intéressantes qui se rapportent au déploiement d'une méthode de contrôle d'admission fondée sur le DSBM dans un environnement consistant en nœuds de réseau avec ou sans capacité RSVP. On fait ci-après une liste de certains de ces scénarios et on explique comment les clients et nœuds à capacité SBM peuvent fonctionner dans ces scénarios.

6.1 Domaine de couche L2 sans capacité RSVP

Il est possible d'envisager des domaines L2 qui n'utilisent pas la signalisation RSVP pour demander des réservations de ressources, mais utilisent plutôt quelque autre mécanisme (par exemple, SNMP ou la configuration statique) pour réserver de la bande passante sur un appareil de réseau particulier comme un routeur. Dans ce cas, la question est de savoir comment fonctionne une méthode de contrôle d'admission fondée sur le DSBM et comment elle interopère avec le mécanisme non RSVP. La méthode fondée sur SBM n'essaye pas de fournir une solution de contrôle d'admission pour un tel environnement. L'approche fondée sur le SBM fait partie d'une approche de signalisation de bout en bout pour établir des réservations de ressources et ne cherche pas à fournir une solution pour un scénario de configuration fondé sur SNMP.

Comme on l'a dit précédemment, l'approche fondée sur SBM peut cependant coexister avec n'importe quel autre mécanisme d'allocation de bande passante non RSVP dans la mesure où les ressources qui sont réservées sont soit partagées de façon statique entre les différents mécanismes, soit sont résolues dynamiquement à travers un répartiteur de bande passante commun afin qu'il n'y ait pas de sur engagement de la même ressource.

6.2 Domaine L2 avec des appareils de couche 2 transparents à SBM

Ce scénario a été visé précédemment dans ce document. La méthode fondée sur SBM est conçue pour fonctionner dans un tel environnement. Lorsque les appareils L2 transparents à SBM s'interconnectent à des appareils à capacité SBM, le segment géré résultant est une combinaison d'un ou plusieurs segments physiques et le DSBM pour le segment géré peut n'être pas aussi efficace dans l'allocation des ressources qu'il le serait si tous les appareils L2 étaient à capacité SBM.

6.3 Domaine de couche 2 sur lequel des envoyeurs fondés sur RSVP ne sont pas des clients DSBM

Tous les envoyeurs qui génèrent des flux de trafic fondés sur RSVP sur un segment géré DOIVENT être à capacité SBM et participer au protocole SBM. L'utilisation de la version standard non SBM de RSVP peut résulter en une sur allocation de ressources, car une telle utilisation court-circuite la fonction de gestion de ressource du DSBM. Tous les autres envoyeurs (c'est-à-dire, les envoyeurs qui n'envoient pas des flux soumis au contrôle d'admission RSVP) devraient être des applications élastiques qui envoient du trafic de priorité inférieure à celle du trafic RSVP, et utilisent des mécanismes d'évitement d'encombrement du style TCP.

Tous les DSBM, SBM, ou clients DSBM sur un segment géré (un segment avec un DSBM actuellement actif) doivent ne pas accepter de messages PATH provenant d'envoyeurs qui ne sont pas à capacité SBM. Les messages PATH provenant de tels appareils peuvent être facilement détectés par les SBM et les clients DSBM car ils ne seront pas envoyés en diffusion groupée à l'adresse ALLSBMAddress (dans le cas de SBM et de clients DSBM) ou à l'adresse DSBMLogicalAddress (dans le cas de DSBM).

6.4 Routeur non SBM qui interconnecte deux domaines de couche 2 gérés par DSBM

Les messages SBM en diffusion groupée (par exemple, messages d'élection et PATH) ont une portée locale et ne sont pas destinés à passer entre les deux domaines. Un routeur non SBM correctement configuré ne va pas passer de tels messages entre les domaines. Une mauvaise mise en œuvre de routeur qui le fait peut causer un fonctionnement incorrect du protocole SBM et il peut s'ensuivre une sur ou sous allocation des ressources.

6.5 Interopérabilité avec des clients RSVP qui utilisent l'encapsulation UDP et ne sont pas capables de recevoir/envoyer de messages RSVP utilisant RAW_IP

Le présent document stipule que les DSBM, les clients DSBM, et les SBM utilisent seulement IP brut pour encapsuler les messages RSVP qui sont transmis vers un domaine L2.La RFC2205 (le projet de norme RSVP) comporte la prise en charge de l'encapsulation aussi bien de IP brut que de UDP . Donc, un nœud RSVP qui n'utilise que l'encapsulation UDP ne sera pas capable d'interopérer avec le DSBM à moins que le DSBM accepte et prenne en charge les messages RSVP encapsulés dans UDP.

7. Lignes directrices pour la mise en œuvre

Dans ce qui suit, on donne des lignes directrices pour la mise en œuvre de différents aspects de la procédure de contrôle d'admission fondée sur SBM, incluant des suggestions pour l'initialisation du DSBM, etc.

7.1 Initialisation du DSBM

Comme on l'a dit précédemment, l'initialisation du DSBM comporte la configuration de la bande passante maximum qui peut être réservée sur un segment géré sous son contrôle. On suggère les lignes directrices suivantes.

Dans le cas d'un segment géré consistant en appareils L2 interconnectés par un seul segment partagé, les entités DSBM sur de tels appareils devraient supposer la bande passante de l'interface comme étant la bande passante totale de la liaison. Dans le cas d'un DSBM situé sur un commutateur L2, il peut de plus devoir être configuré avec une estimation de la capacité de commutation de l'appareil si celle-ci est inférieure à la bande passante de la liaison, et éventuellement avec une estimation des ressources en mémoire tampon du commutateur (voir dans la [RFC2816] le modèle architectural supposé pour les commutateurs L2). Sachant la bande passante totale de la liaison, le DSBM peut de plus être configuré pour limiter la quantité maximum de bande passante pour les flux à capacité RSVP pour assurer une capacité de réserve pour le trafic au mieux.

7.2 Fonctionnement des DSBM dans différentes topologies de couche 2

Selon la topologie L2, un DSBM peut être appelé à gérer les ressources pour un ou plusieurs segments et la mise en œuvre doit tenir compte des implications pour l'efficacité de l'utilisation d'un DSBM dans différentes topologies L2. Les topologies L2 triviales consistent en un seul "segment physique". Dans ce cas, le "segment géré" est équivalent à un seul segment. Les topologies L2 complexes peuvent consister en un certain nombre de segments. Le contrôle d'admission sur un tel segment L2 étendu peut être effectué à partir d'un seul réservoir de ressources, similaire à un seul segment partagé, du point de vue d'un seul DSBM.

Cette configuration compromet l'efficacité avec laquelle le DSBM peut allouer les ressources, parce que un seul DSBM est obligé de prendre les décisions de contrôle d'admission pour toutes les demandes de réservation au sein de la topologie L2, sans savoir quels sont les segments physiques réels qui sont affectés par la réservation.

On peut apporter des améliorations à l'efficacité de l'allocation des ressources en subdivisant le segment complexe en un certain nombre de segment gérés, gérés chacun par leur propre DSBM. Dans ce cas, chaque DSBM gère un segment géré ayant une topologie relativement simple. Comme les segments gérés sont plus simples, le DSBM peut être configuré avec une estimation plus précise des ressources disponibles pour toutes les réservations dans le segment géré. Dans la configuration idéale, chaque segment physique est un segment géré et est géré par son propre DSBM. On ne fera pas d'hypothèses sur le nombre de segments gérés mais on dira simplement que dans les topologies L2 complexes, l'efficacité de l'allocation des ressources s'améliore avec l'augmentation de la granularité des segments gérés.

8. Considérations pour la sécurité

Les règles de formatage et d'utilisation de message décrites dans la présente note soulèvent des problèmes de sécurité, identiques à ceux soulevés par l'utilisation de RSVP et des services intégrés. Il est nécessaire de contrôler et d'authentifier l'accès aux qualités de service améliorées permises par la technologie décrite dans cette RFC. Cette exigence est exposée plus en détails dans les [RFC2205], [RFC2211], et [RFC2212].

La [RFC2747] décrit le mécanisme utilisé pour protéger l'intégrité des messages RSVP qui portent les informations décrites ici. Une mise en œuvre de SBM devrait satisfaire aux exigences de la présente RFC et fournir les mécanismes suggérés comme si c'était une mise en œuvre RSVP conventionnelle. Elle devrait de plus utiliser les mêmes mécanismes pour protéger les objets supplémentaires, spécifiques de SBM, dans un message.

Finalement, il est aussi nécessaire d'authentifier les candidats DSBM durant le processus de choix, et un mécanisme fondé sur un secret partage entre les candidats DSBM peut être utilisé. Le mécanisme défini dans la [RFC2747] devrait être utilisé.

9. Références

[IEEE802Q] "IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks", Draft Standard P802.1Q/D9, 20 février 1998.

[IEEEP8021p] "Information technology - Telecommunications and information exchange between systems - Local and metropolitan area networks - Common specifications - Part 3: Media Access Control (MAC) Bridges: Revision (Incorporating IEEE P802.1p: Traffic Class Expediting and Dynamic Multicast Filtering)", ISO/IEC Final CD 15802-3 IEEE P802.1D/D15, 24 novembre 1997.

[IEEE8021D] "MAC Bridges", ISO/IEC 10038, ANSI/IEEE Std 802.1D- 1993.

[RFC2205] R. Braden, éd., L. Zhang, S. Berson, S. Herzog, S. Jamin, "Protocole de réservation de ressource (RSVP) -- version 1, spécification fonctionnelle", septembre 1997. (*MàJ par* RFC2750, RFC3936, RFC4495) *(P.S.)*

[RFC2206] F. Baker, J. Krawczyk, A. Sastry, "Base de données d'informations de gestion RSVP avec SMIv2", septembre 1997. *(PS)*

[RFC2210] J. Wroclawski, "Utilisation de RSVP avec les services intégrés de l'IETF", septembre 1997. *(P.S.)*

[RFC2211] J. Wroclawski, "Spécification du service d'élément de réseau à charge contrôlée", septembre 1997. *(P.S.)*

[RFC2212] S. Shenker, C. Partridge, R. Guerin, "Spécification de la qualité de service garantie", septembre 1997. *(P.S.)*

[RFC2213] F. Baker et autres, "Base de données d'informations de gestion de services intégrés avec SMIv2", septembre 1997. *(P.S.)*

[RFC2215] S. Shenker, J. Wroclawski, "Paramètres généraux de caractérisation pour éléments de réseau à intégration de service", septembre 1997. *(P.S.)*

[RFC2747] F. Baker, B. Lindell, M. Talwar, "Authentification cryptographique RSVP", janvier 2000. (*MàJ par* RFC3097) *(P.S.)*

[RFC2815] M. Seaman et autres, "Transpositions de services intégrés sur réseaux IEEE 802", mai 2000. *(P.S.)*

[RFC2816] A. Ghanwani et autres, "Cadre pour les services intégrés sur technologies de LAN IEEE 802 partagés et commutés", mai 2000. *(Information)*

Appendice A

A.1 Introduction

Pour simplifier le reste de cet exposé, nous allons supposer qu'il y a un seul DSBM pour le domaine L2 tout entier (c'est-à-dire, on suppose un segment L2 partagé pour le domaine L2 entier). Plus tard, nous exposerons comment un DSBM est choisi pour chaque segment commuté mono ou bidirectionnel.

Pour permettre une récupération rapide des défaillances d'un DSBM, on suppose que des SBM supplémentaires peuvent être actifs dans un domaine L2 pour la tolérance aux fautes. Lorsque plus d'un SBM est actif dans un domaine L2, les SBM utilisent un algorithme d'élection pour désigner un DSBM pour le domaine L2. Après l'élection d'un DSBM et sa prise de fonction, les autres SBM restent passifs en arrière plan pour se manifester lors de l'élection d'un nouveau DSBM lorsque nécessaire. Le protocole pour l'élection et la découverte de DSBM est appelée le "protocole d'élection de DSBM" et est décrit dans la suite de cet Appendice.

A.1.1 Comment un client DSBM détecte un segment géré

Une fois élu, un DSBM envoie périodiquement en diffusion groupée un message I_AM_DSBM *(Je_suis_le_DSBM)* à l'adresse AllSBMAddress pour indiquer sa présence. Le message est envoyé à la périodicité définie (par exemple, toutes les 5 secondes) conformément à la valeur du temporisateur RefreshInterval (un paramètre de configuration). L'absence d'un tel message sur un certain intervalle (appelé "DSBMDeadInterval", un autre paramètre de configuration normalement réglé à un multiple de RefreshInterval) indique que le DSBM est défaillant ou arrêté et déclanche un nouveau tour d'élection de DSBM. Les clients DSBM écoutent toujours les annonces périodiques de DSBM. L'annonce comporte l'adresse IP d'envoi individuel du DSBM (DSBMAddress) et les clients DSBM envoient leurs messages PATH/RESV (ou autres) au DSBM. Lorsque un client DSBM détecte la défaillance d'un DSBM, il attend une annonce I_AM_DSBM suivante avant de reprendre la communication avec le DSBM. Durant la période pendant laquelle un DSBM n'est pas présent, un client DSBM peut transmettre les messages PATH sortants en utilisant les règles standard de transmission RSVP.

Les formats et adresses exacts de messages utilisés pour la communication avec (et entre) les SBM sont décrits dans l'Appendice B.

A.2 Survol de la procédure de choix de DSBM

Lors du premier démarrage d'un SBM, il écoute les annonces entrantes de DSBM pendant un certain temps pour vérifier si il existe déjà un DSBM dans son domaine L2. Si il en existe déjà un (et si une nouvelle élection n'est pas en cours) le nouveau SBM reste silencieux en arrière plan jusqu'à ce qu'une élection de DSBM soit nécessaire. Tous les messages qui se rapportent à l'élection du DSBM et aux annonces de DSBM sont toujours envoyées à AllSBMAddress.

Si il n'existe pas de DSBM, le SBM initie l'élection d'un DSBM par l'envoi d'un message DSBM_WILLING *(volontaire pour être DSBM)* qui indique son adresse IP comme celle d'un candidat DSBM et sa "priorité de SBM". Chaque SBM reçoit une priorité pour déterminer sa préséance relative. Lorsque il existe plus d'un SBM candidat, la valeur de priorité de SBM détermine qui doit être le DSBM sur la base de la priorité relative des candidats. Si il y a égalité des valeurs de priorité, le départage est fait à l'aide de l'adresse IP des candidate en cause (celui qui a l'adresse IP la plus élevée dans l'ordre lexicographique gagne). Les détails du protocole d'élection sont au paragraphe A.4.

A.2.1 Résumé de l'algorithme de choix

Pour les besoins de l'algorithme, un SBM est dans un des quatre états suivants : Idle, DetectDSBM, ElectDSBM, ou IAMDSBM.

Un SBM (qu'on appellera X) commence dans l'état DetectDSBM et attend pendant un intervalle ListenInterval un message entrant I_AM_DSBM (annonce de DSBM) ou DSBM_WILLING. Si une annonce I_AM_DSBM est reçue durant cet état, le SBM note le DSBM en cours (son adresse IP et sa priorité) et entre dans l'état Idle *(repos)*. Si un message DSBM_WILLING est reçu d'un autre SBM (qu'on appellera Y) durant cet état, X entre alors dans l'état ElectDSBM. Avant d'entrer dans le nouvel état, X vérifie d'abord pour voir si il est lui-même un meilleur candidat que Y et, si il en est ainsi, envoie un message DSBM_WILLING puis entre dans l'état ElectDSBM.

Lorsque un SBM (appelons le X) entre dans l'état ElectDSBM, il établit un temporisateur (appelé ElectionIntervalTimer, et est normalement réglé à une valeur au moins égale à la valeur de DSBMDeadInterval) et attend que l'élection se finisse pour découvrir qui est le meilleur candidat. Dans cet état, X garde en mémoire le meilleur candidat vu jusqu'à présent (y compris lui-même). Chaque fois qu'il reçoit un autre message DSBM_WILLING, il met à jour sa notion de "meilleur" candidat sur la base du critère de priorité (et de départage). Durant ElectionInterval, X envoie un message DSBM_WILLING tous les RefreshInterval pour (ré)affirmer sa candidature.

À la fin de ElectionInterval, X vérifie si il est le meilleur candidat jusqu'à présent. Si c'est le cas, il se déclare lui-même le DSBM (en envoyant l'annonce I_AM_DSBM) et entre dans l'état IAMDSBM ; autrement, il décide d'attendre que le meilleur candidat se déclare vainqueur. Pour attendre, X réinitialise son état ElectDSBM et continue d'attendre un autre tour d'élection (chaque tour dure un ElectionTimerInterval).

Un SBM est dans l'état Idle lorsque aucune élection n'est en cours et que le DSBM est déjà élu (et se trouve être quelqu'un d'autre). Dans cet état, il écoute les annonces I_AM_DSBM entrantes et utilise un DSBMDeadIntervalTimer pour détecter la défaillance du DSBM. Chaque fois que l'annonce est reçue, le temporisateur est relancé. Si le temporisateur arrive à expiration, le SBM passe dans l'état DetectDSBM pour préparer l'élection du nouveau DSBM. Si un SBM reçoit un message DSBM_WILLING du DSBM actuel dans cet état, le SBM entre dans l'état ElectDSBM après avoir envoyé un message DSBM_WILLING (pour annoncer sa propre candidature).

Dans l'état IAMDSBM, le DSBM envoie des annonces I_AM_DSBM tous les intervalles de rafraîchissement. si le DSBM souhaite fermer (terminaison en douceur) il envoie un message DSBM_WILLING (avec la valeur de priorité de SBM réglée à zéro) pour initier la procédure d'élection. La valeur de priorité de zéro retire effectivement le DSBM sortant de la procédure d'élection et laisse la place à l'élection d'un DSBM différent.

A.3 Récupération d'une défaillance de DSBM

Lorsque un DSBM a une défaillance (expiration de DSBMDeadIntervalTimer) tous les SBM entrent dans l'état ElectDSBM et commencent le processus d'élection.

À la fin de ElectionInterval, le DSBM élu envoie une annonce I_AM_DSBM et le DSBM est alors opérationnel.

A.4 Annonces DSBM

L'annonce I_AM_DSBM contient les informations suivantes :

1. Informations d'adresse du DSBM -- elles contiennent les adresses IP et L2 du DSBM et sa priorité de SBM (un paramètre de configuration -- priorité spécifiée par un administrateur de réseau). La valeur de priorité est utilisée pour choisir entre les SBM candidats durant l'algorithme d'élection. Les valeurs d'entier les plus élevées indiquent une priorité supérieure et la valeur est dans la gamme 0 à 255. La valeur zéro indique que le SBM n'est pas éligible pour être le DSBM. L'adresse IP est requise et est utilisée pour le départage. L'adresse L2 est pour l'interface du segment géré.

2. RegreshInterval – contient la valeur de RefreshInterval en secondes. La valeur zéro indique que le paramètre a été omis dans le message. Les receveurs peuvent y substituer leur propre valeur par défaut dans ce cas.

3. DSBMDeadInterval – contient la valeur de DSBMDeadInterval en secondes. Si la valeur est omise (ou si la valeur zéro est spécifiée) une valeur par défaut (de la configuration initiale) devrait être utilisée.

4. Diverses informations de configuration à annoncer aux envoyeurs sur le segment géré. Voir l'Appendice C.

A.5 Messages DSBM_WILLING

Lorsque un SBM souhaite déclarer sa candidature au poste de DSBM durant une phase d'élection, il envoie un message DSBM_WILLING. Le message DSBM_WILLING contient les informations suivantes :

1. Informations d'adresse de DSBM – contient les propres adresses du SBM (adresse IP et L2) si il souhaite être DSBM. L'adresse IP est requise et est utilisée pour le départage. L'adresse L2 est l'adresse de l'interface pour le segment géré en question. Aussi, les informations d'adresse de DSBM incluent la priorité correspondante du SBM dont l'adresse est données ci-dessus.

A.6 Variables d'état de SBM

Pour chaque interface réseau, un SBM entretient les variables d'état suivantes en rapport avec l'élection du DSBM pour le domaine L2 sur cette interface :

a) LocalDSBMAddrInfo -- L'adresse IP du DSBM en cours (initialement, 0.0.0.0) et sa priorité. Toutes les adresses IP sont supposées être dans l'ordre des octets du réseau. De plus, l'adresse L2 du DSBM en cours est aussi mémorisée au titre de ces informations d'état.

b) OwnAddrInfo – La propre adresse IP du SBM et l'adresse L2 pour l'interface et sa propre priorité (paramètre de configuration).

c) RefreshInterval en secondes. Lorsque le DSBM n'est pas encore élu, il est réglé à une valeur par défaut spécifiée comme un paramètre de configuration.

d) DSBMDeadInterval en secondes. Lorsque le DSBM n'est pas encore élu, il est réglé initialement à une valeur par défaut spécifiée comme un paramètre de configuration.

f) ListenInterval en secondes -- Un paramètre de configuration qui décide de la durée pendant laquelle un SBM attend dans l'état DetectDSBM (voir ci-dessous).

g) ElectionInterval en secondes -- Un paramètre de configuration qui décide du temps que passe un SBM dans l'état ElectDSBM lorsque il a déclaré sa candidature.

La Figure 3 donne les diagrammes de transition d'état pour le protocole d'élection et pour les divers états décrits ci-dessous. Une description complète de l'automate à états est fournie au paragraphe A.10.

A.7 États de choix de DSBM

DOWN – Le SBM n'est pas opérationnel.

DetectDSBM -- Normalement, c'est l'état initial d'un SBM lorsque il démarre. Dans cet état, il vérifie si il existe déjà un DSBM dans son domaine.

Idle – Le SBM est dans cet état lorsque aucune élection n'est en cours et qu'il n'est pas le DSBM. Dabs cet état, le SBM surveille passivement l'état du DSBM.

ElectDSBM – Le SBM est dans cet état lorsque une élection de DSBM est en cours.

IAMDSBM – Le SBM est dans cet état lorsque il est le DSBM pour le domaine L2.

A.8 Événements qui causent les changements d'état

StartUp – Le SBM commence à fonctionner.

Fin de temporisation ListenInterval – Le temporisateur ListenInterval est arrivé à expiration. Cela signifie que le SBM a surveillé son domaine pour voir si il existe un DSBM ou pour voir si il y a des candidats (autres que lui-même) volontaires pour être le DSBM.

Message DSBM_WILLING reçu – Cela signifie que le SBM a reçu un message DSBM_WILLING d'un autre SBM. Un tel message est envoyé lorsque un SBM souhaite déclarer sa candidature au poste de DSBM.

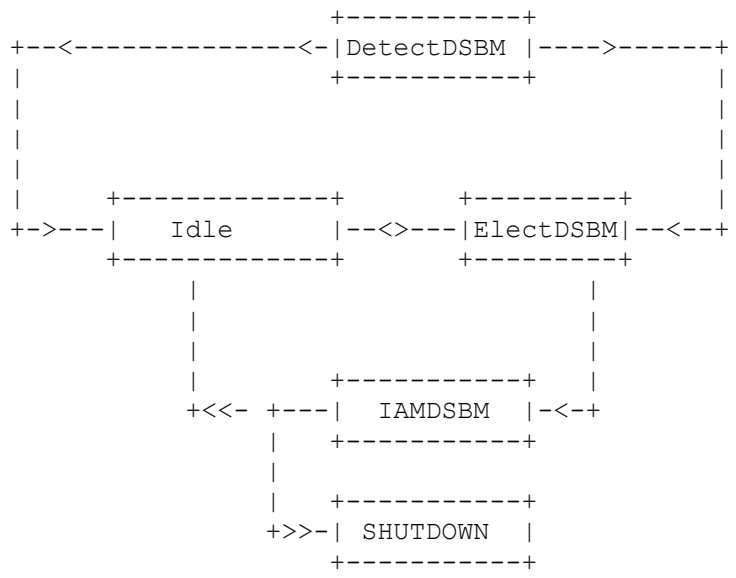
Message I_AM_DSBM reçu – Le SBM a reçu une annonce de DSBM du DSBM de son domaine L2.

Fin de la temporisation de DSBMDeadInterval – Le temporisateur DSBMDeadIntervalTimer est arrivé à expiration. Cela signifie que le SBM n'a pas reçu une seule annonce de DSBM durant cette période et cela indique une possible défaillance du DSBM.

Fin de la temporisation RefreshInterval – Le temporisateur RefreshIntervalTimer est arrivé à expiration. Dans l'état IAMDSBM, cela signifie qu'il est temps d'envoyer la prochaine annonce de DSBM. Dans l'état ElectDSBM, l'événement signifie qu'il est temps d'envoyer un autre message DSBM_WILLING.

Fin de la temporisation ElectionInterval – Le temporisateur ElectionIntervalTimer est arrivé à expiration. Cela signifie que le SBM a attendu assez longtemps après avoir déclaré sa candidature pour déterminer si elle a réussi ou non.

A.9 Diagramme de transition d'état (Figure 3)



A.10 Automate de choix d'état

Sur la base des événements et états décrits ci-dessus, l'état change au SBM comme décrit ci-dessous. Chaque changement d'état est déclanché par un événement et est normalement accompagné par une séquence d'actions. L'automate à états est décrit en supposant un seul fil de mise en œuvre (pour éviter une compétition entre les changements d'état et des événements de temporisation) et sans qu'il survienne de fin de temporisation durant l'exécution de l'automate de choix d'état.

Les sous-programmes suivants seront fréquemment utilisés dans la description de l'automate à états :

ComparePrio(FirstAddrInfo, SecondAddrInfo) – détermine si l'entité représentée par le premier paramètre est meilleure que la seconde entité en utilisant les informations de priorité et les informations d'adresse IP dans les deux paramètres. Si une adresse est zéro, cette entité perd automatiquement ; puis les premières priorités sont comparées; le candidat qui la priorité la plus élevée gagne. Si il y a égalité sur la base de la valeur de la priorité, on fait le départage en utilisant les adresses IP des candidats égaux (celui qui a la plus forte adresse IP dans l'ordre lexicographique gagne). Il retourne VRAI si la première entité est un meilleur choix. FAUX autrement.

```
SendDSBMWillingMessage()
Début
        Envoyer un message DSBM_WILLING me désignant comme candidat DSBM (copier OwnAddr et priorité dans
        les champs appropriés) lancer RefreshIntervalTimer
        passer à l'état ElectDSBM
Fin

AmIBetterDSBM(OtherAddrInfo)
Début
        si (ComparePrio(OwnAddrInfo, OtherAddrInfo))
        retourner VRAI
        changer LocalDSBMInfo = OtherDSBMAddrInfo
        retourner FAUX
Fin

UpdateDSBMInfo()
                                /* invoqué dans une déclaration telle que LocalDSBMInfo = OtherAddrInfo */
Début
        mettre à jour les LocalDSBMInfo telles que adresse IP, adresse L2 de DSBM,
        priorité de DSBM, RefreshIntervalTimer, DSBMDeadIntervalTimer
Fin
```

A.10.1 Changements d'état

Dans ce qui suit, l'action "continue" ou "continue dans l'état en cours" signifie une "sortie" de la séquence d'action en cours sans transition d'état.

```
État :          DOWN
Événement :     StartUp
Nouvel état :   DetectDSBM
Action :        Initialiser les variables d'état local (LocalDSBMADDR et LocalDSBMAddrInfo réglées à 0). Lancer le
                ListenIntervalTimer.

État :          DetectDSBM
Nouvel état :   Idle
Événement :     I_AM_DSBM message received
Action :        régler LocalDSBMAddrInfo = IncomingDSBMAddrInfo
                lancer le temporisateur DeadDSBMInterval
                passer à l'état Idle

État :          DetectDSBM
Événement :     ListenIntervalTimer est arrivé à expiration
Nouvel état :   ElectDSBM
Action :        Lancer ElectionIntervalTimer
                SendDSBMWillingMessage();
```

```
État :          DetectDSBM
Événement :     Le message DSBM_WILLING a été reçu
Nouvel état :   ElectDSBM
Action :        Annuler tout temporisateur en activité
                Lancer ElectionIntervalTimer
                                        /* suis-je un meilleur choix que ce gars là ? */
                Si (ComparePrio(OwnAddrInfo, IncomingDSBMInfo)) {
                                        /* Je suis meilleur */
                EnvoyerLeMessageDSBMWilling()
                } autrement {
                Changer LocalDSBMAddrInfo = IncomingDSBMAddrInfo
                passer à l'état ElectDSBM
                }

État :          Idle
Événement :     DSBMDeadIntervalTimer est arrivé à expiration.
Nouvel état :   ElectDSBM
Action :        lancer ElectionIntervalTimer
                régler LocalDSBMAddrInfo = OwnAddrInfo
                EnvoyerLeMessageDSBMWiliing()

État :          Idle
Événement :     Message I_AM_DSBM reçu.
Nouvel état :   Idle
Action :                                /* Vérifier d'abord si quelque chose a changé */
                si (!ComparePrio(LocalDSBMAddrInfo, IncomingDSBMAddrInfo))
                changer LocalDSBMAddrInfo pour refléter les nouvelles informations
                finDeSi
                relancer DSBMDeadIntervalTimer ;
                continuer dans l'état en cours ;

État :          Idle
Événement :     Le message DSBM_WILLING est reçu
Nouvel état :   Dépend de l'action (ElectDSBM ou Idle)
Action :                                /* vérifier si il est du DSBM lui-même (fermé) */
                si (IncomingDSBMAddr == LocalDSBMAddr) {
                  annuler les temporisateurs actifs
                  Régler LocalDSBMAddrInfo = OwnAddrInfo
                  Lancer ElectionIntervalTimer
                  EnvoyerLeMessageDSBMWilling()   /* passer à l'état ElectDSBM */
                }

                /* autrement, l'ignorer */
                continuer dans l'état en cours ;

État :          ElectDSBMÉvénement :    ElectionIntervalTimer est arrivé à expiration
Nouvel état :   dépend de l'action (état IAMDSBM ou en cours)
Action :        Si (LocalDSBMAddrInfo == OwnAddrInfo) {         /* J'ai gagné */
                envoyer le message I_AM_DSBM
                lancer RefreshIntervalTimer
                passer à l'état IAMDSBM
                } autrement {               /* quelqu'un d'autre a gagné, donc attendre qu'il se déclare comme DSBM */
                régler LocalDSBMAddressInfo = OwnAddrInfo
                lancer ElectionIntervalTimer
                EnvoyerLeMessageDSBMWilling()
                continuer dans l'état en cours ;
                }

État :          ElectDSBM
Événement :     Le message I_AM_DSBM a été reçu
Nouvel état :   Idle
Action :        régler LocalDSBMAddrInfo = IncomingDSBMAddrInfo
                Annuler tout temporisateur actif
```

```
                Lancer le temporisateur DeadDSBMInterval
                passer à l'état Idle

État :          ElectDSBM
Événement :     Le message DSBM_WILLING a été reçu
Nouvel état :   ElectDSBM
Action :        Vérifier si on n'est pas en boucle et si on l'est, éliminer, continuer ;
                si (!AmIBetterDSBM(IncomingDSBMAddrInfo)) {
                Changer LocalDSBMAddrInfo = IncomingDSBMAddrInfo
                Annuler RefreshIntervalTimer
                } autrement si (LocalDSBMAddrInfo == OwnAddrInfo) {
                EnvoyerLeMessageDSBMWilling()
                }
                continuer dans l'état en cours

État :          ElectDSBM
Événement :     RefreshIntervalTimer est arrivé à expiration
Nouvel état :   ElectDSBM
Action :                /* continuer d'envoyer des messages DSBMWilling jusqu'à la fin de l'intervalle d'élection */
                EnvoyerLeMessageDSBMWilling()

État :          IAMDSBM
Événement :     le message DSBM_WILLING a été reçu
Nouvel état :   dépend de l'action (état IAMDSBM ou permanent)
Action :                                        /* vérifier si quelqu'un d'autre est meilleur */
                Si (ComparePrio(OwnAddrInfo, IncomingAddrInfo))  {
                                                /* Je suis meilleur */
                Envoyer le message I_AM_DSBM
                relancer RefreshIntervalTimer
                continuer dans l'état en cours
                } autrement {
                Régler LocalDSBMAddrInfo = IncomingAddrInfo
                Annuler les temporisateurs actifs
                lancer DSBMDeadIntervalTimer
                passer à l'état permanent
                }

État :          IAMDSBM
Événement :     RefreshIntervalTimer est arrivé à expiration
Nouvel état :   IAMDSBM
Action :        envoyer le message I_AM_DSBM
                relancer RefreshIntervalTimer

État :          IAMDSBM
Événement :     message I_AM_DSBM reçu
Nouvel état :   dépend de l'action (IAMDSBM ou Idle)
Action :                                        /* vérifier si quelqu'un d'autre est meilleur */
                Si (ComparePrio(OwnAddrInfo, IncomingAddrInfo))  {
                                                /* Je suis meilleur */
                envoyer le message I_AM_DSBM
                relancer RefreshIntervalTimer
                continuer dans l'état en cours
                } autrement {
                Régler LocalDSBMAddrInfo = IncomingAddrInfo
                annuler les temporisateurs actifs
                lancer DSBMDeadIntervalTimer
                passer à l'état Idle
                }
```

État : IAMDSBM
Événement : Je veux fermer
Nouvel état : DOWN
Action : envoyer le message DSBM_WILLING avec mon adresse, mais la priorité réglée à zéro
passer à l'état Down

A.10.2 Valeurs suggérées pour les temporisateurs d'intervalles

Pour éviter des pannes de DSBM de longue durée, pour assurer une récupération rapide des défaillances du DSBM, et pour éviter la péremption de l'état PATH et RESV dans les appareils de bordures, on suggère les valeurs suivantes pour divers temporisateurs.

En supposant que les mises en œuvre de RSVP utilisent une temporisation de 30 secondes pour les rafraîchissements de PATH et RESV, on suggère que
RefreshIntervalTimer devrait être réglé à environ 5 secondes.
DSBMDeadIntervalTimer réglé à 15 secondes (K = 3, K*RefreshInterval).
DetectDSBMTimer devrait être réglé à une valeur aléatoire comprise entre (DSBMDeadIntervalTimer et 2*DSBMDeadIntervalTimer).
ElectionIntervalTimer devrait être réglé au moins à la valeur de DSBMDeadIntervalTimer pour s'assurer que chaque SBM a une chance d'avoir son message DSBM_WILLING (envoyé à chaque RefreshInterval dans l'état ElectDSBM) délivré aux autres.

A.10.3 Lignes directrices pour le choix des valeurs pour SBM_PRIORITY

Les administrateurs de réseau devraient configurer une entité de protocole SBM à chaque appareil à capacité SBM avec la "priorité de SBM" de l'appareil pour chacune des interfaces rattachées à un segment géré. SBM_PRIORITY est une valeur d'entier non signé de 8 bits (dans la gamme 0 à 255) les plus fortes valeurs d'entier notant les plus fortes priorités. La valeur zéro pour une interface indique que l'entité de protocole SBM sur l'appareil n'est pas éligible pour être un DSBM pour le segment rattaché à l'interface.

Une autre gamme de valeurs est réservée pour chaque type d'appareil à capacité SBM pour refléter la priorité relative entre les différentes classes d'appareils L2/L3. Les appareils L2 obtiennent des priorités plus élevées, suivis par les routeurs, suivis par les hôtes. Les valeurs de priorité dans la gamme de 128 à 255 sont réservées aux appareils L2, les valeurs dans la gamme de 64 à 127 sont réservées aux routeurs, et les valeurs dans la gamme de 1 à 63 sont réservées aux hôtes.

A.11 Choix de DSBM sur des liaisons commutées

L'algorithme d'élection fonctionne comme décrit auparavant dans ce cas, sauf que chaque appareil à capacité SBM restreint la portée de l'élection à son segment local. Comme décrit au paragraphe B.1 ci-dessous, tous les messages qui se rapportent à l'élection du DSBM sont envoyés à l'adresse spéciale de diffusion groupée (AllSBMAddress). AllSBMAddress (son adresse MAC de diffusion groupée correspondante) est configurée dans la base de données permanente des appareils L2 à capacité SBM, de sorte que toutes les trames avec AllSBMAddress comme adresse de destination ne sont pas transmises et sont plutôt dirigées sur l'entité de gestion SBM dans ces appareils. Donc, un DSBM peut être élu séparément sur chaque segment en point à point dans une topologie commutée. Par exemple, dans la Figure 2, un DSBM pour le "segment A" sera élu en utilisant l'algorithme d'élection entre R1 et S1 et aucun des messages en rapport avec l'élection ne sera transmis sur ce segment par S1 au delà du "segment A". De même, une élection distincte aura lieu sur chaque segment dans cette topologie.

Lorsque un segment commuté est un segment unidirectionnel, deux envoyeurs (un envoyeur à chaque extrémité de la liaison) partagent la liaison. Dans ce cas, un des deux envoyeurs va gagner l'élection de DSBM et sera chargé de la gestion du segment.

Si un segment commuté est bidirectionnel, exactement un envoyeur envoie sur la liaison dans chaque direction. Dans ce cas, un ou deux DSBM peuvent exister sur un tel segment géré. Si un envoyeur à chaque extrémité souhaite servir de DSBM pour cette extrémité, il peut se déclarer être le DSBM en envoyant une annonce I_AM_DSBM et commencer à gérer les ressources pour le trafic sortant sur le segment. Si un des deux envoyeurs ne souhaite pas être lui-même le DSBM, l'autre DSBM ne recevra alors aucune annonce de DSBM de son homologue et supposera qu'il est lui-même le DSBM pour le trafic qui traverse dans les deux directions sur le segment géré.

Appendice B Encapsulation et formats de message

Pour minimiser les changements aux mises en œuvre RSVP existantes et assurer un déploiement rapide d'un SBM en conjonction avec RSVP, toutes les communications de et avec un DSBM seront effectuées en utilisant des messages construits selon les règles actuelles pour les formats de message RSVP et l'encapsulation IP brute. Pour plus de détails sur les formats de message RSVP, se référer à la spécification RSVP [RFC2205]. Aucun changement au format de message RSVP n'est proposé, mais de nouveaux types de message et de nouveaux objets spécifiques de L2 sont ajoutés aux formats de message RSVP pour s'accommoder des messages qui se rapportent au DSBM. Ces ajouts sont décrits ci-après.

B.1 Adressage de message

Pour les besoins de l'élection et de la détection du DSBM, AllSBMAddress est utilisé comme adresse de destination lors de l'envoi des messages aussi bien DSBM_WILLING que I_AM_DSBM. Un client DSBM détecte d'abord un segment géré en écoutant les annonces I_AM_DSBM et les enregistrements de DSBMAddress (adresse IP en envoi individuel du DSBM).

B.2 Tailles de message

Chaque message doit occuper exactement un datagramme IP. Si il excède la MTU, un tel datagramme sera fragmenté par IP et réassemblé au nœud receveur. Cela a pour conséquence qu'un seul message ne peut pas excéder la taille maximum de datagramme IP, approximativement de 64 koctets.

B.3 Formats des messages en rapport avec RSVP

Tous les messages RSVP dirigés de et vers un DSBM peuvent contenir divers objets RSVP définis dans la spécification RSVP et les messages continuent de suivre les règles de formatage de la spécification RSVP. De plus, une mise en œuvre de RSVP doit aussi reconnaître les nouvelles classes d'objets qui sont décrites ci-dessous.

B.3.1 Formats d'objet

Tous les objets sont définis en utilisant le format spécifié dans RSVP. Chaque objet a un en-tête de 32 bits qui contient la longueur (de l'objet en octets incluant l'en-tête de l'objet) le numéro de classe de l'objet, et un C-Type. Tous les champs inutilisés devraient être réglés à zéro et ignorés à réception.

B.3.2 Objets spécifiques de SBM

Noter que les valeurs de Class-Num pour les objets spécifiques de SBM (LAN_NHOP, LAN_LOOPBACK, et RSVP_HOP_L2) sont choisis dans l'espace de codes 10XXXXXX. Ce codage assure que les nœuds RSVP sans capacité de SBM vont ignorer les objets sans les transmettre ni générer de message d'erreur.

Au sein de l'espace de code spécifique de SBM, noter l'interprétation suivante du troisième bit de plus fort poids dans le Class-Num :

a) Les objets de la forme 100XXXXX sont à éliminer en silence par les nœuds SBM qui ne les reconnaissent pas.
b) Les objets de la forme 101XXXXX sont à transmettre en silence par les nœuds SBM qui ne les reconnaissent pas.

B.3.3 Format d'adresse canonique IEEE 802

Les adresses MAC de 48 bits utilisées par IEEE 802 étaient à l'origine définies en termes de transmission dans l'ordre du réseau des bits dans les champs Adresse MAC de source et de destination. Le même ordre des octets du réseau s'applique aussi bien à Ethernet qu'aux anneaux à jetons. Comme l'ordre de transmission des bits de l'Ethernet et de l'anneau à jetons diffère – les octets Ethernet sont transmis avec le bit de moindre poids en premier, ceux de l'anneau à jeton avec le bit de poids fort en premier – les valeurs numériques associées naturellement à la même adresse sur des supports 802 différents diffèrent. Pour faciliter la communication des valeurs d'adresse dans les protocoles de couche supérieure qui peuvent s'étendre sur des systèmes d'anneau à jetons et d'Ethernet rattachés connectés par des ponts, il était nécessaire de définir un format de référence – ce qu'on appelle le format canonique de ces adresses. Formellement, le format canonique définit la valeur de l'adresse, distinct des règles de codage utilisées pour la transmission. Il comporte une séquence d'octets déduite comme suit de l'ordre original de transmission des bits dans l'ordre du réseau. Le bit de moindre poids du premier octet est le premier bit transmis, le bit de moindre poids suivant est le second bit, et ainsi de suite jusqu'au bit de plus fort poids du premier octet qui est le huitième bit transmis ; le bit de moindre poids du second octet est le neuvième bit transmis, et ainsi de suite jusqu'au bit de plus fort poids du sixième octet du format canonique qui est le dernier bit de l'adresse transmise.

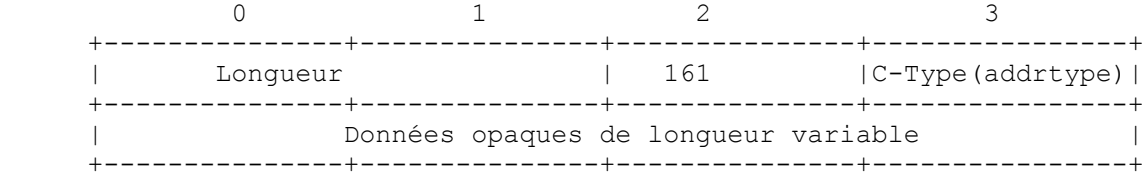
Ce format canonique correspond à la valeur naturelle des octets d'adresse pour Ethernet. L'ordre de transmission réel, ou les règles formelles de codage pour les adresses sur des supports qui ne transmettent pas les bits en série, est déduit des valeurs d'octet au format canonique.

Le présent document exige que toutes les adresses L2 utilisées en conjonction avec le protocole SBM soient codées dans le format canonique comme une séquence de 6 octets. Dans ce qui suit, on définit les formats d'objet pour les objets qui contiennent des adresses L2 qui se fondent sur la représentation canonique.

B.3.4 Objet RSVP_HOP_L2

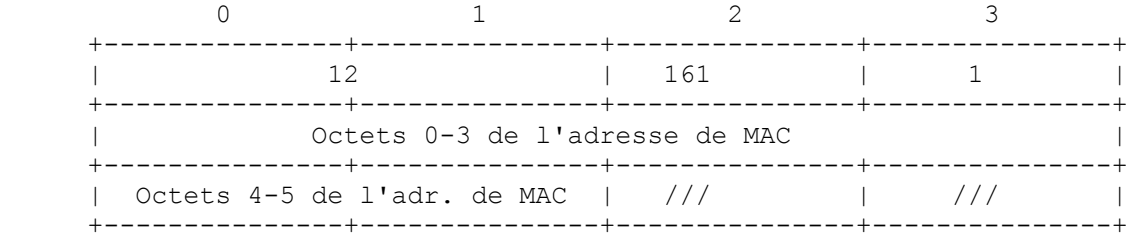
L'objet RSVP_HOP_L2 utilise la classe d'objet = 161; il contient l'adresse L2 de l'appareil L3 du bond précédent sous le format d'adresse canonique IEEE exposé ci-dessus.

Objet RSVP_HOP_L2 : classe = 161, le C-Type représente le format d'adressage utilisé. Dans notre cas, C-Type=1 représente le format d'adresse canonique IEEE.



C-Type = 1 (format d'adresse canonique IEEE)

Lorsque C-Type = 1, le format de l'objet est :



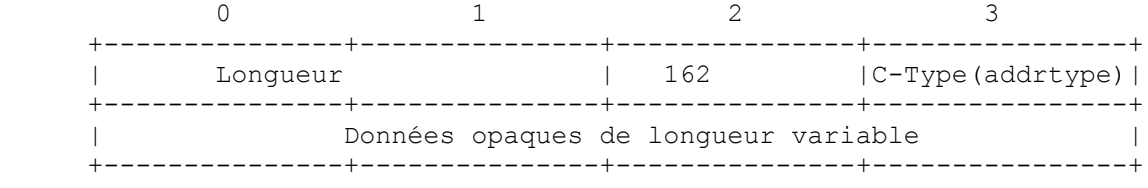
/// -- non utilisés (mis à zéro)

B.3.5 Objet LAN_NHOP

L'objet LAN_NHOP représente deux objets, à savoir, l'objet d'adresse LAN_NHOP_L3 et l'objet d'adresse LAN_NHOP_L2.

<objet LAN_NHOP> ::= <objet LAN_NHOP_L2> <objet LAN_NHOP_L3>

L'objet d'adresse LAN_NHOP_L2 utilise la classe d'objet = 162 et utilise le même format (mais un numéro de classe différent) que l'objet RSVP_HOP_L2. Il fournit l'adresse L2 MAC de l'appareil L3 de prochain bond.

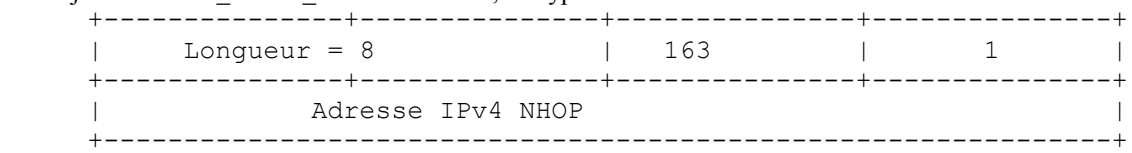


C-Type = 1 (format d'adresse canonique IEEE 802 comme défini ci-dessous) Voir l'objet d'adresse RSVP_HOP_L2 pour d'autres détails.

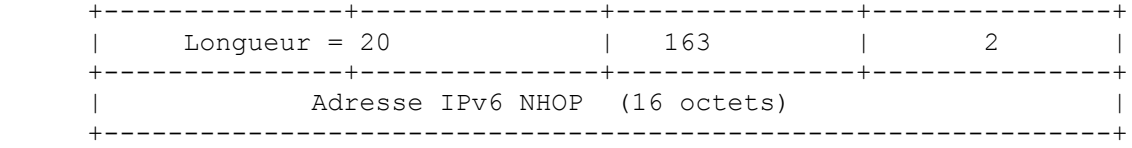
L'objet LAN_NHOP_L3 utilise la classe d'objet = 163 et donne l'adresse IP ou L3 de l'appareil L3 de prochain bond.

Objet LAN_NHOP_L3 : classe = 163, le C-Type spécifie la famille d'adresse s IPv4 ou IPv6 utilisée.

Objet IPv4 LAN_NHOP_L3 : classe =163, C-Type = 1



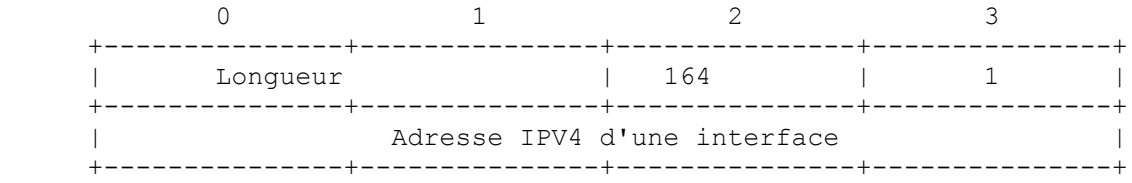
Objet IPv6 LAN_NHOP_L3 : classe =163, C-Type = 2



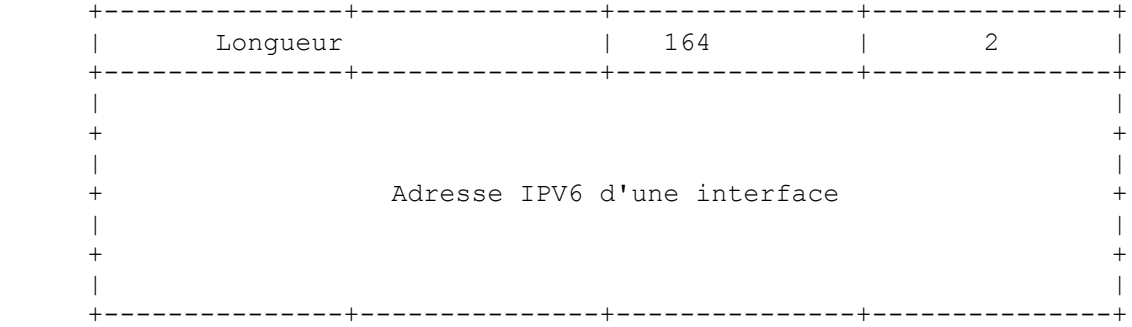
B.3.6 Objet LAN_LOOPBACK

L'objet LAN_LOOPBACK donne l'adresse IP de l'interface sortante pour le message PATH et utilise la classe d'objet = 164; les deux formats IPv4 et IPv6 sont spécifiés.

Objet IPv4 LAN_LOOPBACK : classe = 164, C-Type = 1

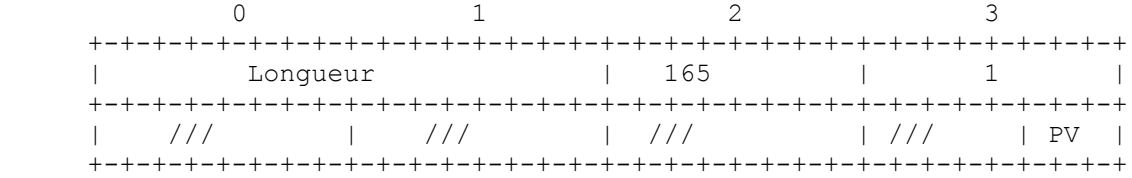


Objet IPv6 LAN_LOOPBACK : classe = 164, C-Type = 2



B.3.7 Objet TCLASS

L'objet TCLASS (classe de trafic fondée sur IEEE 802.1p) utilise la classe d'objet = 165.



Seuls 3 bits de données contiennent la valeur user_priority (PV).

B.4 Formats de message RSVP PATH et PATH_TEAR

Comme spécifié dans RSVP, les messages PATH et PATH_TEAR contiennent l'en-tête RSVP commun et les objets RSVP pertinents.

Pour l'en-tête RSVP commun, se référer à la spécification RSVP [RFC2205]. Les améliorations à un message PATH RSVP comportent les objets supplémentaires spécifiés ci-dessous.

<Message PATH> ::= <En-tête RSVP commun> [<INTEGRITY>]
<RSVP_HOP_L2> <LAN_NHOP>
<LAN_LOOPBACK> [<TCLASS>] <SESSION><RSVP_HOP>
<TIME_VALUES> [<POLICY DATA>] <descripteur de l'envoyeur>

<Message PATH_TEAR> ::= <En-tête RSVP commun> [<INTEGRITY>]
<LAN_LOOPBACK> <LAN_NHOP> <SESSION> <RSVP_HOP> [<descripteur de l'envoyeur>]

Si l'objet INTEGRITY est présent, il doit suivre immédiatement l'en-tête RSVP commun. Les objets spécifiques de L2 doivent toujours précéder l'objet SESSION.

B.5 Format de message RESV RSVP

Comme spécifié dans RSVP, un message RESV RSVP contient l'en-tête RSVP commun et les objets RSVP pertinents. De plus, il peut contenir un objet TCLASS facultatif comme décrit précédemment.

B.6 Types supplémentaires de message RSVP pour traiter les interactions SBM

De nouveaux types de message RSVP sont introduits pour permettre les interactions entre un DSBM et un nœud RSVP (hôte/routeur) pour les besoins de la découverte et de la liaison avec un DSBM. Les nouveaux types de message RSVP nécessaires sont les suivants :

Type de message RSVP (8 bits)	**Valeur**
DSBM_WILLING	66
I_AM_DSBM	67

Tous les messages spécifiques de SBM sont formatés comme des messages RSVP avec un en-tête RSVP commun suivi par des objets spécifiques de SBM.

<SBMP_MESSAGE> ::= <En-tête SBMP commun> <objets spécifiques de SBM>

où <En-tête SBMP commun> ::= <En-tête RSVP commun> [<INTEGRITY>]

Pour chaque type de message SBM, il y a un ensemble de règles pour les choix de types d'objet permis. Ces règles sont spécifiées en utilisant le format Backus-Naur augmenté (ABNF) avec des crochets angulaires qui entourent les sous séquences facultatives. L'ABNF implique un ordre des objets dans un message. Cependant, dans de nombreux cas (mais pas tous) l'ordre des objets ne fait pas de différence logique. Une mise en œuvre devrait créer des messages avec les objets dans l'ordre indiqué ici, mais accepter les objets dans tout ordre permissible. Toutes les exceptions à cette règle seront mentionnées dans les formats des message spécifiques.

Message DSBM_WILLING

<message DSBM_WILLING> ::= <En-tête SBM commun> <ADRESSE IP DSBM>
<Adresse L2 DSBM> <PRIORITÉ DE SBM>

Message I_AM_DSBM

<I_AM_DSBM> ::= <En-tête SBM commun> <ADRESSE IP DSBM> <Adresse L2 DSBM>
<Adresse L2 DSBM> <DSBM Timer Intervals> [<NON_RESV_SEND_LIMIT>]

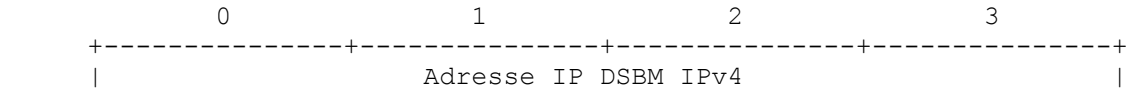
Pour des raisons de compatibilité, les receveurs du message I_AM_DSBM doivent être prêts à recevoir des objets supplémentaires du type de classe Inconnue [RFC2205].

Tous les messages I_AM_DSBM sont en diffusion groupée à l'adresse bien connue AllSBMAddress. La priorité par défaut d'un SBM est 1 et les valeurs de priorité supérieures représentent une préséance plus élevée. La valeur de priorité zéro indique que le SBM n'est pas éligible au poste de DSBM.

Objets pertinents

Les objets ADRESSE IP DSBM utilisent la classe d'objet = 42 ; l'objet IPv4 ADRESSE IP DSBM utilise <Classe = 42, le C-Type = 1> et l'objet IPv6 ADRESSE IP DSBM utilise la <Classe = 42, et le C-Type = 2>.

Objet IPv4 ADRESSE IP DSBM : classe = 42, C-Type = 1



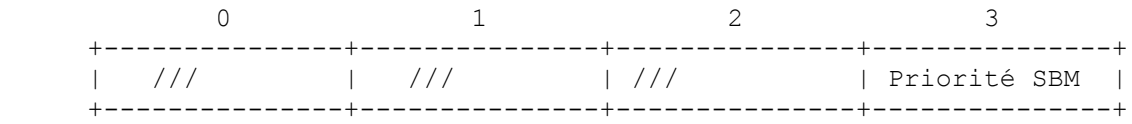
Objet IPv6 DSBM IP ADDRESS : Classe = 42, C-Type = 2

L'objet <Adresse L2 DSBM> est le même que l'objet <RSVP_HOP_L2> avec le C-Type = 1 pour le format canonique d'adresse IEEE.

<Adresse L2 DSBM> ::= <RSVP_HOP_L2>

Un SBM peut omettre cet objet en incluant un objet Adresse L2 NUL. Pour le C-Type = 1 (format canonique d'adresse IEEE) une telle version de l'objet Adresse L2 contient la valeur zéro dans les six octets qui correspondent à l'adresse MAC (voir au paragraphe B.3.4 le format exact).

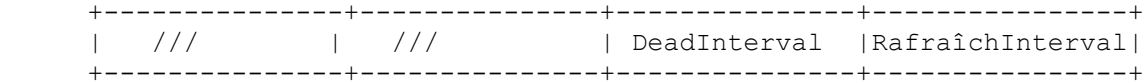
Objet SBM_PRIORITY : classe = 43, C-Type = 1



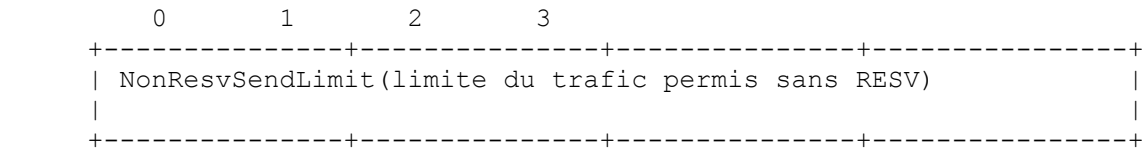
Valeurs d'intervalle de temporisateur.

Les deux intervalles de temporisateur, à savoir, Intervalle de DSBM mort et Intervalle de rafraîchissement de DSBM, sont spécifiés comme des valeurs d'entier chacune dans la gamme de 0 à 255 secondes. Les deux valeurs sont incluses dans un seul objet "Intervalles de temporisateur DSBM" décrit ci-dessous.

Objet Intervalles de temporisateur DSBM : classe = 44, C-Type =1



Objet NON_RESV_SEND_LIMIT : classe = 45, C-Type = 1



<NonResvSendLimit> ::= <Objet Intserv Sender_TSPEC> (classe = 12, C-Type =e2)

L'objet NON_RESV_SEND_LIMIT spécifie une limite par flux au profil de trafic qu'il est permis à un hôte d'envoyer sur un segment géré sans une réservation RSVP valide (voir les détails de l'usage de cet objet à l'Appendice C). L'objet contient le paramètre NonResvSendLimit. Ce paramètre est équivalent au SENDER_TSPEC de Intserv (voir les règles de contenu et de codage dans la [RFC2210]). La SENDER_TSPEC comporte cinq paramètres qui décrivent un profil de trafic (r, b, p, m et M). Les hôtes envoyeurs comparent la SENDER_TSPEC qui décrit le flux de trafic d'un envoyeur à la

SENDER_TSPEC annoncée par le DSBM. Si la SENDER_TSPEC du flux de trafic en question est inférieure ou égale à la SENDER_TSPEC annoncée par le DSBM, il est permis d'envoyer du trafic sur le flux correspondant sans une réservation RSVP valide installée. Autrement, ce n'est pas permis.

L'administrateur de réseau peut configurer le DSBM de façon à refuser tout trafic envoyé en l'absence d'une réservation RSVP en configurant une NonResvSendLimit dans laquelle r = 0, b = 0, p = 0, m = infini et M = 0. De même, l'administrateur de réseau peut permettre tout envoi de trafic en l'absence d'une réservation RSVP en configurant une NonResvSendLimit dans laquelle r = infini, b = infini, p = infini, m = 0 et M = infini. Bien sûr, tous ces paramètres peuvent être réglés à des valeurs entre zéro et l'infini pour annoncer des limites par flux finies.

L'objet NON_RESV_SEND_LIMIT est facultatif. Les envoyeurs sur un segment géré devraient interpréter l'absence de l'objet NON_RESV_SEND_LIMIT comme équivalent à une SENDER_TSPEC infiniment grande (il est permis d'envoyer tout profil de trafic en l'absence d'une réservation RSVP).

Appendice C DSBM comme source centralisée d'informations de configuration

Il y a certains paramètres de configuration qu'il peut être utile de distribuer aux envoyeurs de couche 3 sur un segment géré. Le DSBM peut servir de point de gestion centralisé à partir duquel de tels paramètres peuvent facilement être distribués. En particulier, il est possible à l'administrateur de réseau qui configure un DSBM de faire que certains paramètres de configuration soient distribués comme des objets ajoutés aux messages I_AM_DSBM. L'objet de configuration suivant est défini à ce moment. D'autres pourront être définis à l'avenir. Voir à l'Appendice B des détails supplémentaires concernant l'objet NON_RESV_SEND_LIMIT.

C.1 NON_RESV_SEND_LIMIT

Lorsque on active pour la qualité de service des segments de couche 2, on s'attend à une évolution de la part des sous-réseaux composés de segments partagés traditionnels (sans moyen de séparation du trafic et sans DSBM) en sous-réseaux comportant des segments dédiés commutés par des commutateurs sophistiqués (avec à la fois un DSBM et des capacités de séparation de trafic 802.1p).

Un ensemble de configurations intermédiaires consiste en un groupe d'hôtes à capacité de qualité de service qui envoient sur un segment partagé traditionnel. Un appareil de couche 3 (ou un appareil de couche 2) agit comme DSBM pour le segment partagé, mais ne peut pas mettre en application la séparation des trafics. Dans une telle configuration, le DSBM peut être configuré de façon à limiter le nombre de réservations approuvées pour les envoyeurs sur le segment, mais ne peut pas les empêcher d'envoyer. Il en résulte que les envoyeurs peuvent encombrer le segment même si un administrateur de réseau a configuré une limite appropriée pour le contrôle d'admission dans le DSBM.

Une solution à ce problème qui pourrait donner à l'administrateur de réseau le contrôle sur le segment, est d'exiger des applications (ou des systèmes d'exploitation au nom des applications) de ne pas envoyer jusqu'à ce qu'ils aient obtenu une réservation. Ceci est problématique car la plupart des applications sont utilisées pour envoyer aussitôt qu'elles souhaitent le faire et s'attendent à avoir la qualité de service que le réseau est capable de leur accorder à ce moment là. De plus, il peut être souvent acceptable de permettre à certaines applications d'envoyer avant qu'une réservation soit reçue. Par exemple, sur un segment comportant un seul Ethernet à 10 Mbit/s et dix hôtes, il peut être acceptable de permettre un flux de téléphonie à 16 kbit/s mais pas un flux vidéo à 3 Mbit/s.

Une solution plus pragmatique est alors de permettre à l'administrateur de réseau de régler une limite par flux sur la quantité de trafic non adaptatif qu'il est permis à un envoyeur de générer sur un segment géré en l'absence d'une réservation valide. Cette limite est annoncée par le DSBM et est reçue par les hôtes qui envoient. Une API chez l'hôte envoyeur peut alors approuver ou refuser la demande de qualité de service d'une application sur la base des ressources demandées.

L'objet NON_RESV_SEND_LIMIT peut être utilisé pour annoncer une Flowspec qui décrit la forme du trafic qu'il est permis à un envoyeur de générer sur un segment géré lorsque ses demandes de réservation RSVP ne sont pas encore terminées ou ont été rejetées.

10. Remerciements

Les auteurs témoignent de leur gratitude à l'égard de Eric Crawley (Argon), Russ Fenger (Intel), David Melman

(Siemens), Ramesh Pabbati (Microsoft), Mick Seaman (3COM), Andrew Smith (Extreme Networks) pour leurs commentaires constructifs sur le concept de SBM et sur les précédentes versions de ce document.

11. Adresse des auteurs

Raj Yavatkar
Intel Corporation
2111 N.E. 25th Avenue,
Hillsboro, OR 97124
USA
téléphone : +1 503-264-9077
mél : yavatkar@ibeam.intel.com

Don Hoffman
Teledesic Corporation
2300 Carillon Point
Kirkland, WA 98033
USA
téléphone : +1 425-602-0000

Yoram Bernet
Microsoft
1 Microsoft Way
Redmond, WA 98052
USA
téléphone : +1 206 936 9568
mél : yoramb@microsoft.com

Fred Baker
Cisco Systems
519 Lado Drive
Santa Barbara, California 93111
USA
téléphone : +1 408 526 4257
mél : fred@cisco.com

Michael Speer
Sun Microsystems, Inc
901 San Antonio Road UMPK15-215
Palo Alto, CA 94303
USA
téléphone : +1 650-786-6368
mél : speer@Eng.Sun.COM

12. Déclaration complète de droits de reproduction

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