Date: 21 Juillet 2004

Chapitre du livre « IPv6 : Théorie et Pratique », Quatrième édition

Multicast IPv6

Version 13

Editeur :

Jérôme Durand (Renater)
Jerome.Durand@renater.fr

Auteurs :

Bernard Cousin (IRISA, INRIA, Rennes)
Bernard.Cousin@irisa.fr

Mickael Hoerdt
hoerdt@clarinet.u-strasbg.fr

Konstantin Kabassanov (LIP6, Paris)
Konstantin.Kabassanov@lip6.fr

Christophe Janneteau
Christophe.Janneteau@motorola.com

Emmanuel Riou
Emmanuel.Riou@motorola.com

Imed Romdhani
Imed.Romdhani@motorola.com

Table des Matières

1 Introduction

Une communication multicast est une communication dans laquelle un même paquet de données peut être envoyé à un groupe de récepteurs, quelque soit leur localisation. Dans le modèle Internet IPv6, une station peut potentiellement émettre un paquet multicast vers n'importe quel groupe. Comparé aux communications *point à point* (unicast), le multicast évite la duplication des paquets de données au niveau de la source, et minimise l'utilisation de la bande passante au niveau du réseau. De plus, il offre un service insensible à l'augmentation du nombre et la localisation des membres d'un groupe. Le multicast peut être utilisé pour la distribution de logiciels, la téléconférence, les applications d'enseignement à distance, la radio ou la télévision sur Internet, les simulations interactives distribuées, les jeux multimédia interactifs, les applications militaires, etc. Ce chapitre insistera sur les différences par rapport au multicast IPv4, tout en donnant une vue d'ensemble des protocoles mis en jeu.

Pour le multicast, on distingue deux modèles de communication : le modèle ASM (*Any-Source Multicast*) et le modèle SSM (*Source-Specific Multicast*). Dans le modèle ASM, un récepteur s'abonne à un groupe, et reçoit les données émises par n'importe quelles sources pour ce groupe. Ce modèle s'applique par exemple dans le cas de visioconférences avec de nombreux participants qui ne sont pas connus à l'avance. Dans le modèle SSM, les sources sont connues à l'avance et les récepteurs s'abonnent à un groupe et un ensemble de sources. Ce modèle s'applique par exemple à la diffusion de la télévision ou radio sur Internet, où il n'y a qu'une seule source connue de tous.

Les étapes suivantes interviennent dans l'établissement d'une session multicast IPv6 :

- ***Choix de l'adresse multicast pour la session :*** L'architecture de l'adressage multicast IPv6 est décrite dans la section 2. Dans cette section sont aussi présentés les mécanismes permettant l'allocation des adresses multicast.
- ***Description et annonce de la session multicast à tous les participants :*** La fin de la section 2 en donne un bref aperçu.
- ***Gestion des membres du groupe sur le lien-local :*** Elle est réalisée par le protocole MLD (*Multicast Listener Discovery*), détaillé dans la section 3
- ***Construction de l'arbre multicast :*** Elle est assurée par le protocole PIM *(Protocol Independant Multicast)* défini dans la section 4.

La cinquième section de ce chapitre est consacrée au multicast interdomaine IPv6. L'état actuel du déploiement du multicast IPv6 est ensuite sommairement décrit. Les deux sections suivantes portent sur les applications multicast IPv6 et la coexistence avec le multicast IPv4. La dernière section présente un cas pratique.

2 Adressage multicast

Pour initier une session multicast, le groupe de récepteurs intéressés, appelé aussi groupe multicast, doit être formé. Un groupe multicast est identifié par une adresse IP multicast. Chaque adresse a une portée spécifique, qui limite la propagation du trafic multicast.

Dans ce chapitre, nous commençons par détailler le format des adresses multicast IPv6 Nous examinons ensuite successivement l'allocation des adresses multicast IPv6 puis l'annonce des sessions.

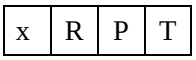
2.1 Format des adresses multicast IPv6

2.1.1 Généralités

Cette section décrit le système d'adressage multicast IPv6. Le RFC 3513 [27] (IP Version 6 Addressing Architecture) définit une adresse IPv6 multicast de la façon suivante :

Structure de l'adresse IPv6 multicast

Les adresses multicast IPv6 sont dérivées du préfixe FF00::/8. Le champ *flags* de 4 bits est défini de la manière suivante :



Le champ flags

Seul le bit T (comme Transient) du champ *flags* est décrit dans le RFC 3513. Les bits P et R sont décrits dans le RFC3306 [15] et l'Internet draft sur embedded-RP [24]. Le bit de poids fort du champ *flags* n'est pas encore attribué. Le champ *flags* permet de définir plusieurs types d'adresses multicast IPv6 qui seront décrits dans les sections suivantes.

Les champs *scope* et *group-ID* sont également détaillés plus tard.

2.1.2 Adresses multicast IPv6 permanentes

Le RFC 3513 indique qu'une adresse multicast IPv6 avec le bit T du champ *flags* à 0 correspond à une adresse multicast permanente, allouée par l'IANA (Internet Assigned Number Authority).

Structure des adresses IPv6 multicast permanentes

Quand le multicast IPv6 sera déployé à grande échelle, certains organismes pourraient avoir des émissions permanentes. Des chaînes de télévision ou stations de radio pourront par exemple se voir attribuer des adresses permanentes par l'IANA dans le préfixe FF00::/12.

Le RFC 2375 [26] définit déjà certaines adresses IPv6 multicast. Deux types d'adresses multicast permanentes sont à distinguer : des adresses correspondant à des services de niveau réseau (comme NTP, DHCP, cisco-rp-announce, SAP, …) et des adresses correspondant d'avantage à des services applicatifs commerciaux permanents comme la distribution des chaînes de télévision. Le RFC 3307 [14] définit des procédures pour l'allocation des adresses multicast permanentes. Celles-ci seront décrites par la suite.

2.1.3 Adresses temporaires

Les adresses temporaires sont des adresses multicast IPv6 dont le bit T est positionné à 1. Il existe plusieurs types d'adresses temporaires : celles qui sont générales, celles dérivées d'un préfixe unicast, les adresses multicast "Embedded-RP" et les adresse SSM.

2.1.3.1 Adresses temporaires générales

Ce sont des adresses avec tous les bits du champ flag à 0 sauf le bit T positionné à 1.Il semble qu'il n'y ait pas vraiment de recommandations pour l'utilisation de ces adresses. Des scénarios d'utilisation peuvent être par exemple les visioconférences ponctuelles.

2.1.3.2 Adresses multicast dérivées d'un préfixe unicast IPv6

Le RFC 3306 (Unicast-Prefix-based IPv6 Multicast Addresses) [15] définit une méthode pour dériver une adresse multicast IPv6 à partir d'un préfixe unicast :

FF	x011	scope	res	Plen	prefix	group-ID
8 bits	4 bits	4 bits	8 bits	8 bits	64 bits	32 bits

Structure d'une adresse multicast IPv6 dérivée d'un préfixe unicast

- *res* (reserved) : tous les bits de ce champ doivent être positionnés à 0.
- *Plen* (prefix length) : ce champ contient la longueur du préfixe unicast utilisé pour en dériver une adresse multicast.
- *prefix* : ce champ contient la valeur du préfixe du réseau utilisé pour en dériver une adresse multicast.
- *group-ID* : ce champ de 32 bits contient l'identifiant de groupe, détaillé dans la section 2.1.6.

Voici un exemple de dérivation d'une adresse multicast à partir du préfixe de RENATER (2001:660::/32). Le champ *prefix* a la valeur 2001:0660:0000:0000 et le champ *Plen* a la valeur 0x20 (32 en décimal). Les adresses multicast IPv6 à choisir seront de type FF3X:20:2001:660::aabb:ccdd (aabb:ccdd étant le *group-ID* choisi dans l'exemple).

Cette méthode permet la création potentielle de 2^{32} adresses par préfixe.

2.1.3.3 Adresses multicast "Embedded-RP"

Embedded-RP [24] définit une méthode pour inclure l'adresse du RP (Point de Rendez-Vous qui sert à la construction de l'arbre multicast) dans l'adresse multicast IPv6. Le schéma suivant montre la structure d'une telle adresse, aussi appelée adresse "embedded-RP" :

FF	x111	scope	res	RPad	Plen	prefix	Group-ID
8 bits	4 bits	4 bits	4 bits	4 bits	8 bits	64 bits	32 bits

Structure d'un adresse multicast IPv6 "embedded RP"

L'exemple suivant décrit la méthode proposée dans [24]. Pour un RP qui possède l'adresse 2001:660:3307:125::3, une adresse multicast correspondante peut être dérivée de la façon suivante :

- *res* (Reservé) : Les 4 bits de ce champ sont positionnés à 0.
- *RPad* : Ce champ contient les 4 derniers bits de l'adresse du RP. Dans cet exemple, *RPad* prend la valeur 3.
- *Plen* (Longueur du préfixe) : Ce champ contient la longueur du préfixe réseau du RP à prendre en compte. Dans cet exemple, la valeur est de 0x40 (soit 64 en décimal).
- *prefix* (Préfixe) : Ce champ contient le préfixe réseau du RP. Ici, cette valeur est 2001:660:3007:125
- *group-ID* : ce champ de 32 bits contient l'identifiant de groupe, détaillé dans la section 2.1.6.

Une adresse multicast dérivée de ce RP sera donc de la forme FF7X:340:2001:660:3007:125:aabb:ccdd (aabb:ccdd étant le *group-ID* choisi dans cet exemple).

2.1.3.4 Adresses SSM

Les adresses SSM (Source Specific Multicast) sont décrites également dans le RFC 3306. Si le préfixe FF3X::/32 a été réservé pour les adresses multicast SSM, seules les adresses dérivées du préfixe FF3X::/96 doivent être utilisées dans un premier temps. Ce sont des adresses multicast basées sur le préfixe unicast où les champs *Plen* et *prefix* sont positionnés à 0.

FF	x011	scope	zeros	Group-ID
8 bits	4 bits	4 bits	80 bits	32 bits

Structure d'une adresse multicast IPv6 SSM

2.1.4 Portée des adresses multicast

Le champ *scope* de l'adresse multicast IPv6 permet d'en limiter la portée (scope en anglais). Les valeurs suivantes sont définies :

1 - node-local
2 - link-local
3 - subnet-local
4 - admin-local
5 - site-local
8 - organisation-local
E - global

Les portées 0 et F sont réservées.

En IPv4 la portée d'un paquet est limitée par le champ TTL (Time To Live), de même des préfixes peuvent être définis pour identifier des adresses à portée réduite.

2.1.5 Récapitulatif des types d'adresses multicast

Préfixe	Usage
FF0X::/16	Adresses IPv6 multicast permanentes
FF1X::/16	Adresses IPv6 multicast temporaires générales
FF3X::/16	Adresses multicast dérivées d'un préfix unicast (temporaires)
FF3X::/96	Adresses SSM (temporaires)
FF7X::/16	Adresses IPv6 multicast "Embedded-RP" (temporaires)

Récapitulatif des types d'adresses multicast définis

2.1.6 Identifiant de groupe

Le RFC 3307 [14] décrit des procédures de création d'un identifiant de groupe (Group-ID). La taille de ce champ dans le document en question est de 32 bits conformément au RFC 2373. Ce dernier a été rendu obsolète par le RFC 3513 [27] qui a défini la nouvelle taille du champ Group-ID à 112 bits.

Le RFC 3307 [14] définit la correspondance entre les adresses IPv6 multicast et les adresses de niveau 2 : les 32 derniers bits de l'adresse multicast IPv6 sont ajoutés au préfixe MAC 33-33. Par exemple, l'adresse FF0E:30:2001:660:3001:4002:**AE45**:**2C56** correspondra à l'adresse MAC 33-33-**AE**-**45**-**2C**-**56**. La probabilité que 2 adresses multicast IPv6 utilisées sur un même lien correspondent à la même adresse MAC existe mais est très faible et les conséquences minimes. Restreindre le champ group-ID à 32 bits a toutefois un intérêt car cela apporte une homogénéité entre les différents types d'adresses décrits précédemment. En effet, dans le cas des adresses dérivées d'un préfixe unicast, ce champ a une longueur de 32 bits.

Le RFC 3307 définit aussi les adresses IPv6 multicast et identifiants de groupe qui seront gérés par l'IANA, où réservés pour des allocations dynamiques.

	Description	Valeur minimale de l'identifiant de groupe	Valeur maximale de l'identifiant de groupe
Adresse multicast permanente	C'est une adresse allouée par l'organisme IANA. Les bits P et T doivent être initialisés à zéro.	0x00000001	0x3FFFFFFF
Identifiant de groupe permanent	Le but de ces identifiants de groupe est de pouvoir identifier un service donné dans un réseau. Ces services sont définis par des Group-ID alloués par l'IANA et devraient être utilisées pour des adresses IPv6 multicast dérivées d'un préfixe unicast (RFC 3306). Avec cette méthode, il est théoriquement possible d'atteindre un service donné dans n'importe quel réseau.	0x40000000	0x7FFFFFFF
Adresse multicast dynamique	Les adresses multicast allouées dynamiquement doivent avoir un group-ID compris entre 0x80000000 et 0xFFFFFFFF. Ces adresses ont le bit T du champ *flags* positionné à 1.	0x80000000	0xFFFFFFFF

Récapitulatif des usages des identifiants de groupe

2.2 Allocation des adresses multicast IPv6

Un mécanisme d'allocation est aujourd'hui nécessaire pour assurer d'une part l'unicité de l'adresse sur le réseau, et d'autre part pour simplifier le processus de création d'une session multicast. En effet, l'existence des adresses multicast de format non trivial interdit la possibilité de laisser l'utilisateur choisir ces adresses. Le problème de l'allocation des adresses multicast n'apparait que dans le modèle ASM car le choix d'une adresse dans le modèle SSM est local à la source. Un canal SSM est en effet défini par un couple d'adresses (une adresse de source et une adresse de groupe). Deux sources SSM peuvent par conséquence utiliser indépendamment la même adresse multicast.

2.2.1 MADCAP

MADCAP (Multicast Address Dynamic Client Allocation Protocol) [22] est un protocole de type client-serveur utilisé pour l'allocation des adresses multicast (IPv4 et IPv6). MADCAP fait partie d'une architecture complète qui inclut les protocoles MASC et AAP. Ces protocoles permettent de diviser l'espace d'adressage multicast en segments attribués aux serveurs MADCAP, pour allocations aux stations sur les réseaux locaux.

Le RFC 3306 et la technologie Embedded-RP permettent de segmenter l'espace d'adressage suivant le préfixe unicast ou l'adresse du RP considéré. Dès lors il apparaît que MADCAP peut être utilisé seul pour attribuer ce type d'adresses.

Cependant, il faut noter la complexité du protocole et le manque total de déploiement de celui-ci depuis sa standardisation. Aussi, aucune implémentation de MADCAP ne supporte IPv6 à ce jour.

2.2.2 SAP

SAP (Session Announcement Protocol) [21] n'est pas un mécanisme d'allocation d'adresses multicast. Il permet, comme il sera expliqué dans la section 2.3.1, d'annoncer des sessions multicast. Ce mécanisme est cependant largement utilisé en IPv6 pour obtenir des adresses multicast. Le client SAP découvre

l'existence de toutes les sessions annoncées et peut ainsi choisir une adresse qui n'est pas utilisée. Ceci implique que :

- Un nombre restreint de sessions soient annoncées.
- Toutes les sessions qui obtiennent une adresse multicast IPv6 soient annoncées. Si ce n'est pas le cas, il est possible qu'un client choisisse une adresse multicast déjà utilisée puisqu'il n'a pas reçu l'annonce.
- Un point de rendez-vous soit déployé pour les adresses dérivées du préfixe multicast `FF0X::/16` puisque les annonces SAP sont émises avec l'adresse multicast `FF0X::2:7FFE`. Ceci n'est pas compatible par exemple avec Embedded-RP.

SAP ne peut être utilisé que sur des portées restreintes puisqu'il ne peut pas gérer des dizaines de milliers d'adresses.

2.2.3 Choix de l'adresse au hasard

L'utilisateur peut aussi spécifier l'adresse multicast IPv6 qu'il souhaite utiliser pour la session multicast qu'il souhaite créer. Il choisira alors l'adresse au hasard et n'aura aucun moyen de s'assurer que l'adresse choisie n'est pas déjà utilisée. Néanmoins, devant le grand nombre d'adresses disponibles, la probabilité que l'adresse choisie soit déjà utilisée est pratiquement nulle.

Le vrai problème posé par cette approche est qu'il n'est pas possible de laisser l'utilisateur construire des adresses dérivées d'un préfixe unicast ou des adresses de type embedded-RP. Cette solution n'est pas imaginable car elle ne supporte donc pas le passage à l'échelle et n'est pas transparente pour l'utilisateur qui doit configurer manuellement l'adresse multicast.

2.2.4 Conclusion

Il n'existe donc pas aujourd'hui de méthode d'allocations d'adresse multicast IPv6 satisfaisant les critères recherchés : transparence pour l'utilisateur, passage à l'échelle et implémentation disponible. A ce stade du déploiement du multicast IPv6, les adresses multicast IPv6 sont soit choisies manuellement ou alors allouées par SAP.

2.3 Annonce des sessions multicast IPv6

Il n'y a pas de différence fondamentale entre une annonce de session multicast IPv6 et IPv4.

2.3.1 SAP - Session Announcement Protocol

SAP est décrit dans le RFC 2974 [21]. Ce protocole permet à un client d'annoncer des sessions multicast sur une portée prédéfinie. Les annonces sont émises périodiquement sur le groupe multicast IPv6 `FF0X:0:0:0:0:0:2:7FFE`. Lorsqu'un client désire connaître l'ensemble des sessions annoncées, il s'abonne au groupe des annonces SAPv1 et reçoit toutes les annonces.

```
FF0X:0:0:0:0:0:2:7FFE    SAPv1 Announcements
FF0X:0:0:0:0:0:2:7FFF    SAPv0 Announcements (deprecated)
FF0X:0:0:0:0:0:2:8000 - FF0X:0:0:0:0:0:2:FFFF  SAP Dynamic Assignments
```

Adresses réservées pour le protocole SAP

Il est difficile d'imaginer qu'un tel protocole puisse convenir pour un usage intense du multicast, avec annonces de toutes parts de plusieurs milliers de sessions multicast. SAP remplit les besoins actuels car l'utilisation du multicast IPv6 est encore faible, mais si l'utilisation devient intense, SAP pourrait mieux convenir pour des sites, ou pour l'annonce de sessions propres à un fournisseur d'accès à Internet.

2.3.2 Messagerie électronique, pages web...

Une solution simple pour annoncer des sessions est d'utiliser des moyens plus traditionnels comme la messagerie ou les pages web. Comme les pages web des principales stations de radio permettent aujourd'hui d'écouter leurs programmes sur Internet en unicast, il paraît simple d'imaginer la même chose en multicast IPv6. La messagerie électronique s'applique d'avantage aux sessions avec un nombre d'abonnés restreint.

3 Le multicast IPv6 sur le lien-local

3.1 Gestion des abonnements sur le lien-local : MLD

Pour offrir un service de distribution multicast, deux composants sont nécessaires : un protocole de gestion de groupe multicast et un protocole de construction d'arbre multicast. Le protocole de gestion de groupe multicast réalise la signalisation entre l'hôte et son routeur d'accès à l'Internet. En IPv6, ce protocole est MLD (*Multicast Listener Discovery*). Il est utilisé par un routeur de bordure IPv6 pour découvrir la présence de récepteurs multicast sur ses liens directement attachés, ainsi que les adresses multicast concernées.

MLD est un protocole asymétrique qui spécifie un comportement différent pour les hôtes et les routeurs multicast. Toutefois, pour les adresses multicast sur lesquelles un routeur lui-même écoute, il doit exécuter les deux parties du protocole et répondre à ses propres messages.

Comme MLD est un sous-protocole d'ICMPv6, les messages MLD sont des messages ICMPv6 particuliers. Ils sont envoyés avec :

- une adresse source IPv6 lien-local ;
- le champ "nombre de sauts" fixé à 1 ;
- l'option "IPv6 Router Alert" activée.

Cette dernière option est nécessaire afin de contraindre les routeurs à examiner les messages MLD envoyés à des adresses multicast par lesquelles les routeurs ne sont pas intéressés. La version d'origine du protocole MLD [RFC2710] (que nous appellerons également MLDv1) présente les mêmes fonctionnalités que le protocole IGMPv2 en IPv4.

Trois types de messages sont utilisés. Leur format est donné sur la figure 4-17 :

- recensement des récepteurs multicast (type = 130) avec deux sous-types de messages :
  - recensement général émis à l'adresse de diffusion générale sur le lien (`FF02::1`)
  - recensement spécifique à une adresse multicast, l'adresse de destination est l'adresse multicast du groupe en question
- rapport d'abonnement multicast (type = 131), l'adresse de destination est l'adresse multicast du groupe en question
- résiliation d'abonnement multicast (type = 132), émis à l'adresse du groupe multicast "tous les routeurs du lien local" (`FF02::2`).

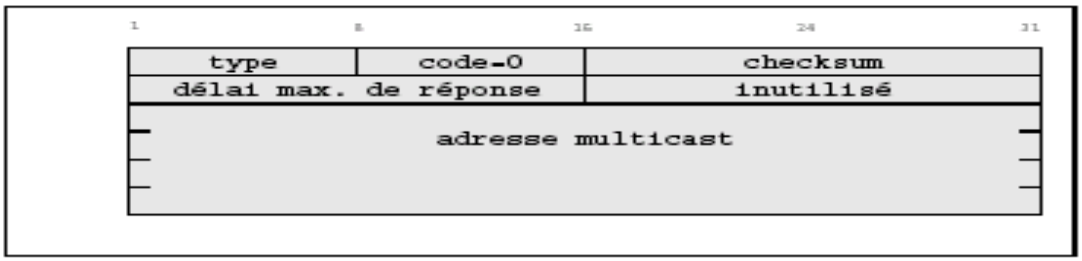


Figure 4-17 . Format d'un message de gestion de groupe multicast

Les champs ont la signification suivante :

- `type`
- `code:` mis à zéro par l'émetteur et ignoré par les récepteurs
- `checksum:` celui du protocole ICMPv6 standard, couvrant tout le message MLD auquel s'ajoutent les champs du pseudo-en-tête IPv6
- `délai maximal de réponse:`
  - utilisé seulement dans les messages de recensement. Il exprime le retard maximal autorisé (en millisecondes) pour l'arrivée des rapports d'abonnement
  - dans les messages de rapport ou de résiliation d'abonnement ce champ est mis à zéro par l'émetteur et ignoré par les récepteurs
- `inutilisé :` mis à zéro par l'émetteur et ignoré par les récepteurs
- `adresse multicast:`
  - pour un message de recensement général ce champ est mis à zéro
  - pour un message de recensement spécifique il contient l'adresse multicast en question
  - pour les messages de rapport et de résiliation d'abonnement, le champ contient l'adresse multicast sur laquelle l'hôte souhaite écouter ou cesser d'écouter

3.1.1 Messages de recensement et rapports d'abonnement périodiques MLD

Le routeur envoie régulièrement des messages de recensement général à l'adresse de diffusion générale sur le lien (`FF02::1`). Les hôtes arment un temporisateur pour chaque adresse multicast qui les concerne. Si un temporisateur expire sans que l'hôte ait entendu une réponse d'un de ses voisins concernant la même adresse, il envoie un rapport d'abonnement à l'adresse multicast du groupe. Ce système de temporisateurs permet aux hôtes de surveiller les rapports des autres hôtes sur le lien et d'annuler leurs propres rapports concernant les mêmes adresses. Ainsi la quantité du trafic MLD peut être minimisée.

3.1.2 Rapports d'abonnements MLD non-sollicités

Les changements d'état des hôtes sont notifiés par des messages non-sollicités :

- Pour souscrire à une adresse multicast spécifique, un hôte envoie un rapport d'abonnement non-sollicité ;
- Pour cesser d'écouter sur une adresse multicast, l'hôte peut simplement ne plus répondre aux messages de recensement du routeur. S'il est le seul récepteur de cette adresse multicast sur le lien, après un certain temps l'état du routeur concernant cette adresse expire. Le routeur arrêtera de faire suivre les paquets multicast envoyés à l'adresse en question, s'il s'avère que l'hôte était le dernier concerné par l'adresse multicast sur le lien;
- La résiliation rapide est aussi une possibilité offerte par MLDv1. L'hôte envoie un message de résiliation d'abonnement à l'adresse multicast de "tous les routeurs du lien local" (`FF02::2`). Le

routeur répond avec un message de recensement spécifique à l'adresse en question. S'il n'y a plus de récepteur pour répondre à ce recensement, le routeur efface l'adresse multicast de sa table de routage.

Il est possible d'avoir plusieurs routeurs multicast sur le même lien local. Dans ce cas un mécanisme d'élection est utilisé pour choisir le routeur recenseur. Celui-ci sera le seul responsable pour l'envoi des messages de recensement.

3.1.3 Exemples de fonctionnement de MLDv1

Les paquets suivants ont été capturés lors de l'exécution d'un programme (multi2out6, dont le code est donné après). Ce programme prend comme arguments une interface de la machine et une adresse multicast. Dans cet exemple, l'adresse choisie `ff12::1234:5678`, représente un groupe éphémère (valeur `0x1` du drapeau) sur le lien local (valeur `0x02`).

L'interface se joint à ce groupe multicast et commence par émettre un rapport d'abonnement :

```
En-tête IPv6
Version : 6 Classe : 00 Label : 00000
Longueur : 32 octets (0x0020) Proto. : 0 (0x0) "Proche-en-proche"
Nombre de sauts : 1
Source : fe80::0a00:20ff:fe18:964c
Desti. : ff12::1234:5678 (adresse du groupe multicast)
Proche-en-proche
En-tête Suivant : 58 (0x3a) ICMPv6/MLD
Type : 5 (0x5) Router Alert longueur : 2 valeur : 0


ICMPv6/MLD
Type : 131 (0x83) rapport d'abonnement
Code : 0
Checksum : 0xef48
Délai maximal de réponse : 0
Adresse multicast : ff12::1234:5678 (adr du grp multicast en question)


0000: 60 00 00 00 00 20 00 01 fe 80 00 00 00 00 00 00
0010: 0a 00 20 ff fe 18 96 4c ff 12 00 00 00 00 00 00
0020: 00 00 00 00 12 34 56 78 3a 00 05 02 00 00 00 00
0030: 83 00 ef 48 00 00 00 00 ff 12 00 00 00 00 00 00
0040: 00 00 00 00 12 34 56 78
```

En arrêtant le programme, l'interface en question se désabonne du groupe multicast et en s'apercevant qu'elle est la dernière à avoir envoyé un rapport concernant ce groupe, elle émet un message de fin d'abonnement :

```
En-tête IPv6
Version : 6 Classe : 00 Label : 00000
Longueur : 32 octets (0x0020) Proto. : 0 (0x0) "Proche-en-proche"
Nombre de sauts : 1
Source : fe80::0a00:20ff:fe18:964c
Desti. : ff12::1234:5678
Proche-en-proche
En-tête Suivant : 58 (0x3a) ICMPv6/MLD
Type : 5 (0x5) Router Alert longueur : 2 valeur : 0
```

```
ICMPv6/MLD
Type : 132 (0x84) Fin d'abonnement
Code : 0
Checksum : 0x5703
Délai maximal de réponse : 0
Adresse multicast : ff12::1234:5678 (adr du grp multicast en question)

0000: 60 00 00 00 00 20 00 01 fe 80 00 00 00 00 00 00
0010: 0a 00 20 ff fe 18 96 4c ff 02 00 00 00 00 00 00
0020: 00 00 00 00 00 00 00 02 3a 00 05 02 00 00 00 00
0030: 84 00 57 03 00 00 00 00 ff 12 00 00 00 00 00 00
0040: 00 00 00 00 12 34 56 78
```

3.2 Gestion des abonnements sur le lien-local : MLD version 2

La nouvelle version du protocole de gestion de groupe multicast, MLDv2 est décrite dans le RFC 3810 [28]. Elle implante les fonctionnalités du protocole IGMPv3 défini pour IPv4, la plus importante étant l'introduction du filtrage des sources. Un hôte peut désormais spécifier les sources qu'il veut ou qu'il ne veut pas écouter pour une adresse multicast donnée. Cette information peut être utilisée par les protocoles de routage multicast afin d'éviter l'acheminement des paquets multicast provenant de certaines sources vers des liens où il n'y a pas de récepteur intéressé.

Pour être en mesure de supporter les fonctionnalités de MLDv2, l'API de l'hôte doit permettre l'opération suivante (ou un équivalent logique de celle-ci) :

```
EcouteIPv6Multicast (socket, interface, adresse multicast IPv6, mode de
filtrage, liste de sources)
```

Par cet appel, une application demande, pour une certaine adresse multicast, la réception de paquets sur une certaine interface, en tenant compte du mode de filtrage et de la liste des sources spécifiées. Le mode de filtrage peut être soit INCLUDE, soit EXCLUDE :

- En mode INCLUDE, la réception des paquets envoyés à l'adresse multicast spécifiée est demandée seulement pour ceux en provenance des sources présentes dans la liste qui suit
- En mode EXCLUDE, la réception des paquets est demandée pour toutes les sources, à l'exception de celles spécifiées dans la liste de sources

Il existe deux types de messages MLDv2 :

- recensement des récepteurs multicast (type=130)
- rapport d'abonnement multicast version 2 (type=143)

Pour garder l'interopérabilité avec la version précédente de MLD, les messages de rapport d'abonnement multicast version 1 et de résiliation d'abonnement multicast sont également supportés.

3.2.1 Messages de recensement MLDv2

Un message de recensement des récepteurs en MLDv2 est donné sur la figure suivante :

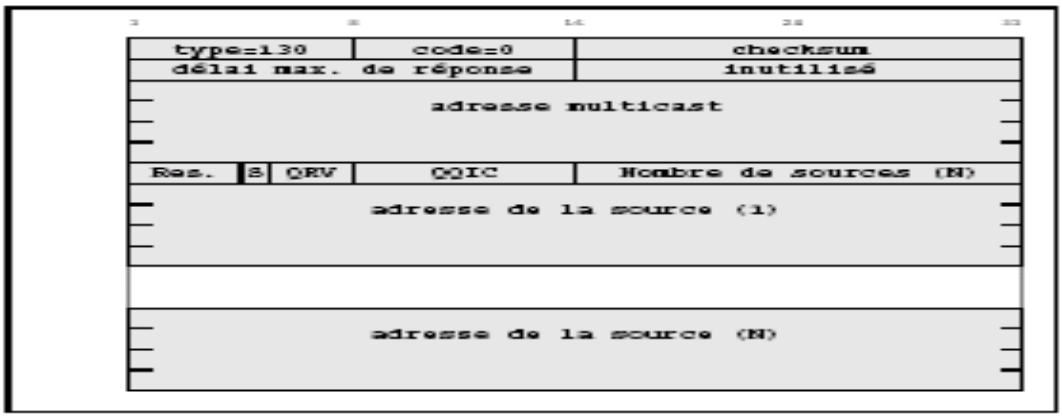


Figure 4-18 . Format d'un message de recensement MLDv2

Les champs ont la signification suivante :

- `type`, le même type qu'en MLDv1
- `code`, mis à zéro par l'émetteur et ignoré par les récepteurs
- `checksum`, calculé de la même façon que pour la version précédente du protocole
- `délai max. de réponse`, utilisé pour calculer le délai maximal de réponse durant lequel le récepteur doit envoyer éventuellement son rapport d'abonnement
- `inutilisé`, mis à zéro par l'émetteur et ignoré par les récepteurs
- `adresse multicast`
- `réservé`, mis à zéro par l'émetteur et ignoré par les récepteurs
- drapeau `S`, indique aux routeurs multicast qui reçoivent ce message s'ils doivent ou pas supprimer la mise à jour des temporisateurs, effectuée normalement au moment de la réception d'un message de recensement
- `QRV`, contient la variable de robustesse utilisée par le recenseur (le nombre de fois qu'un récepteur envoie un rapport pour être robuste aux pertes dans le réseau)
- `QQIC`, code utilisé pour calculer l'intervalle de recensement
- `nombre de sources`
- `adresse de la source [N]`, vecteur contenant la liste éventuelle des sources.

Trois types de messages de recensement de récepteurs multicast sont utilisés :

- recensement général envoyé par un routeur multicast afin de découvrir les adresses multicast pour lesquelles il y a des récepteurs sur ses liens directs. Dans un tel message les champs "adresse multicast" et "nombre de sources" sont mis à zéro
- recensement spécifique à une adresse multicast envoyé par un routeur multicast afin de découvrir l'existence de récepteurs pour une adresse multicast spécifique. Le champ "adresse multicast" contient l'adresse en question, tandis que le champ "nombre de sources" est mis à zéro
- recensement spécifique à une adresse multicast et à une source envoyé par un routeur multicast afin de découvrir l'existence de récepteurs pour une adresse multicast et une source spécifiques. Le champ "adresse multicast" contient l'adresse en question, tandis que les champs "adresse

source [i]" forment un vecteur de N adresses unicast (valeur spécifiée dans le champ "nombre de sources").

Les messages de recensement général sont envoyés à l'adresse de diffusion générale sur le lien (`FF02::1`). Les autres messages de recensement sont envoyés à l'adresse multicast spécifiée dans l'entête MLDv2.

3.2.2 Rapports d'abonnement MLDv2

Un rapport d'abonnement multicast en MLDv2 est donné sur la figure suivante :

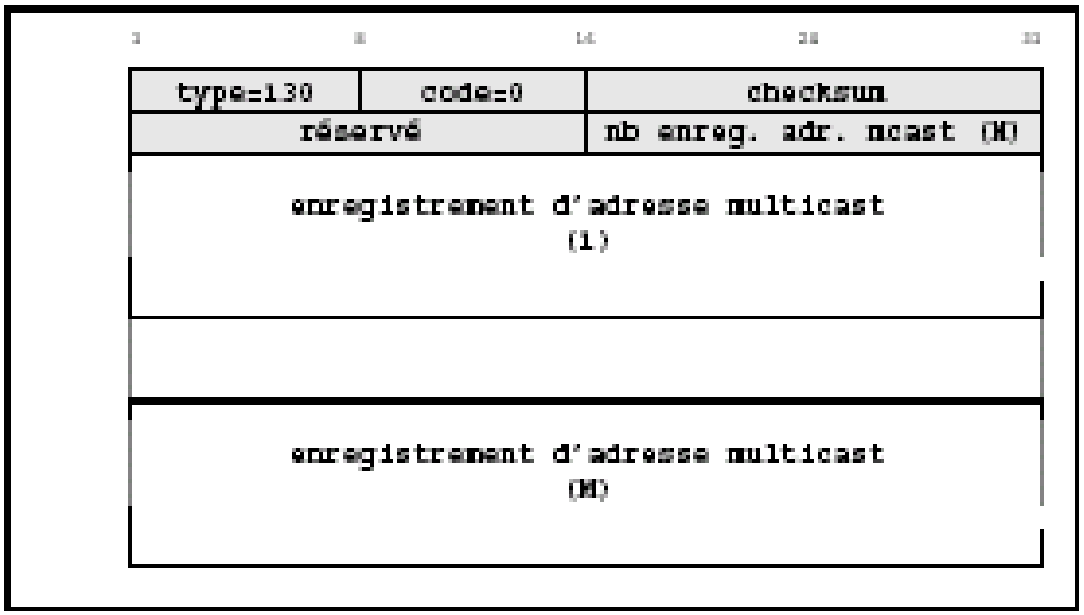


Figure 4-19 . Format d'un message de rapport d'abonnement MLDv2

Les champs ont la signification suivante :

- `type,` type=143
- `réservés,` mis à zéro par l'émetteur et ignorés par les récepteurs
- `checksum,` calculé de la même façon que pour la version précédente du protocole
- `nombre d'enregistrements d'adresse multicast`
- `enregistrement d'adresse multicast :` chaque enregistrement d'adresse multicast a la forme donnée sur la figure suivante :

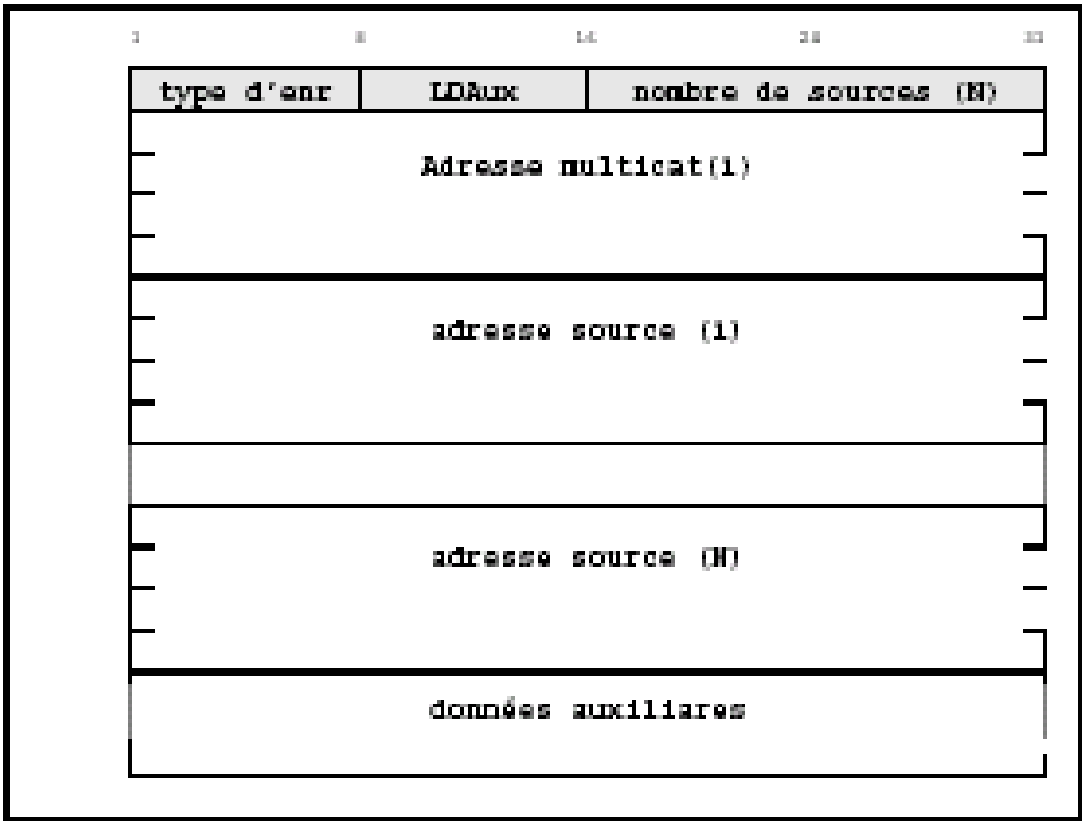


Figure 4-20 . Format d'un enregistrement d'adresse multicast

Les champs ont la signification suivante :

- `type` : plusieurs types d'enregistrements d'adresse multicast peuvent être inclus dans un rapport d'abonnement :
    - un "enregistrement d'état actuel" est envoyé par un hôte en réponse à un message de recensement. Il décrit l'état de l'hôte concernant une adresse multicast spécifique .Le champ "type d'enregistrement" peut dans ce cas avoir les valeurs MODE_IS_INCLUDE ou MODE_IS_EXCLUDE
    - un "enregistrement de changement de mode de filtrage" est envoyé par un hôte chaque fois qu'un appel `EcouteIPv6Multicast` modifie son mode de filtrage pour une adresse multicast précise. Le champ "type d'enregistrement" peut dans ce cas avoir les valeurs CHANGE_TO_INCLUDE_MODE ou CHANGE_TO_EXCLUDE_MODE
    - un "enregistrement de changement de la liste des sources" est envoyé par un hôte quand un appel `EcouteIPv6Multicast` modifie la liste des sources qu'il souhaite ou ne souhaite pas écouter pour une adresse multicast précise. Le champ "type d'enregistrement" peut dans ce cas avoir les valeurs ALLOW_NEW_SOURCES ou BLOCK_OLD_SOURCES
- `LDAux` contient la longueur du champ `données auxiliaires`
- `adresse multicast`
- `nombre de sources`
- `adresse source [i]`, vecteur contenant la liste des sources qui doivent être désormais autorisées ou bloquées

- `données auxiliaires`, si présentes, elles peuvent contenir des informations supplémentaires concernant l'enregistrement d'adresse multicast en question

Les rapports d'abonnement sont envoyés par les hôtes à l'adresse "tous les routeurs MLDv2" (ff02::16). Ainsi les récepteurs ne reçoivent pas les rapports des autres, chacun étant obligé d'envoyer son propre rapport. Le mécanisme d'économie d'émission des rapports du protocole MLDv1 n'est donc plus utilisé. Le risque de surcharge du routeur à cause du trop grand nombre de rapports reçus est toutefois évité en fixant des temporisateurs avec des valeurs différentes pour chaque rapport.

3.2.3 Fonctionnement de MLDv2

Comme dans le cas du MLDv1, s'il existe plusieurs routeurs multicast sur le même lien local, un seul routeur recenseur va être désigné à l'aide d'un mécanisme spécifique. Le recenseur envoie régulièrement des messages de recensement général auxquels les récepteurs répondent avec des rapports d'abonnement contenant des enregistrements d'état actuel.

Chaque hôte, ainsi que chaque routeur, gardent un état contenant le mode de filtrage et une liste des sources pour chaque adresse multicast. Dans le cas d'un hôte ce sont les adresses multicast sur lesquelles il écoute. Dans le cas d'un routeur ce sont les adresses multicast que ses récepteurs écoutent. Une application sur une machine hôte demande la modification du mode de filtrage ou de la liste des sources pour une adresse multicast précise à travers d'un appel `EcouteIPv6Multicast`. L'hôte inclut par conséquent ces modifications dans un rapport d'abonnement non-sollicité. Ce rapport contient des enregistrements de changement de mode de filtrage ou de la liste des sources, en conformité avec les changements qui ont eu lieu dans l'état interne de l'hôte.

Au moment de la réception des changements communiqués, le routeur met à jour son état pour l'adresse multicast concernée. Si, suite à ces changements, il constate qu'une certaine source ne doit plus être acceptée, le routeur envoie un message de recensement spécifique pour cette source, afin de vérifier l'existence d'éventuels récepteurs qui souhaitent toujours l'écouter. Un temporisateur est déclenché, et s'il expire sans que le routeur ait reçu un rapport d'abonnement concernant la source, celle-ci est éliminée de l'état local du routeur. Si le routeur détecte que plus aucune source n'est sollicitée pour une certaine adresse, il envoie un message de recensement spécifique pour cette adresse. Si des rapports d'abonnement concernant l'adresse en question ne sont pas reçus en temps dû, l'adresse est effacée de l'état du routeur.

3.2.4 Exemples de fonctionnement de MLDv2

Voici quelques exemples pour illustrer le fonctionnement du protocole MLDv2.

Le routeur recenseur envoie un message de recensement général :

```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 36 octets (0x0024)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::240:95ff:fe49:ba9
Adresse destination : ff02::1 (adresse de diffusion générale sur le
lien)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient à 2 octets de bourrage)
Router alert : 0x05 Longueur : 0x02 Valeur : 0x0000 (pour les messages
MLD)
ICMPv6 :
```

```
Type: 130 (0x82) - message de recensement
Code : 0 (0x00)
Somme de contrôle : 0xb464
Code de réponse maximal : 10000 (Ox2710) Réservé : 0x0000
Adresse multicast : 0::0 Réservé : 0x0
Drapeau S : 0
QRV : 2
QQIC : 125 (0x7d)
Nombre de sources: 0 (il s'agit d'un recensement général)
0x0000 6000 0000 0024 0001 fe80 0000 0000 0000
0x0010 0240 95ff fe49 0ba9 ff02 0000 0000 0000
0x0020 0000 0000 0000 0001 3a00 0100 0502 0000
0x0030 8200 b464 2710 0000 0000 0000 0000 0000
0x0040 0000 0000 0000 0000 027d 0000
```

Un hôte envoie un rapport d'abonnement avec des enregistrements d'état actuel :

```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 76 octets (0x004c)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::203:47ff:fe7c:b9c5
Adresse destination : ff02::16 (tous les routeurs MLDv2 sur le lien)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient a 2 octets de bourrage)
Router alert : 0x0502 Valeur: 0x0000 (pour les messages MLD)
ICMPv6 :
Type: 143 (0x8f) - rapport d'abonnement
Réservé : 0x00
Somme de contrôle : 0x9454
Réservé : 0x0000
Nombre d'enregistrements : 0x0003
Type d'enregistrement : 0x02 (MODE_IS_EXCLUDE)
Longueur des données auxiliaires : 0x00
Nombre de sources : 0x0000
Adresse multicast : ff02::9

Type d'enregistrement : 0x02 (MODE_IS_EXCLUDE)
Longueur des données auxiliaires : 0x00
Nombre de sources : 0x0000
Adresse de la source : ff02::2:816a:9e88
Type d'enregistrement : 0x02 (MODE_IS_EXCLUDE)
Longueur des données auxiliaires : 0x00
Nombre de sources : 0x0000
Adresse multicast : ff02::1:ff7c:b9c5
0x0000 6000 0000 004c 0001 fe80 0000 0000 0000
0x0010 0203 47ff fe7c b9c5 ff02 0000 0000 0000
0x0020 0000 0000 0000 0016 3a00 0100 0502 0000
0x0030 8f00 9454 0000 0003 0200 0000 ff02 0000
0x0040 0000 0000 0000 0000 0000 0009 0200 0000
0x0050 ff02 0000 0000 0000 0000 0002 816a 9e88
```

```
0x0060 0200 0000 ff02 0000 0000 0000 0000 0001
0x0070 ff7c b9c5
```

Un hôte rajoute une source dans la liste des sources qu'il veut écouter :

```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 52 octets (0x0034)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::2e0:29ff:fe3e:db03
Adresse destination : ff02::16 (tous les routeurs MLDv2 sur le lien)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient a 2 octets de bourrage)
Router alert : 0x0502 Valeur: 0x0000 (pour les messages MLD)
ICMPv6 :
Type: 143 (0x8f) - rapport d'abonnement
Réservé : 0x00
Somme de contrôle : 0x6b59
Réservé : 0x0000
Nombre d'enregistrements : 0x0001
Type d'enregistrement : 0x05 (ALLOW_NEW_SOURCES)
Longueur des données auxiliaires : 0x00
Nombre de sources : 0x0001
Adresse multicast : ff34::17
Adresse source : 2001:660:10d:4105:50:fcff:fe0b:9966
0x0000 6000 0000 0034 0001 fe80 0000 0000 0000
0x0010 02e0 29ff fe3e db03 ff02 0000 0000 0000
0x0020 0000 0000 0000 0016 3a00 0100 0502 0000
0x0030 8f00 6b59 0000 0001 0500 0001 ff34 0000
0x0040 0000 0000 0000 0000 0000 0017 2001 0660
0x0050 010d 4105 0050 fcff fe0b 9966
```

Un hôte ne désire plus écouter une source donnée :

```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 52 octets (0x0034)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::2e0:29ff:fe3e:db03
Adresse destination : ff02::16 (tous les routeurs MLDv2 sur le lien)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient a 2 octets de bourrage)
Router alert : 0x0502 Valeur: 0x0000 (pour les messages MLD)
ICMPv6 :
Type: 143 (0x8f) - rapport d'abonnement
Réservé : 0x00
```

```
Somme de contrôle : 0x6a59
Réservé : 0x0000
Nombre d'enregistrements : 0x0001
Type d'enregistrement : 0x06 (BLOCK_OLD_SOURCES)
Longueur des données auxiliaires : 0X00
Nombre de sources : 0x0001
Adresse multicast : ff34::17
Adresse source : 2001:660:10d:4105:50:fcff:fe0b:9966

0x0000 6000 0000 0034 0001 fe80 0000 0000 0000
0x0010 02e0 29ff fe3e db03 ff02 0000 0000 0000
0x0020 0000 0000 0000 0016 3a00 0100 0502 0000
0x0030 8f00 6a59 0000 0001 0600 0001 ff34 0000
0x0040 0000 0000 0000 0000 0000 0017 2001 0660
0x0050 010d 4105 0050 fcff fe0b 9966
```

Un routeur envoie un message de recensement spécifique à une adresse multicast et à une source :

```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 52 octets (0x0034)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::240:95ff:fe49:ba9
Adresse destination : ff34::17 (adresse multicast concernée)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient a 2 octets de bourrage)
Router alert : 0x0502 Valeur: 0x0000 (pour les messages MLD)
ICMPv6 :
Type: 130 (0x82) - message de recensement
Code : 0 (0x00)
Somme de contrôle : 0xdab1
Code de réponse maximal : 1000 (0x03e8)
Réservé : 0x0000
Adresse multicast : ff34::17
Réservé : 0x0
Drapeau S : 0
QRV : 2
QQIC : 125 (0x7d)
Nombre de sources : 0x0001
Adresse de la source : 2001:660:10d:4105:50:fcff:fe0b:9966
0x0000 6000 0000 0034 0001 fe80 0000 0000 0000
0x0010 0240 95ff fe49 0ba9 ff34 0000 0000 0000
0x0020 0000 0000 0000 0017 3a00 0100 0502 0000
0x0030 8200 dab1 03e8 0000 ff34 0000 0000 0000
0x0040 0000 0000 0000 0017 027d 0001 2001 0660
0x0050 010d 4105 0050 fcff fe0b 9966
```

Un hôte envoie un rapport d'abonnement contenant des enregistrements de changement de mode de filtrage :

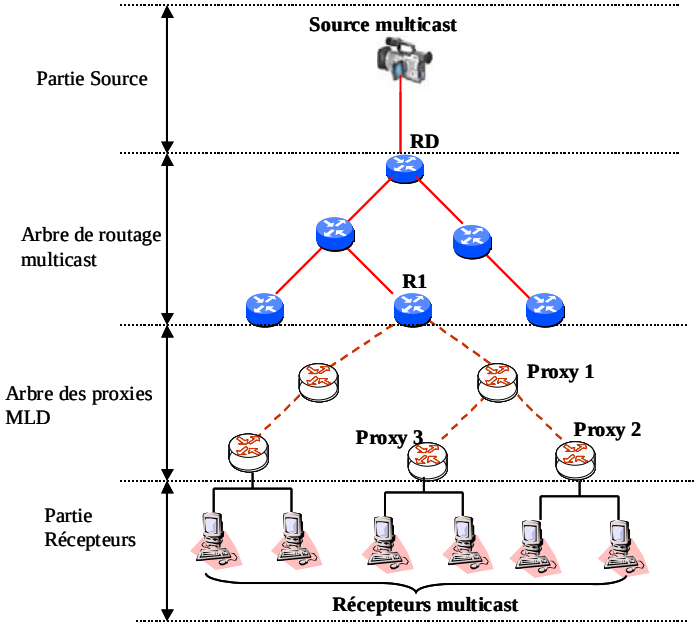
```
En-tête IPv6 :
Version : 6 Classe de trafic : 0x00 Identifiant de flux : 0x00000
Longueur des données : 52 octets (0x0034)
En-tête suivant : extension proche-en-proche (0x00) Nombre de sauts :
0x01
Adresse source : fe80::2e0:29ff:fe3e:db03
Adresse destination : ff02::16 (tous les routeurs MLDv2 sur le lien)
Extension proche-en-proche :
En-tête suivant : ICMPv6 (0x3a)
Longueur : 0x00 (nombre de mots de 64 bits -1)
PadN : 0x01 Longueur : 0x00 (ce qui revient a 2 octets de bourrage)
Router alert : 0x0502 Valeur: 0x0000 (pour les messages MLD)
ICMPv6 :
Type: 143 (0x8f) – rapport d'abonnement
Réservé : 0x00
Somme de contrôle : 0x6c49
Réservé : 0x0000
Nombre d'enregistrements : 0x0001
Type d'enregistrement : 0x04 (CHANGE_T0_EXCLUDE_MODE)
Longueur des données auxiliaires : 0x00
Nombre de sources : 0x0001
Adresse multicast : ff44::17
Adresse source : 2001:660:10d:4105:50:fcff:fe0b:9966
0x0000 6000 0000 0034 0001 fe80 0000 0000 0000
0x0010 02e0 29ff fe3e db03 ff02 0000 0000 0000
0x0020 0000 0000 0000 0016 3a00 0100 0502 0000
0x0030 8f00 6c49 0000 0001 0400 0001 ff44 0000
0x0040 0000 0000 0000 0000 0000 0017 2001 0660
0x0050 010d 4105 0050 fcff fe0b 9966
```

3.3 MLD Fowarding Proxy

Afin d'éviter de déployer des routeurs multicast dans un domaine donné, l'IETF a récemment proposé l'utilisation de proxies MLD. Comme le montre la Erreur ! Source du renvoi introuvable., les proxies peuvent former un arbre de gestion de groupe enraciné sur un routeur multipoint. Seul le routeur R1 est responsable de joindre le groupe et établir la branche multipoint vers l'arbre. Chaque proxy (Proxy 2 et Proxy 3) collecte localement les informations de gestion de groupe sur ses propres liens et transmet un rapport d'abonnement (*Host Membership Report*) vers son proxy hiérarchiquement supérieur (Proxy 1). Proxy 1 se charge de transmettre un autre rapport d'abonnement qui reflète exactement l'information de gestion de groupe des proxies 2 et 3.

Pour maintenir une base d'information des abonnements, chaque proxy maintient un ensemble d'enregistrement d'abonnement pour chacune de ses interfaces. Chaque enregistrement a la forme suivante :

```
(adresse multicast IPv6, mode de filtrage, liste de sources)
```



Gestion de groupe avec des Proxies MLD

L'avantage du recours à des proxies est d'éviter l'utilisation d'un protocole de construction d'arbre multicast dans certains types de topologies de réseau simples comme les DSLAM (*Digital Subscriber Line Acess Multiplexer*). Dans une telle topologie, seul le routeur de bordure du réseau est sensé implémenter les fonctionnalités de construction d'arbre multicast ce qui simplifie l'architecture et le coût des équipements d'accès. De même la charge du réseau est considérablement réduite grâce à la suppression du trafic de signalisation pour la maintenance de l'arbre multicast.

Cependant, l'utilisation de proxies peut avoir des inconvénients. Par exemple, les proxies ne permettent pas une tolérance des défaillances de liens ou de routeurs puisque les proxies ne peuvent pas reconstruire un arbre multicast en fonction de l'état du réseau. La panne d'un proxy d'une hiérarchie donnée, entraîne la panne de tous les proxies du niveau inférieur. Par conséquent les récepteurs multicast ne peuvent plus recevoir du trafic multicast ni envoyer les rapports d'abonnements au routeur multicast.

3.4 MSNIP, un nouvelle extension à MLDv2

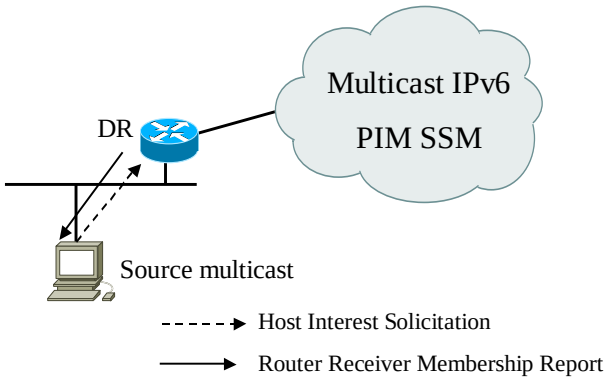
Comme décrit précédemment, MLD est un protocole asymétrique qui spécifie un comportement différent pour les hôtes et pour les routeurs. Dans les sections précédentes, nous avons détaillé le comportement d'un hôte récepteur. Cependant, des nouvelles extensions sont proposées par l'IETF pour prendre en considération le cas où un hôte est une source multicast. Nous rappelons ici qu'une source multicast n'a pas besoin de connaître les membres des groupes auxquels elle veut émettre. Toutefois, il est important qu'une source multicast soit informée de l'arrivée et des départs des récepteurs multicast afin d'envoyer les données multicast au moment opportun.

Les nouvelles propositions consistent à l'introduction d'un nouveau protocole dit « *Protocole de Notification d'Intérêt d'une Source Multicast* » ou MSNIP (*Multicast Source Notification of Interest Protocol*). Ce protocole est une extension du protocole MLDv2. Il est dédié aux sources qui utilisent une plage d'adressage multicast SSMv6. Il a pour but principal de synchroniser l'émission du trafic multicast entre un hôte source et son routeur multicast afin d'éviter toute émission inutile.

Le protocole MSNIP comporte deux parties :

- *Partie hôte* : cette partie gère les opérations d'enregistrement, d'émission des données multicast et la résiliation de l'hôte source. Avant d'émettre, la source se fait connaître du Designated Router avec un message "Host Interest Sollicitation" et attend l'annonce de récepteurs pour commencer à émettre.

- *Partie routeur* : cette partie notifie la source de l'arrivée du premier membre multicast ainsi que du départ du dernier membre grâce à un message "Router Received Membership Report".



3.5 La diffusion du multicast IPv6 sur le lien-local

Sur le lien-local, un segment ethernet par exemple, les paquets IPv6 multicast sont le plus généralement broadcastés. L'adresse MAC utilisée pour transporter des paquets multicast est une adresse MAC multicast (le 8ème bit de poids fort est positionné à 1). Cette adresse MAC est la concaténation de 33-33 et des 32 bits de poids faible de l'adresse IPv6 multicast. Si l'on considère l'adresse multicast IPv6 FF1E::12:AC21:6521, l'adresse MAC correspondante sera 33-33-AC-21-65-21

Les mécanismes qui existent en IPv4 comme IGMP snooping ou CGMP (Cisco Group Multicast Protocol) qui permettent de limiter la diffusion des paquets multicast aux hosts abonnés n'existent pas encore pour IPv6. MLD snooping n'est toujours pas implémenté sur les commutateurs. La conséquence directe est que si le réseau n'est pas segmenté correctement, à l'aide de VLANs par exemple, le flux multicast inondera tout le réseau et chaque host recevra des paquets multicast non souhaités.

4 La construction d'arbre multicast - PIM

4.1 Introduction

Sur le lien-local, le protocole MLD permet aux stations de travail d'exprimer leur intérêt pour un groupe multicast (et d'un ensemble de sources pour MLDv2). Il reste ensuite à acheminer les paquets multicast IPv6 entre les sources et les abonnés. Ceci est réalisé par le protocole PIM *(Protocol Independant Multicast)*. Le fonctionnement du protocole PIM en IPv6 est le même que pour IPv4. Aussi l'objectif de cette section est d'expliquer les bases du protocole de construction d'arbre multicast PIM pour permettre une meilleure compréhension du chapitre multicast.

4.2 Le protocole PIM SM - Sparse-Mode

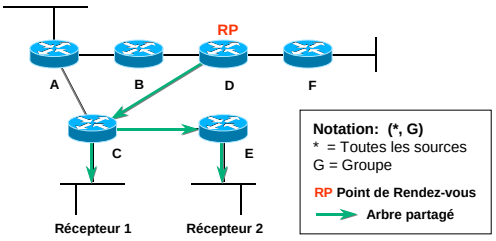
Le protocole PIM-SM *(Protocol Independent Multicast - Sparse Mode)* permet la construction d'arbres multicast (RFC 2362 [5]). Ce protocole peut utiliser la base d'information de routage unicast sous-jacente, ou une autre base d'information de routage multicast comme BGP IPv6 multicast SAFI : dans ce sens il est indépendant. Il construit pour chaque groupe un arbre de diffusion unidirectionnelle, chaque arbre prenant racine sur un nœud spécifique appelé point de rendez-vous ou RP *(Rendez-vous Point)*. Lorsqu'il y a plusieurs sources alimentant le même groupe, les paquets en provenance des différentes sources convergent vers le RP associé au groupe, puis à partir de celui-ci les paquets empruntent (et donc partagent) l'arbre associé au groupe, ce qui leur permet d'atteindre tous les destinataires membres du groupe. La construction de l'arbre de diffusion peut se décomposer en 3 étapes.

4.2.1 Etape 1 : l'arbre partagé

4.2.1.1 Un récepteur s'abonne à un groupe

Une station de travail exprime son désir de recevoir le trafic multicast associé à un groupe en utilisant le protocole MLD vu dans la section précédente. Le routeur PIM en charge du lien-local ou DR *(Designated Router)*, envoie un message `PIM (*,G) Join` vers le RP associé à ce groupe. On utilise la notation (*,G) car cela concerne n'importe quelle source pour le groupe G.

Ce message va être propagé de routeur en routeur vers le RP de ce groupe. A chaque routeur traversé un état associé à l'arbre multicast du groupe G est créé. Finalement le message `PIM (*,G) Join` va atteindre soit le RP, soit un routeur possédant déjà un état (*,G) associé au groupe. Quand plusieurs récepteurs adhèrent à un groupe, les messages `PIM Join` envoyés par les DR convergent vers le RP de ce groupe, ce qui forme l'arbre multicast pour ce groupe. Cet arbre est appelé RPT (*Rendez-vous Point Tree*) et il est qualifié d'arbre partagé puisqu'il sera utilisé pour atteindre tous les destinataires du groupe quelque soit l'émetteur du paquet multicast.

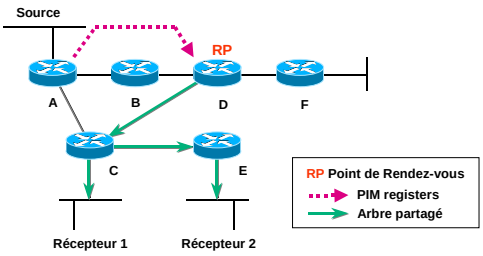


Arbre partagé

Des messages PIM d'adhésion sont envoyés périodiquement tant qu'au moins un destinataire est membre du groupe. Quand tous les destinataires situés sur un lien-local quittent un groupe, le DR peut envoyer un message PIM d'élagage (`PIM Prune`). Une durée limite de validité étant associée à chaque adhésion, si aucun message PIM ne parvient, l'adhésion sera résiliée.

4.2.1.2 Une source émet des paquets multicast

Dès qu'une station émet un paquet multicast vers un groupe, le DR de son lien-local encapsule ce paquet dans un datagramme unicast ayant pour adresse de destination le RP associé au groupe. Lorsque le RP reçoit ce datagramme, il décapsule le paquet multicast, et le propage sur le RPT associé au groupe. Le paquet est dupliqué aux nœuds qui forment de nouvelles branches, et donc parvient à l'ensemble des destinataires membres du groupe. Les datagrammes ayant encapsulé les paquets multicast sont appelés messages `PIM Register`.



Envoi des messages PIM Register

4.2.2 Etape 2 : l'acheminement spécifique

L'encapsulation dans les messages `PIM Register` est doublement inefficace :

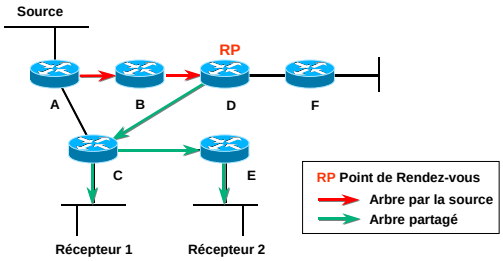
- L'encapsulation et la décapsulation sont des opérations qui sont coûteuses pour un routeur surtout s'il ne possède pas de matériel spécifique pour accélérer ces opérations.
- Le chemin de l'émetteur vers le RP puis à travers l'arbre RPT pour les récepteurs qui sont placés près de l'émetteur peut engendrer un large détour, produisant des délais et pouvant surcharger inutilement le réseau.

Le RP pourra choisir de basculer vers un acheminement natif (sans encapsulation) entre la source et le RP. Dans ce cas, lorsque le RP reçoit un message `PIM Register` contenant un paquet multicast provenant d'un émetteur `S` pour un groupe `G`, il peut envoyer un message `PIM (S,G) Join` vers S.

Ce message provoque dans les routeurs traversés la création d'un état multicast spécifique `(S,G)`. Ces états ne seront utilisés que pour transmettre les paquets multicast émis par `S` vers le groupe `G`. Finalement, ce message `PIM (S,G) Join` arrivera au DR associé à la source.

Dès que le `PIM (S,G) Join` arrive au DR de la source, ce dernier émet les données à la fois nativement vers le RP et en les encapsulant dans les messages `PIM Register`. Quand des paquets multicast natifs arrivent simultanément avec des paquets multicast encapsulés en provenance de la même source et pour le même groupe, le RP reçoit alors deux copies de chaque message. À partir de cet instant le RP doit détruire la copie des paquets multicast encapsulés, et il envoie un message `PIM Register-Stop` vers le DR de la source. Lorsque le DR reçoit un message `PIM Register-Stop`, il cesse d'encapsuler les paquets multicast dans des messages `PIM Register`.

A la fin de cette phase, le trafic émis par la source `S` pour le groupe `G` suit l'arbre centré sur la source jusqu'au RP, puis utilise l'arbre RPT (associé au groupe `G`) pour atteindre tous les destinataires du groupe `G`.



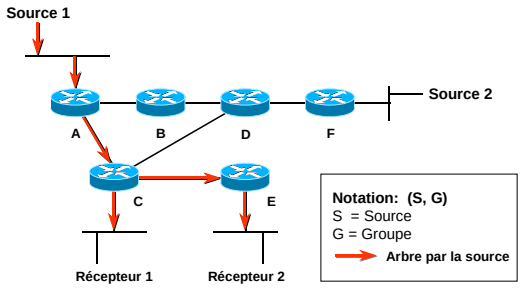
Le RP rejoint l'arbre centré sur la source

On notera que l'émetteur peut commencer à émettre avant ou après qu'un destinataire s'abonne à un groupe, et donc que cette deuxième phase peut être mise en place avant que l'arbre RPT soit construit.

4.2.3 Etape 3 : l'arbre des plus courts chemins

L'étape 2 supprime le surcoût introduit par l'encapsulation entre l'émetteur et le RP, cependant cela n'optimise pas complètement le chemin suivi par les paquets multicast. Pour de nombreux destinataires, le transit par le RP provoque un détour important si on compare ce chemin avec le chemin le plus court entre l'émetteur et chaque destinataire.

Une fois que le DR d'un récepteur reçoit les paquets émis par la source S vers le groupe G, il peut rejoindre l'arbre centré sur la source. On désignera cet arbre par SPT (*Shortest Path Tree*).



Arbre centré sur la source

Dans ce cas, le DR émet un message `PIM (S,G) Join` vers l'émetteur. Cela crée des états spécifiques `(S,G)` dans les routeurs rencontrés sur le chemin vers l'émetteur. Le message atteint le DR de la source ou un autre routeur ayant déjà l'état (S,G). Les paquets multicast dorénavant émis par l'émetteur `S` suivront les états `(S,G)`.

À partir de cet instant le DR du destinataire peut recevoir deux copies de chaque paquet multicast : une provenant du RP et ayant suivi l'arbre RPT associé, l'autre provenant directement de l'émetteur en ayant emprunté l'arbre centré sur la source (SPT). Dès que le premier paquet multicast est reçu en provenance de l'arbre centré sur la source, le DR du destinataire détruit les paquets qui arrivent via le RPT. Il envoie un message d'élagage `PIM (S,G) Prune` qui est propagé de routeur en routeur vers le RP. Dans chaque routeur rencontré ce message place un état indiquant que le trafic multicast `(S,G)` ne doit plus être propagé sur l'arbre partagé.

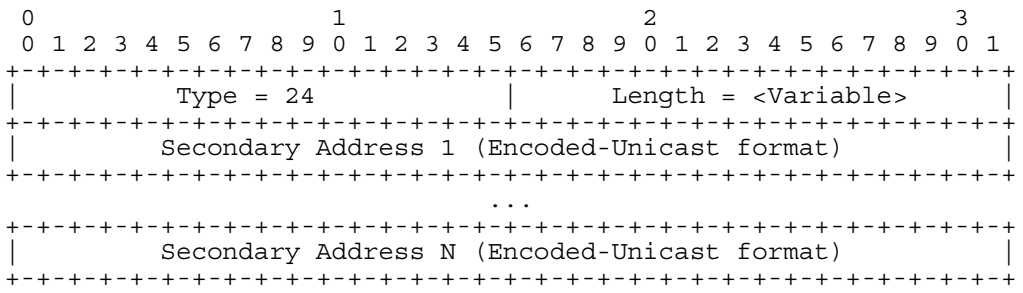
4.3 Le protocole PIM SSM - Source Specific Multicast

MLDv2 permet à un récepteur de spécifier le groupe auquel il veut s'abonner ainsi qu'un ensemble de sources pour ce groupe. La détermination d'un RP dans ce cas n'est pas nécessaire puisque les sources sont connues à l'avance par les destinataires. Prenons l'exemple d'une station qui s'abonne au groupe G en indiquant son intérêt pour les sources S1 et S2 uniquement. Le DR du récepteur peut envoyer directement des messages `PIM (S1,G) Join` vers S1 et `PIM (S2,G) Join` vers S2 et joindre ainsi les 2 arbres par la source associés. PIM SSM ne définit aucun nouveau message, c'est un sous-ensemble de PIM Sparse-Mode. Cependant, des adresses multicast dédiées pour PIM SSM doivent être utilisées. Ce sont des adresses dérivées du préfixe FF3X::/96

Il est à noter qu'il y a une forte préférence vers le modèle SSM pour le multicast. En effet, ce modèle permet de pallier tous les problèmes rencontrés dans le modèle et non résolus de manière complètement satisfaisante à ce jour : interdomaine multicast, allocation des adresses multicast, annonce des sessions...

4.4 Différences principales avec IPv4

- Les messages PIM sont échangés avec l'adresse de destination FF02::D (adresse de tous les routeurs PIM du lien)
- L'adresse source utilisée pour les messages PIM est l'adresse lien-local de l'interface d'où est émis le message. La conséquence directe est que cette adresse ne permet pas de réaliser le RPF check sur les messages PIM. Ainsi une option a été définie pour les messages `PIM Hello` afin de pouvoir spécifier toutes les adresses globales de l'interface. Ce sont ces adresses globales qui seront utilisées pour faire le RPF check sur le message. La structure de cette option qui est décrite ci-dessous est détaillée dans [23] :



Option "Address list" dans les messages PIM Hello

- La technologie embedded-RP est une différence majeure avec IPv4 en ce qui concerne le protocole PIM SM. Embedded-RP est décrit dans la section suivante sur l'interdomaine multicast (5.2.2)

5 Multicast IPv6 inter-domaine

5.1 Introduction

L'Internet est comme son nom l'indique une interconnexion de réseaux sous la direction d'entités administratives différentes (Autonomous system) qu'on appelle domaines. Il faut définir des mécanismes qui permettent à ces domaines de dialoguer, tout en préservant leur autonomie. Ces mécanismes sont déjà pleinement déployés pour l'unicast, mais sont encore en plein développement pour ce qui concerne le multicast.

Aujourd'hui, les protocoles multicast interdomaine pour IPv4 sont considérés comme non extensibles comme on le verra dans la section suivante sur MSDP. Ainsi pour IPv6, de nouveaux mécanismes ont été définis, prenant en compte l'existence de deux modèles de diffusion pour les applications : ASM et SSM. Alors que le modèle ASM interdomaine était implémenté par MSDP en IPv4, la solution qui semble privilégiée aujourd'hui est embedded-RP. Pour le modèle SSM, l'utilisation du protocole MLDv2 est indispensable afin d'informer le réseau des sources d'intérêt pour la construction des arbres. Dans cette partie, nous présentons deux solutions qui pourraient être déployées à grande échelle pour le multicast IPv6 : PIM-SM associé à embedded-RP pour l'ASM et PIM-SSM associé à MLDv2 pour SSM.

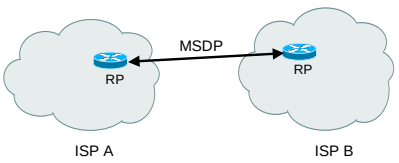
5.2 ASM

5.2.1 Rappel IPv4

En IPv4, un domaine PIM correspond à un ensemble de routeurs PIM gérés par une même entité. Tous les routeurs du domaine PIM sont configurés avec le même ensemble de points de rendez-vous qui appartiennent aussi à ce domaine. Pour permettre à des sources et des récepteurs répartis sur différents domaines de participer à une session multicast, un protocole été standardisé et déployé : il s'agit de MSDP *(Multicast Source Discovery Protocol)* [24].

Des peerings MSDP sont déployés entre les RP des différents domaines PIM comme indiqué sur la figure suivante. Ils permettent aux RP de s'échanger les informations quant aux sources actives dans les différents domaines. Chaque RP envoie à ses peers MSDP les sources qui émettent et les groupes destinataires.

Des filtres peuvent être appliqués sur les peerings pour permettre les annonces de certaines sources pour certains groupes uniquement.



Peering MSDP

Ce protocole classé expérimental ne permet pas une utilisation massive de la technologie multicast puisque les RP doivent s'échanger toutes les sources actives sur l'Internet. De plus, il s'agit d'un protocole compliqué, peu implémenté et difficile à administrer. Aussi, depuis plusieurs années, l'IETF recommande l'utilisation du modèle

SSM afin de rendre le modèle plus simple, même si le service rendu est différent avec SSM. Pour toutes ces raisons MSDP n'est pas défini pour IPv6 et il n'existe pas de protocole équivalent.

5.2.2 Embedded-RP

En l'absence de MSDP, la construction de l'arbre multicast avec PIM-SM nécessite que tous les routeurs PIM soient configurés avec le même ensemble de RP. Un groupe doit correspondre à un seul RP unique dans tout l'Internet puisque les RP ne peuvent pas s'échanger les informations sur les sources actives en IPv6.

Il est difficile d'imaginer un protocole permettant d'échanger les informations sur les RP existants et les adresses multicast qu'ils gèrent. Un tel protocole ne serait pas meilleur que MSDP.

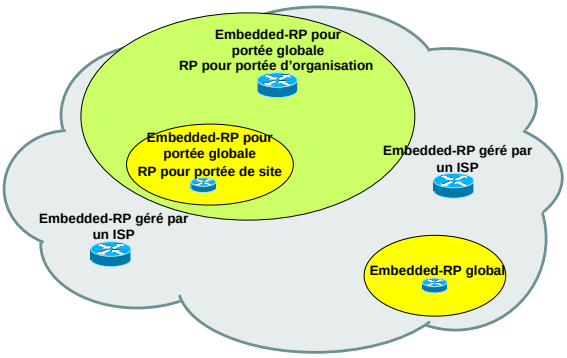
Une proposition simple a émergé : embarquer l'adresse du RP dans l'adresse multicast (ou embedded-RP). Ceci semble impossible car les 2 adresses ont une taille identique (de 128 bits). Cependant, en faisant certaines hypothèses sur l'identifiant d'interface du RP, il est possible de parvenir à cette solution. La méthode de construction d'une adresse embedded-RP est décrite dans la section adressage multicast de ce chapitre (voir section 2.1.3.3).

Modifications sur le protocole PIM SM

Embedded-RP [24] nécessite une modification de l'algorithme de correspondance entre les adresses multicast et les RP (ou *group-to-RP mapping)*. Pour un paquet à destination d'une adresse dérivée du préfixe FF70::/12, l'adresse du RP doit être retrouvée à l'aide des mécanismes décrits dans la section adressage multicast de ce chapitre (voir section 2.1.3.3). Embedded-RP doit être supporté sur tous les routeurs de l'arbre partagé, le RP et le DR des sources et des récepteurs. Le support d'embedded-RP sur l'arbre centré sur la source et sur les routeurs entre la source et le RP n'est pas nécessaire puisque ce sont des messages `PIM (S,G) prune/join` qui sont utilisés. Cependant, il est à noter qu'une implémentation sur tous les routeurs PIM SM du réseau simplifie l'utilisation et la gestion de la technologie embedded-RP.

Impact sur le modèle multicast

L'interdomaine multicast en IPv6 apparaît donc très différent de ce qui est réalisé en IPv4. Notamment la notion de domaine PIM disparaît et le terme interdomaine multicast ne correspond plus vraiment. L'Internet IPv6 multicast est un unique domaine PIM dans lesquels sont configurés des multiples points de rendez-vous.



Le modèle embedded-RP

Il est encore difficile à la date de rédaction de ce chapitre de déterminer si ce modèle va être adopté. Si des tests ont montré que la technologie embedded-RP fonctionnait, il reste néanmoins des questions sur les impacts causés par les différences avec le modèle connu et déployé à ce jour pour IPv4.

5.3 Problématique de déploiement de SSM sur plusieurs domaines

Le modèle SSM, implémenté par PIM-SSM constitue un sous-ensemble simplifié du modèle ASM. En effet, il a été défini historiquement pour répondre aux problèmes de l'interdomaine ASM n'ayant pas été résolus en IPv4. Par conséquent, les mécanismes permettant de réaliser l'interdomaine SSM en IPv6 sont très similaires à ceux utilisés en IPv4.

Au niveau du protocole PIM-SSM, il n'y a aucune disposition à prendre pour permettre à ce protocole de fonctionner entre plusieurs domaines. Les messages `PIM Join` sont acheminés de routeur en routeur entre les DR des récepteurs et les sources spécifiées par les récepteurs. Peu importe donc que les sources soient dans le même domaine ou non que les récepteurs.

Le problème du déploiement du protocole de construction d'arbre multicast PIM-SSM sur plusieurs domaines se situera donc au niveau du routage.

6 Déploiement du multicast

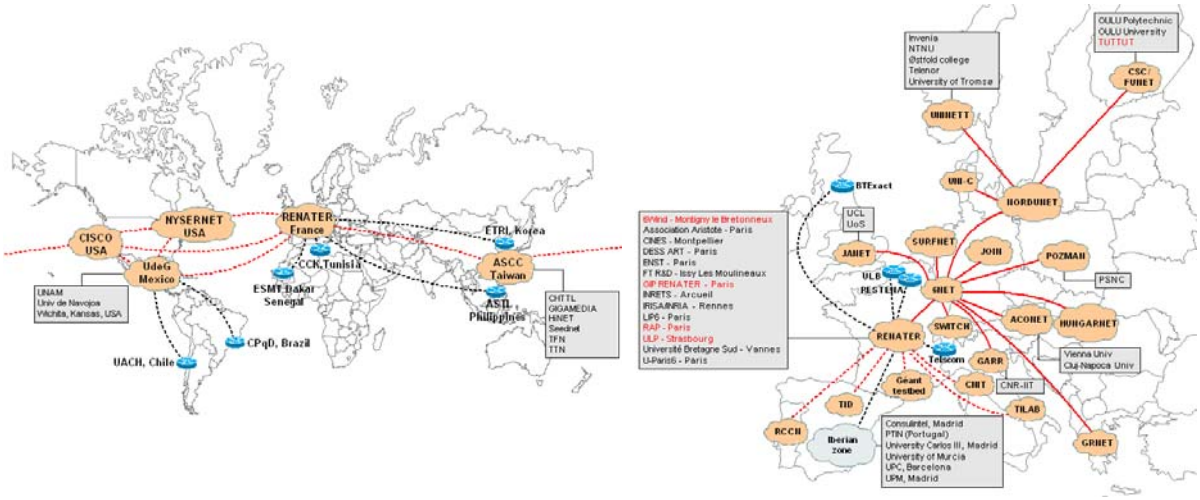
6.1 Le M6Bone

Qu'est-ce que le M6Bone ?

Le M6Bone est un réseau de test multicast IPv6. Le projet a débuté en Juillet 2001, sous l'impulsion en France de l'association Aristote, du G6 et de RENATER. Le but du projet est d'offrir une connectivité multicast IPv6 aux sites voulant expérimenter cette technologie. Le M6Bone permet aussi de valider des applications ou matériels relatifs au multicast IPv6.

Topologie du M6Bone

La figure ci-dessous présente des cartes de la topologie actuelle du M6Bone, avec les sites et réseaux connectés.



Carte mondiale et européenne du M6Bone

La majeure partie des liens du M6Bone est constituée de tunnels (IPv6 multicast dans IPv6 unicast ou alors IPv6 multicast dans IPv4). Si, au départ, des équipement différents devaient être utilisés pour le

multicast et l'unicast (absence de table de routage IPv6 multicast), l'implémentation de MBGP et de PIM sur certains routeurs commerciaux permet aujourd'hui de considérer un déploiement à grande échelle.

Le protocole utilisé dans le M6Bone est PIM sparse-mode. Le point de rendez-vous global est géré par RENATER. Des questions sur l'interdomaine multicast subsistent. MSDP ne verra pas le jour pour IPv6 et des solutions sont à l'étude. Le M6Bone permet de tester à grande échelle les solutions envisagées.

Services disponibles grâce au M6Bone

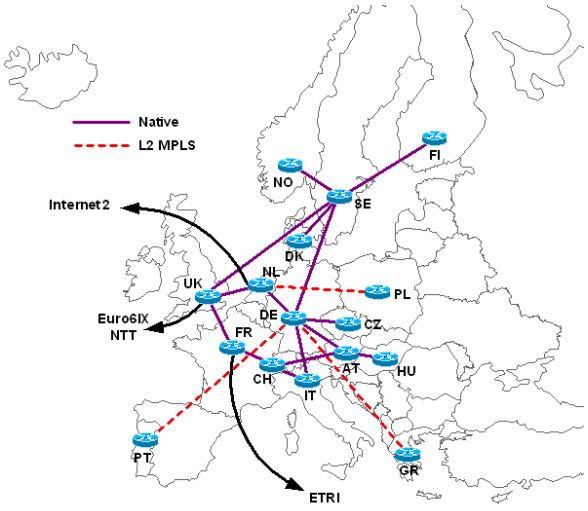
Presque tous les systèmes d'exploitation supportent IPv6 aujourd'hui et permettent d'utiliser des applications multicast. Les outils du MBone (vic, rat, sdr, nte, whiteboard...) supportent aujourd'hui IPv6 et il est possible de réaliser des visioconférences sur le M6Bone. Des outils comme freeamp permettent aussi de diffuser des stations de radio sur le M6Bone à haut débit. Des passerelles (ou réflecteurs) avec le réseau IPv4 multicast ont aussi été développées. Il est ainsi possible pour des personnes sur le M6Bone de joindre des sessions IPv4 et vice-versa. Des réflecteurs unicast/multicast permettent à des personnes ne disposant que d'une connectivité unicast de rejoindre des sessions multicast. Tous ces outils sont régulièrement utilisés pour la diffusion d'événements (conférences, causeries de Renater, séminaires Aristote, ...)

La communauté M6Bone

Une mailing-list libre (m6bone@ml.renater.fr), permet aujourd'hui à plus de 180 personnes d'échanger leurs connaissances sur le multicast IPv6 (routage, applications...) Un site web (http://www.m6bone.net) collecte les informations principales pour tout savoir sur les technologies multicast et les évolutions du réseau M6Bone. La configuration des différents équipements est également détaillée.

6.2 6NET

Le projet 6NET est un projet IST (Information Society Technology) du 5ème programme cadre de la Commission Européenne, regroupant principalement des partenaires du monde académique d'une quinzaine de pays européens. Il a entre autres permis le déploiement d'un backbone de test pan-européen IPv6 permettant d'interconnecter les réseaux de recherche partenaires. La figure ?? ci-dessous présente la topologie du réseau 6NET.



Topologie du réseau 6NET

Dès mars 2003, il a été possible de déployer le multicast IPv6 sur l'ensemble des routeurs du réseau 6NET et d'interconnecter nativement tous les réseaux nationaux de la recherche partenaires du projet. Les protocoles PIM SM, PIM SSM, MBGP (SAFI multicast IPv6) ont été déployés. BSR *(Bootstrap Router)* a été configuré sur les routeurs de cœur afin de permettre l'échange d'information quant aux points de rendez-vous configurés. De plus, tous les nœuds du réseau supportent Embedded-RP.

Le réseau 6NET a été interconnecté au M6Bone, ce qui a d'une part permis l'établissement de sessions multicast avec des entités non partenaires du projet ; et d'autre part cela a facilité la dissémination de l'expérience acquise à travers le projet vers tous les acteurs du M6Bone.

7 Applications multicast IPv6

Il existe un certain nombre d'applications qui fonctionnent déjà en multicast IPv6. La liste suivante n'est pas exhaustive mais montre l'étendue des services déjà disponibles.

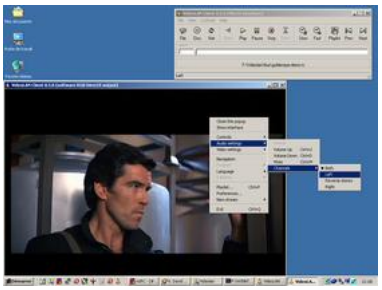
7.1 Diffusion de vidéo ou audio

DVTS

DVTS (Digital Video Transport System) permet la diffusion de flux vidéo de très bonne qualité à travers un réseau multicast IPv6. Des débits de plus de 20 Mbit/s sont utilisés pour transmettre la vidéo. Du matériel dédié est nécessaire comme des caméras numériques NTSC avec un port IEEE 1394.

Les informations sur DVTS sont disponibles en anglais sur le site http://www.dvts.jp/en/

VideoLAN



Cette application GNU est un projet de l'Ecole centrale de Paris. VLC (VideoLan Client) supporte un grand nombre de formats audio et vidéo (MPEG-1, MPEG2, MPEG-4, DivX, mp3..), et aussi les DVDs, VCDs et de nombreux protocoles de streaming. Cette application peut tout aussi bien être source ou récepteur de flux multicast IPv6.

Toutes les informations sur VideoLAN client sont disponibles sur le site web http://www.videolan.org

Windows Media Player 9

Les versions 9 et suivantes du client Windows Media Player permettent de recevoir des flux audio et vidéos multicast IPv6.

Freeamp

Freeamp est une application libre qui permet de recevoir des flux audio diffusés en streaming. Le format MP3 est le plus utilisé avec cette application qui correspond parfaitement à l'écoute de radio sur Internet. Le multicast IPv6 est supporté ce qui peut permettre aux sources de diffuser la radio avec une excellente qualité sans saturation des ressources du réseau.

7.2 Télé-enseignement

Isabel

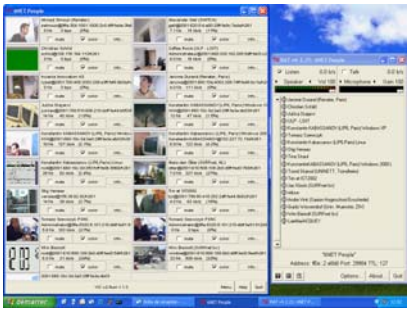
L'application Isabel est entièrement conçue pour le télé-enseignement. L'application permet la transmission de vidéo, d'audio et de transparents entre les multiples participants d'une session. Ceux-ci peuvent demander la parole pour poser des questions, et les intervenants peuvent clarifier certains points

en rajoutant des éléments sur les transparents à l'aide d'un stylo virtuel. Il est aussi possible d'utiliser des modes simplifiés pour des visioconférences simples.

Le site http://isabel.dit.upm.es donne des informations complémentaires sur l'application Isabel.

7.3 Visioconférence

VIC (VIdeoconference Tool)



VIC est l'application GNU traditionnellement la plus utilisée en multicast pour la visioconférence. Cet outil permet a plusieurs participants de s'échanger du trafic vidéo de manière simple, avec différents formats permettant une optimisation des débits disponibles pour la session. Il est possible avec VIC de créer des sessions avec un très grand nombre de participants comme il est possible de le voir sur l'image ci-contre. La page http://www-mice.cs.ucl.ac.uk/multimedia/software/vic/ donne des informations complémentaires sur VIC.

RAT (Robust Audio Tool)

RAT est l'équivalent de l'application VIC mais permet l'échange de l'audio. Pour assurer une visioconférence, il faut utiliser VIC et RAT en parallèle. D'avantages d'informations sont disponibles sur http://www-mice.cs.ucl.ac.uk/multimedia/software/rat/

7.4 Autres

WB (White Board)

WB est un tableau blanc partagé. Cette application peut permettre lors de visioconférences effectuées avec VIC et RAT de clarifier certains points à l'aide de schémas explicatifs, pouvant inclure du texte. Des informations complémentaires sont disponibles sur http://www-mice.cs.ucl.ac.uk/multimedia/software/wb/

NTE (Network Text Exchange)

NTE permet à des utilisateurs de communiquer en échangeant des messages. Chaque personne du groupe a une couleur personnelle ce qui permet de distinguer les messages écrits par chacun. La page http://www-mice.cs.ucl.ac.uk/multimedia/software/nte/ donne des renseignements complémentaires sur cet outil.

MAD-FLUTE

MAD-FLUTE permet le transfert de fichierq en utilisant la technologie multicast. MAD-FLUTE est une implémentation du protocole FLUTE qui est au moment de la rédaction de cet ouvrage un Internet-draft. A l'aide de cette application, il est possible de s'abonner à un groupe de diffusion de fichiers. MAD-FLUTE peut permettre la mise à jour de logiciels en téléchargeant les nouvelles versions régulièrement.

SDR

SDR permet de créer et d'annoncer des sessions multicast en utilisant les protocoles SDP et SAP. Les personnes désirant participer à une session n'ont qu'à retrouver l'annonce dans la liste créée par SDR. Les bonnes applications sont ensuite automatiquement lancées, avec les paramètres associés adresse multicast IPv6, numéro de port, codecs... La page http://www-mice.cs.ucl.ac.uk/multimedia/software/sdr/ donne des renseignements complémentaires sur cet outil.

8 Supervision du réseau multicast IPv6

La supervision d'un réseau multicast comprend plusieurs niveaux. Ils à toutes les composantes intervenants pour la mise en place du service multicast.

8.1 Supervision des liens

C'est le premier niveau de supervision du service multicast. Un lien cassé ou saturé ne permettra évidemment pas de rendre le service multicast. Les outils à mettre en oeuvre sont les mêmes que ceux utilisés pour l'unicast.

Il est important de comprendre que la supervision des liens est le premier niveau intervenant dans la gestion du multicast. Il ne sert à rien d'entrer dans une procédure longue et compliquée de debugging du service multicast quand l'origine est un lien cassé, qui peut être détecté très simplement.

8.2 Supervision du routage

Nous avons vu que le protocole PIM repose sur des informations de routage indépendantes (routes statiques multicast, MBGP, table de routage unicast...) De ce fait la supervision du routage est capitale. Une grande partie des problèmes rencontrés interviennent à ce niveau et peuvent être très simplement détectés.

Ici aussi les outils à mettre en oeuvre diffèrent peu des outils déjà mis en place pour la gestion du réseau unicast. Des scripts pourront détecter l'état des peerings MBGP, et relever des incohérence de routage. En effet, comme vu dans le chapitre sur la supervision du réseau, les MIBs ne sont pas encore implémentées pour la supervision des peerings BGP. Aussi, un outil comme AS-Path-tree peut être déployé pour relever des incohérence de routage sur l'ensemble du réseau. Si ce dernier outil ne permet pas d'effectuer une supervision en temps réel du réseau, et d'avoir des remontées d'alarme, il est très efficace pour étudier l'état global du routage et des politiques BGP mises en oeuvre.

8.3 Supervision de l'arbre multicast

Ce dernier niveau de supervision est le plus difficile à debugguer, c'est pour cette raison qu'il est plus simple de commencer par s'assurer qu'il n'y a déjà pas de problèmes rencontrés sur les liens, ou sur les informations de routage avant de se pencher sur la résolution des problèmes au niveau PIM.

La MIB PIM pour IPv6 n'est pas encore standardisée et n'est pas encore implémentée. Les informations quant à l'arbre multicast pour certains groupes et certaines sources particulières ne peuvent donc être récupérées sur les routeurs qu'en CLI, rendant les opérations compliquées. Des outils existants en IPv4 permettant de créer des cartes indiquant le trafic multicast sur l'ensemble du réseau pour certains groupes et sources ne sont donc pas disponibles en IPv6.

Néanmoins l'outil "Multicast Beacon" permet de donner des indications sur l'état du service multicast dans tout le réseau. Cet outil, basé sur perl ou java a un fonctionnement très simple : chaque Beacon installé émet du trafic multicast vers un groupe réservé pour la supervision du réseau, et reçoit aussi tous les paquets émis vers ce groupe. Chaque Beacon peut donc collecter localement les informations sur tous les Beacons desquels il reçoit du trafic. Ces informations sont envoyées vers un serveur central qui synthétise les résultats dans une matrice. Un exemple est représenté ci-après.

Loss [%]	S0	S1	S2	S3	S4	S5	S6
R0 zephyr.ipv6.unige.ch							
R1 UoS			0.0	0.0	2.0	0.0	
R2 merapi.switch.ch		0.0		0.0	0.0	0.0	
R3 UdeG-Mexico		0.0	0.0		0.0	0.0	
R4 tut.fi_telecom_lab		0.0	0.0	0.0		0.0	
R5 RENATER		0.0	0.0	0.0	0.0		
R6 beacon-test.geant.net		0.0		0.0	0.0	0.0	

Exemple de matrice Beacon

La matrice peut indiquer les pertes de paquets, gigue, délai entre les différents Beacons (tous émetteurs et récepteurs). Une simple lecture de la matrice montre où se trouvent les problèmes dans le réseau, et permet d'être proactif pour la résolution des problèmes. Pour avoir des informations sur le délai entre les différents Beacons, il convient de synchroniser en temps les clients Beacons, avec un protocole comme NTP par exemple.

L'utilisation de Beacon est simple et permet de gérer simplement les procédures de debugging : toute anomalie est détectée rapidement, et la résolution du problème peut être réalisée avant que le service soit nécessaire (conférence programmée...) Aussi, cela évite de demander aux personnes rencontrant des problèmes d'émettre et de recevoir du trafic sur certains groupes avec des applications multicast (par exemple des outils de visioconférence) et de demander à chacun leurs observations. Un temps précieux est ainsi gagné.

Il existe plusieurs manières pour déployer Beacon : l'opérateur peut fournir le serveur et demander à chaque site d'installer au moins un client. Ceci permettra de s'assurer que le service est rendu de bout en bout, et permettra de faciliter la résolution des problèmes comme expliqué précédemment. Néanmoins, ceci ne permettra pas de voir l'état du multicast dans le réseau (entre les différents routeurs gérés par l'opérateur). La deuxième manière consiste à installer des clients Beacons près des nœuds du réseau (ou du moins directement attachés aux nœuds de grande importance). Il est ensuite possible de générer des alarmes sur le serveur Beacon dès que 2 clients ne peuvent plus communiquer. Il est évident qu'installer Beacon des deux manières présentées permet d'être le plus efficace. En effet si l'état de la matrice interne est normal alors qu'un site ne peut émettre ou recevoir du trafic multicast, le problème se trouvera vraisemblablement sur le réseau du site.

9 Coexistence avec le multicast IPv4

L'objectif des passerelles IPv6/IPv4 est de permettre à des utilisateurs répartis dans les réseaux IPv6 et IPv4 multicast de participer à une même session multicast. Il existe des passerelles statiques aussi appelées réflecteurs qui permettent à des groupes IPv4 et IPv6 prédéfinis de communiquer. Une passerelle dynamique a aussi été conçue et implémentée permettant une complète interaction entre les réseaux multicast IPv6 et IPv4.

9.1 Passerelles statiques IPv6/IPv4 multicast (réflecteurs)

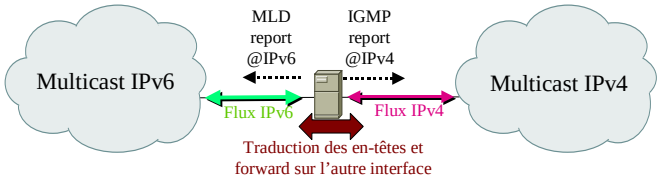
Usage des réflecteurs IPv4/IPv6 multicast

Ce type de passerelle permet d'assurer la correspondance entre un groupe IPv4 et un groupe IPv6 multicast. Sur chaque passerelle, les adresses des groupes IPv4 et IPv6 sont configurées statiquement. Ces réflecteurs peuvent donc être utilisés pour des sessions périodiquement utilisées sur des adresses IPv4 et IPv6 qui ne changent pas. La passerelle déployée sur RENATER traduit d'IPv4 à IPv6 (et vice versa) des sessios d'enseignement à distance, ainsi que des conférences organisées sur les thèmes des réseaux.

Mode de fonctionnement

Une passerelle s'abonne aux groupes IPv4 et IPv6 multicast qui sont rentrés en paramètre. Pour cela, elle envoie un message IGMP report sur l'interface vers le réseau multicast IPv4 et un message MLD report vers le réseau multicast IPv6. Il est bien évidemment possible d'utiliser une seule interface physique lorsque le multicast IPv4 et IPv6 sont configurés sur le même lien-local.

Une fois les messages IGMP et MLD report envoyés, la passerelle reçoit les flux multicast correspondants et n'a plus qu'à faire la traduction des en-têtes.



Réflecteur IPv6/IPv4 multicast

Dans le paquet traduit, l'adresse source est l'adresse (IPv4 ou IPv6) de la passerelle et l'information de la source est perdue. Cependant, l'expérience acquise montre que beaucoup d'applications utilisent la couche applicative pour identifier les différents participants de la session. L'utilisation d'une passerelle devient donc transparente pour les différents utilisateurs.

Un problème peut survenir lors du déploiement de réflecteurs si 2 passerelles sont configurées pour les mêmes groupes. Une boucle est alors créée dans le réseau ce qui sature les liens du réseau. Pour limiter la probabilité d'avoir ce genre de problème, il est possible de configurer une passerelle avec les ports UDP utilisés. La probabilité d'avoir 2 passerelles pour les mêmes adresses et numéros de ports devient quasi nulle. Il ne faut pas non plus voir ce problème comme une faille de sécurité car une personne malveillante peut émettre une forte quantité de trafic dans les groupes considérés pour saturer les liaisons et les équipements, et n'a pas besoin de passerelle pour cela. Un contrôle du débit maximal utilisable sur chaque groupe est une bonne solution pour éviter des conséquences trop importantes de ces attaques par déni de service.

9.2 Passerelles dynamiques

L'intérêt d'une telle passerelle est de pouvoir traduire n'importe quelle session IPv4 en IPv6 et vice-versa sans configuration préalable. Un utilisateur peut déployer le multicast IPv6 et arrêter le multicast IPv4 car il pourra avoir accès à tous les services IPv4 multicast grâce à la passerelle.

Le principe de base de cette passerelle est d'utiliser l'adresse IPv4 multicast comme identifiant de groupe de l'adresse multicast IPv6.

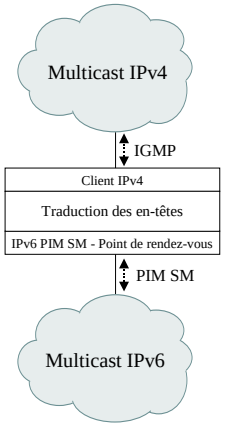
Préfixe multicast associé à la passerelle	Adresse IPv4 multicast

96 bits	32 bits

Structure des adresses utilisées par les passerelles dynamiques

La passerelle dynamique est :

- Un point de rendez-vous dans le réseau multicast IPv6 pour le préfixe qui lui est associé
- Un client terminal dans le réseau multicast IPv4



Principe de la passerelle dynamique IPv6/IPv4 multicast

Quand un client multicast IPv6 veut recevoir du flux multicast IPv4, il envoie un message MLD report pour l'adresse IPv6 multicast dérivée du préfixe multicast associé à la passerelle et de l'adresse IPv4 multicast. Le DR du lien envoie alors un message PIM join pour cette adresse qui sera propagé jusqu'à la passerelle puisqu'elle est configurée comme RP pour le préfixe utilisé. La passerelle retrouvera l'adresse IPv4 dans le paquet et enverra un message IGMP report pour cette adresse IPv4. Le trafic IPv4 multicast atteindra ensuite la passerelle qui traduira les paquets comme cela est expliqué dans la partie précédente sur les passerelles statiques.

Lorsqu'une source émet du trafic multicast IPv6 avec des adresses associées à la passerelle, le DR sur le lien-local encapsule ce trafic vers la passerelle puisqu'elle est configurée comme point de rendez-vous pour le préfixe utilisé. La passerelle va traduire ensuite les en-têtes de la même manière qu'une passerelle statique en utilisant l'adresse IPv4 embarquée dans l'adresse IPv6. Les paquets traduits sont envoyés sur l'interface multicast IPv4. Comme expliqué dans la section consacrée à PIM, les paquets pourront être transmis de maniére native entre la source et le RP afin d'éviter le coût lié à l'encapsulation dans les messages PIM register.

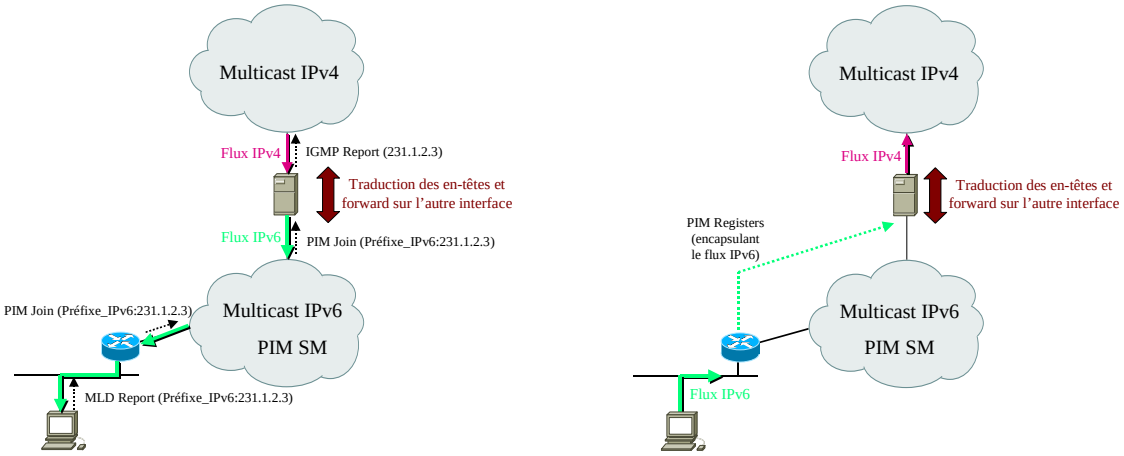


Schéma explicatif de la passerelle dynamique

Pour rendre l'usage de ces passerelles encore plus facile, il est possible de traduire d'IPv4 à IPv6 les annonces SAP de toutes les sessions IPv4. Ainsi une session IPv4 annoncée par SAP sera visible par les clients multicast IPv6, avec des adresses automatiquement traduites pour permettre d'utiliser la passerelle. Il est ainsi possible pour un site de se passer entièrement du multicast IPv4, en gardant la certitude d'avoir accès à tous les services associés.

10 Etude pratique du déploiement du multicast IPv6

L'objectif est ici de détailler les différentes étapes de la mise en place d'un service multicast IPv6 dans un réseau existant. Pour avoir d'avantage de détails sur la configuration des équipements, le lecteur pourra lire le chapitre détaillant la configuration des routeurs.

10.1 Choisir le service (ou les applications) multicast IPv6 à déployer

La toute première chose est de comprendre quelles sont les caractéristiques principales du service multicast que l'on souhaite mettre en place dans le réseau. Celles-ci dépendent principalement des applications désirées.

Comme expliqué dans une section précédente de ce chapitre, il existe 2 modèles pour le multicast IPv6: le modèle ASM (Any Source Multicast) et le modèle SSM (Source Specific Multicast). Pour des applications de vidéoconférence multi-utilisateurs, le modèle ASM doit être déployé. Si les applications désirées sont de type diffusion de contenu d'une source connue vers un ensemble de récepteurs (par exemple la radio ou la télévision sur internet), le modèle choisi sera SSM. Il est bien évidemment possible de déployer les 2 modèles multicast dans un même réseau.

Le protocole PIM SM/SSM permet l'implémentation de ces 2 modèles :

- Utilisé avec MLDv1 ou MLDv2 et configuré avec un certain nombre de points de rendez-vous, PIM permet de fonctionner en mode ASM.
- Utilisé avec MLDv2, PIM peut fonctionner en mode SSM, et dans ce cas il n'est pas nécessaire de configurer des points de rendez-vous. Les clients grâce au protocole MLDv2 peuvent spécifier les sources multicast et la configuration des RP devient inutile.

10.2 Choisir la topologie du réseau

La topologie du réseau multicast dépend du support des protocoles multicast choisis dans les équipements du réseau. Il est plus simple de déployer le multicast IPv6 lorsque ce service est supporté par tous les équipements du réseau. Dans ce cas, les topologies unicast et multicast sont congruentes ce qui simplifie le design et l'administration du réseau. Lorsque des routeurs ne sont pas capables de gérer le multicast IPv6, plusieurs solutions peuvent être envisagées : tunnels multicast, VLANs ethernet dédiés, changement des équipements réseaux, mise à jour du logiciel des routeurs, changement de la topologie du réseau, PVC ATM ou LSP MPLS dédiés...

10.2.1 Topologies unicast et multicast congruentes

Dans le cas où tous les équipements du réseau supportent les protocoles multicast nécessaires (PIM SM/SSM, MLDv2...) la topologie multicast pourra être la même que la topologie unicast déployée : on parle alors de topologies unicast et multicast congruentes. Cela signifie que les routes entre les sources et récepteurs multicast sont identiques aux routes unicast.

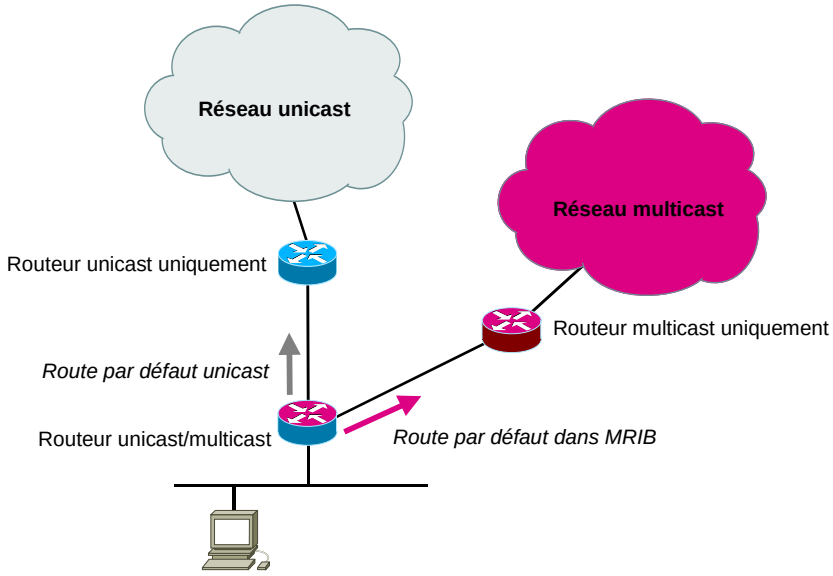
Le protocole PIM pourra alors utiliser la table de routage unicast. L'implantation de MBGP (SAFI IPv6 multicast) ou de routes statiques multicast n'est alors pas nécessaire.

10.2.2 Topologies unicast et multicast non congruentes

Dans ce cas, le déploiement et l'administration du service multicast deviennent plus complexes. Les cas de topologies non-congruentes sont multiples : déploiement d'équipements spécifiques pour le multicast, routage du multicast sur des liens dédiés...

L'administrateur du réseau devra alors utiliser une base d'information de routage différente de la base unicast. Le protocole PIM devra reposer sur des informations de routage correspondant à la topologie multicast déployée. Des routes statiques multicast ou alors MBGP IPv6 multicast SAFI doivent alors être considérés.

La figure suivante montre le cas d'un site de petite taille (un seul routeur et un seul lien local) connecté à l'Internet v6 unicast et au multicast (M6Bone) par 2 liens différents. La connexion vers le M6Bone peut être par exemple de type tunnel. Si le site souhaite déployer du routage statique, une route par défaut devra être configurée vers le réseau unicast. Pour le service multicast, le protocole PIM utilise des informations de routage pour connaître la topologie multicast et envoyer les messages protocolaires sur les bonnes interfaces. Si aucune autre information n'est ajoutée, le routeur du site va envoyer tous les messages PIM Join/Prune vers le réseau unicast car la route par défaut pointe dans cette direction. Il faut donc indiquer à ce routeur la topologie multicast. Ceci est fait en ajoutant une route statique multicast qui ne sera utilisée que pour le multicast. Cette route va peupler la table de routage multicast du routeur ou MRIB (Multicast Routing Information Base). Cette route n'est pas utilisée pour le forwarding des paquets unicast mais sera utilisée pour la construction de l'arbre et pour le routage des paquets multicast. En effet, le RPF check sera fait en utilisant les informations de routage contenues dans la MRIB. MBGP peut aussi être utilisé pour peupler la MRIB. La sub-address Family IPv6 multicast permet d'échanger les informations de routage pour le multicast.



10.3 Comment déployer un service fiable et efficace ?

Afin de déployer le protocole de routage multicast, il est important de déterminer quelle base d'information de routage unicast doit être utilisée par les routeurs multicast pour la vérification RPF. Si le service multicast n'est déployé qu'à l'intérieur d'un seul domaine (intra-domaine), la table de routage unicast sous-jacente peut alors être utilisée dans la mesure où les topologies unicast et multicast sont congruentes. Si toutefois le service multicast doit couvrir plusieurs domaines (inter-domaine) les routeurs multicast, tels que les routeurs PIM-SM, peuvent s'appuyer sur une autre base d'information de routage qui sera alors utilisée uniquement pour la vérification RPF du routage multicast (on non pour la routage unicast). MBGP (ou BGP IPv6 multicast SAFI, voir RFC2858) peut être utilisé pour générer de telles bases.

Lors du déploiement du routage multicast, il est aussi important de s'assurer de l'efficacité du routage mis en place. Par exemple dans le cas de PIM-SM, le placement des « Rendez-vous Points » (RP) peut avoir un impact important sur la charge du réseau. Afin d'éviter les surcharges sur un nœud particulier, il est possible de configurer plusieurs RP, chacun gérant uniquement un sous-ensemble d'adresses multicast.

La robustesse du service mis en place doit aussi être considérée. Par exemple dans le cas de PIM-SM, il est aussi possible de configurer plusieurs RP pour servir le même ensemble d'adresses multicast. Si le RP principal devient indisponible, un RP secondaire prend alors le relais afin de maintenir le service.

10.4 Quels services supplémentaires supporter ?

Le routage multicast à lui seul ne suffit pas pour assurer un service multicast complet. Certains services supplémentaires comme l'annonce des sessions ou encore la gestion des adresses multicast peuvent être nécessaire.

L'annonce des sessions multicast peut se faire par simple publication sur page web, ou avec le protocole SAP et l'application SDR décrits plus haut dans ce chapitre.

La gestion des adresses multicast IPv6 dans un réseau est comme nous l'avons vu plus haut un problème complexe et il n'existe pas encore de solutions aujourd'hui pour permettre d'allouer des adresses multicast IPv6 pour les applications.

10.5 Comment interconnecter son réseau à l'Internet multicast IPv6 ?

Plusieurs réseaux multicast IPv6 régionaux ou internationaux, tel le M6Bone, sont d'ores et déjà déployés. Toute organisation souhaitant expérimenter le multicast IPv6 à grande échelle peut s'y rattacher, en attendant l'extension du service multicast IPv6 à tous les ISP.

Comme discuté précédemment, le protocole PIM-SM semble actuellement la solution la plus appropriée pour supporter le déploiement massif du multicast IPv6. En effet, ce seul protocole, associé à l'utilisation d'adresses multicast IPv6 « embedded RP », offre une solution globale permettant de s'affranchir du modèle traditionnel « intra-domaine / inter-domaine » du multicast IPv4.

Références

[1] Linux, http://www.linux.org/

[2] FreeBSD, http://www.freebsd.org/

[3] S. Deering, W. Fenner, and B. Haberman, “Multicast Listener Discovery (MLD) for IPv6”, RFC 2710, October 1999.

[4] D. Estrin et al., “Protocol Independent Multicast-Sparse Mode (PIM-SM) Protocol Specification”, RFC 2362, June 1998.

[5] Bill Fenner, Mark Handley, Hugh Holbrook, Isidor Kouvelas, “Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)”, Internet Draft, daft-ietf-pim-sm-v2-new-05.txt, 1 March 2002.

[6] VideoLAN VLC media player, http://www.videolan.org/vlc/.

[7] Video Conferencing Tool, VIC, http://www-mice.cs.ucl.ac.uk/multimedia/software/vic/

[8] B. Carpenter and K. Moore "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.

[9] S. Thomson and T. Narten, “IPv6 Stateless Address Autoconfiguration”, RFC 2462, December 1998.

[10] FreeBSD ports, pim6sd, http://www.freshports.org/net/pim6sd/

[11] M6Bone, http://www.m6bone.net/

[12] Ethereal, http://www.ethereal.com/

[13] Narten T., Nordmark E., and W. Simpson, “Neighbor Discovery for IP version 6 (IPv6)”, RFC 2461, December 1998.

[14] B. Haberman, “Dynamic Allocation Guidelines for IPv6 Multicast Addresses”, RFC 3307, June 2002.

[15] B. Haberman, and D. Thaler, “Unicast-Prefix-based IPv6 Multicast Addresses”, RFC 3306, Auguest 2002.

[16] P. Savola, “IPv6 Multicast Deployment Issues”, IETF Internet Draft, draft-savola-v6ops-multicast-issues-03.txt, February 2004.

[17] David B. Johnson, C. Perkins, "Mobility support in IPv6", IETF Internet Draft, draft-ietf-mobileip-ipv6-24.txt, June 2003.

[18] Bill Fenner, Haixiang He, Brian Haberman, Hal Sandick, “IGMP/MLD-based Multicast Forwarding (“IGMP/MLD Proxying”)”, Internet Draft, draft-ietf-magma-igmp-proxy-04.txt, September 2003.

[19] B. Haberman, H. Sandick, G. Kump, “Protocol Independent Multicast Routing in the Internet Protocol Version 6 (IPv6)”, Internet Draft, draft-ietf-pim-ipv6-03.txt, March 2000.

[20] Mark Handley, Van Jacobson, and Colin Perkins, “SDP: Session Description Protocol”, draft-ietf-mmusic-sdp-new-13.txt, 22 May 2003.

[21] Mark Handley, Colin Perkins, and E. Whelan, “Session Announcement Protocol”, IETF Standards Track, RFC 2974, October 2000.

[22] S. Hanna, B. Patel, and M. Shah, “Multicast Address Dynamic Client Allocation Protocol (MADCAP)”, IETF Standards Track, RFC 2730, December 1999

[23] Bill Fenner, Mark Handley, Hugh Holbrook, Isidor Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", IETF Internet Draft, draft-ietf-pim-sm-v2-new-09.txt ,16 February 2004

[24] B. Fenner, D. Meyer, "Multicast Source Discovery Protocol (MSDP)", RFC3618, October 2003.

[25] Embedded-RP

[26] R. Hinden, S. Deering, "IPv6 Multicast Address Assignments", RFC 2375 , July 1998.

[27] R. Hinden, S. Deering, "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC 3513, April 2003

[28] R. Vida, L.Costa, "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, June 2004

[29] Bill Fenner, Haixiang He, Brian Haberman, Hal Sandick, "IGMP/MLD-based Multicast Forwarding ('IGMP/MLD Proxying')", IETF Internet Draft, draft-ietf-magma-igmp-proxy-06.txt, Avril 2004.

[30] B. Fenner, B. Haberman, H. Holbrouk, I.Kouvelas, "Multicast Source Notification of Interest Protocol (MSNIP)" , IETF Internet Draft, draft-ietf-magma-msnip-05.txt, 10 mars 2004.

Glossaire

..

Index

If necessary.